III General Relativity

Ishan Nath, Michaelmas 2024

Based on Lectures by Prof. Claude Warnick

February 9, 2025

Contents

0   Introduction

Office hours: 8:40AM MWF, in MR2. Normal room E1.14. Will follow roughly Reall's course.

General relativity is our best theory of gravitation on the largest scales. It is:

- Classical : No quantum effects.

- Geometrical: Space and time are combined in a curved spacetime.

- Dynamical: In contrast to Newton's theory of gravity, Einstein's gravitational field has its own non-trivial dynamics.

1   Differentiable Manifolds

The basic object of study in differential geometry is the (differentiable) manifold. This is an object which 'locally looks like $\mathbb{R}^n$', and has enough structure to let us do calculus.

Definition 1.1. A *differentiable manifold* of dimension n is a set M, together with a collection of coordinate charts (O_α, ϕ_α), where:

- $O_\alpha \subseteq M$ are subsets of M such that

$$\bigcup_\alpha O_\alpha = M.$$

- ϕ_α is a bijective map from O_α to U_α, an open subset of $\mathbb{R}^n$.

- If $O_\alpha \cap O_\beta \neq \emptyset$, then

$$\phi_\beta \circ \phi_\alpha^{-1}$$

is a smooth map from $\phi_\alpha(O_\alpha \cap O_\beta) \subseteq U_\alpha$ to $\phi_\beta(O_\alpha \cap O_\beta) \subseteq U_\beta$.
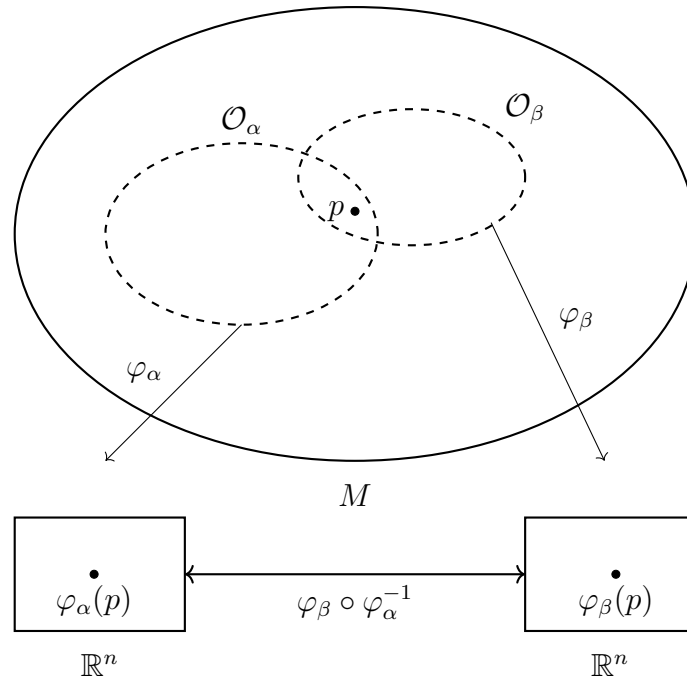
Remark.

- We could replace smooth with finite differentiability (e.g. k-times differentiable).

- The charts define a topology on M: $U \subseteq M$ is open if and only if $\phi_\alpha(U \cap O_\alpha)$ is open in $\mathbb{R}^n$ for all α. Every open subset of M is itself a manifold, by restricting the charts to U.

The collection $\{(O_\alpha, \phi_\alpha)\}$ is called an *atlas*. Two atlases are *compatible* if their union is an atlas.

An atlas A is *maximal* if there exists no atlas B which is compatible with A, and strictly larger than A. Every atlas is contained in a maximal atlas (by taking the union of all compatible atlases). Hence we can assume without loss of generality that we work with a maximal atlas.

Example 1.1.

1. If $U \subseteq \mathbb{R}^n$ is open, we can take $O = U$, and $\phi : U \to \mathbb{R}^n$ to be the identity on U. Then $\{(O, \phi)\}$ is an atlas.

2. Take S^1. If $p \in S^1 \setminus \{(-1, 0)\} = O_1$, there is a unique $\theta_1 \in (-\pi, \pi)$ such that
$$p = (\cos\theta_1, \sin\theta_1).$$

   If $p \in S^1 \setminus \{(1, 0)\} = O_2$, there is a unique $\theta_2 \in (0, 2\pi)$ such that
$$p = (\cos\theta_2, \sin\theta_2).$$

   These maps from $(-\pi, \pi)$ and $(0, 2\pi)$ to O_1, O_2 give ϕ_1^{-1}, ϕ_2^{-1} respectively. Note that $\phi_1(O_1 \cap O_2) = (-\pi, 0) \cup (0, \pi)$, and the transition function is

$$\phi_2 \circ \phi_1^{-1}(\theta) = \begin{cases} \theta & \theta \in (0, \pi), \\ \theta + 2\pi & \theta \in (-\pi, 0). \end{cases}$$

   This is smooth where defined, and similarly for $\phi_2 \circ \phi_1^{-1}$. Hence S^1 is a manifold.

3. More generally, we can consider S^n, and can define charts by stereographic projections. If $\{\mathbf{E}_1, \ldots, \mathbf{E}_{n+1}\}$ is the standard basis for $\mathbb{R}^{n+1}$, and $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the standard basis for $\mathbb{R}^n$, write

$$\mathbf{P} = P^1\mathbf{E}_1 + \cdots + P^{n+1}\mathbf{E}_{n+1}.$$

   Set $O_1 = S^n \setminus \{\mathbf{E}_{n+1}\}$, and write

$$\phi_1(\mathbf{P}) = \frac{1}{1 - P^{n+1}}(P^1\mathbf{e}_1 + \cdots + P^n\mathbf{e}_n).$$

In a similar way we may define O_2, ϕ_2, for $-\mathbf{E}_{n+1}$. The transition map is then
$$\phi_2 \circ \phi_1^{-1}(\mathbf{x}) = \frac{\mathbf{x}}{|\mathbf{x}|^2},$$
which is smooth on $\mathbb{R}^n \setminus \{0\} = \phi_1(O_1 \cap O_2)$.

"Nice" surfaces in $\mathbb{R}^n$ are manifolds with no cusps, cornered or self-intersections, for example $S^n \subseteq \mathbb{R}^{n+1}$.

1.1   Smooth Functions on Manifolds

Suppose M, N are manifolds of dimension n, n' respectively. Let $f : M \to N$.

Then let $p \in M$, and pick charts (O_α, ϕ_α) for M, and (O'_β, ϕ'_β) for N, with $p \in O_\alpha$ and $f(p) \in O_\beta$. Then,
$$\phi'_\beta \circ f \circ \phi_\alpha^{-1}$$
maps a neighbourhood of $\phi_\alpha(p)$ in $U_\alpha \subseteq \mathbb{R}^n$, to $U'_\beta \in \mathbb{R}^{n'}$. If this function is smooth for all possible choices of this chart, we say that $f : M \to N$ is smooth.
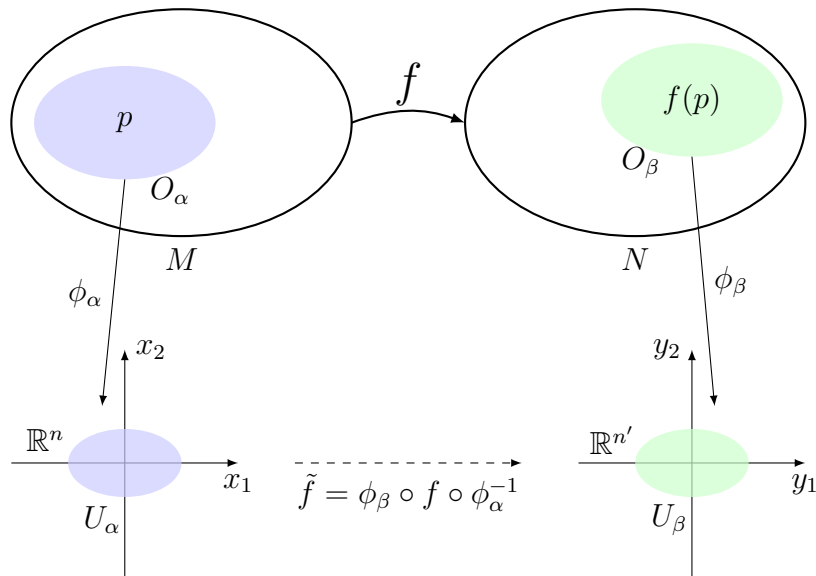
Remark.

- A smooth map $\psi : M \to N$ which has a smooth inverse is called a *diffeomorphism*.

- If $N = \mathbb{R}$ or $\mathbb{C}$, we sometimes call f a *scalar field*.

- If $M = I \subseteq \mathbb{R}$, an open interval, then $f : I \to N$ is a *smooth curve* in N.

- If f is smooth in one atlas, then it is smooth in all compatible atlases.

Example 1.2.

1. Recall S^1. Let $f(x, y) = x$, $f : S^1 \to \mathbb{R}$. Using our previous charts,

$$f \circ \phi_1^{-1} : (-\pi, \pi) \to \mathbb{R}$$
$$\theta_1 \mapsto \cos\theta_1.$$

   Similarly,

$$f \circ \phi_2^{-1} : (0, 2\pi) \to \mathbb{R}$$
$$\theta_2 \mapsto \cos\theta_2.$$

   Hence f is smooth.

2. If (O, ϕ) is a coordinate chart on M, write

$$\phi(p) = (x^1(p), x^2(p), \dots, x^n(p)),$$

   for $p \in O$. Then $x^i(p)$ defines a map from O to $\mathbb{R}$. This is smooth for each $i = 1, \dots, n$. If (O', ϕ') is another overlapping coordinate chart, then $x^i \circ (\phi')^{-1}$ is the i'th component of $\phi \circ (\phi')^{-1}$, hence smooth.

3. We can define a smooth function chart-by-chart. For simplicity, let $N = \mathbb{R}$, and let $\{(O_\alpha, \phi_\alpha)\}$ be an atlas on M. Define smooth functions

$$F_\alpha : U_\alpha \to \mathbb{R},$$

   and suppose that $F_\alpha \circ \phi_\alpha = F_\beta \circ \phi_\beta$ on $O_\alpha \cap O_\beta$, for all α, β.

   Then for $p \in M$, we can define

$$f(p) = F_\alpha \circ \phi_\alpha(p),$$

   for any chart (O_α, ϕ_α) with $p \in O_\alpha$. Now f is smooth, as

$$f \circ \phi_p^{-1} = F_\alpha \circ \phi_\alpha \circ \phi_\beta^{-1}.$$

   In practice, we often don't distinguish between f and its coordinate chart representations F_α.

1.2   Curves and Vectors

For a surface in $\mathbb{R}^3$, we have a notion of the 'tangent space' at a point p, which consists of all vectors tangent to the surface at that point.

The tangent spaces are vector spaces (copies of $\mathbb{R}^2$). Different points have different tangent spaces.

In order to define the tangent space for a manifold, we first consider tangent vectors of a curve.

Recall that $\lambda : I \to M$ a smooth map, is a smooth curve in M.

If $\lambda(t)$ is a smooth curve in $\mathbb{R}^n$, and $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth function, the chain rule gives

$$\frac{\mathrm{d}}{\mathrm{d}t}[f(\lambda(t))] = X(t) \cdot \nabla f(\lambda(t)),$$

where $X(t) = \mathrm{d}\lambda/\mathrm{d}t(t)$ is the tangent vector to λ at t.

Definition 1.2. Let $\lambda : I \to M$ be a smooth curve with $\lambda(0) = p$. The *tangent vector* to λ at p is the linear map X_p from the space of smooth functions $f : M \to \mathbb{R}$, given by

$$X_p(f) = \frac{\mathrm{d}}{\mathrm{d}t} f(\lambda(t)) \Big|_{t=0}.$$

We observe that:

(i) X_p is linear: $X_p(f + ag) = X_p(f) + aX_p(g)$, for f, g smooth, and $a \in \mathbb{R}$.

(ii) X_p satisfies the Leibniz rule:

$$X_p(fg) = X_p(f)g(p) + f(p)X_p(g).$$

If (O, ϕ) is a chart and $p \in O$, write

$$\phi(p) = (x^1(p), \ldots, x^n(p)).$$

Let $F = f \circ \phi^{-1}$, and $x^i(t) = x^i(\lambda(t))$. Now $F : \mathbb{R}^n \to \mathbb{R}$, and $x : \mathbb{R} \to \mathbb{R}^n$. Then,

$$f \circ \lambda(t) = f \circ \phi^{-1} \circ \phi \circ \lambda(t) = F \circ x(t),$$

and by applying the regular chain rule for functions $\mathbb{R}^m \to \mathbb{R}^n$,

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\lambda(t)) \Big|_{t=0} = \frac{\partial F}{\partial x^\mu}(x) \cdot \frac{\mathrm{d}x^\mu}{\mathrm{d}t} \Big|_{t=0}.$$

Proposition 1.1. *The set of tangent vectors to curves at p forms a vector space, T_pM, of dimension $n = \dim M$. We call T_pM the* tangent space *to M at p.*

Proof: Given X_p, Y_p tangent vectors, we need to show that $\alpha X_p + \beta Y_p$ is a tangent vector for $\alpha, \beta \in \mathbb{R}$.

Let λ, κ be smooth curves with $\lambda(0) = \kappa(0) = p$, and whose tangent vectors at p are X_p, Y_p, respectively.

Let (O, ϕ) be a chart with $p \in O$ and $\phi(p) = 0$, and define

$$\gamma(t) = \phi^{-1}(\alpha\phi(\lambda(t)) + \beta\phi(\kappa(t))).$$

This exists, as this is just the sum of elements in $\mathbb{R}^n$.

Now $\gamma(0) = \phi^{-1}(0) = p$. If Z_p is the tangent to v at p, then

$$
\begin{aligned}
Z_p(f) &= \frac{\mathrm{d}}{\mathrm{d}t} f(v(t)) \Big|_{t=0} = \frac{\partial F}{\partial x^\mu}\Big|_0 \frac{\mathrm{d}}{\mathrm{d}t}[\alpha x^\mu(\lambda(t)) + \beta x^\mu(\kappa(t))]\Big|_{t=0} \\
&= \alpha \frac{\partial F}{\partial x^\mu}\Big|_0 \frac{\mathrm{d}}{\mathrm{d}t} x^\mu(\lambda(t))\Big|_{t=0} + \beta \frac{\partial F}{\partial x^\mu}\Big|_0 \frac{\mathrm{d}}{\mathrm{d}t} x^\mu(\kappa(t))\Big|_{t=0} \\
&= \alpha X_p(f) + \beta Y_p(f).
\end{aligned}
$$

Thus $T_p M$ is a vector space.

To see that $T_p M$ is n-dimensional, consider the curves

$$\lambda_\mu(t) = \phi^{-1}\underbrace{(0, \ldots, 0, t, 0, \ldots, 0)}_{\mu\text{'th component}}.$$

We denote the tangent vector to λ_μ at p by $(\partial/\partial x^\mu)_p$. To see why, note that

$$\left(\frac{\partial}{\partial x^\mu}\right)_p f = \frac{\partial F}{\partial x^\mu}\Big|_{\phi(p)=0}.$$

These vectors are linearly independent. Otherwise there exists α^μ not all zero such that

$$\alpha^\mu \left(\frac{\partial}{\partial x^\mu}\right)_p = 0 \implies \alpha^\mu \frac{\partial F}{\partial x^\mu}\Big|_0 = 0,$$

for all F. Setting $F = x^\nu$ gives $\alpha^\nu = 0$.

Further, these vectors form a basis for $T_p M$, since if λ is any curve with tangent X_p at p, then

$$X_p(f) = \frac{\partial F}{\partial x^\mu}\Big|_0 \frac{\mathrm{d}}{\mathrm{d}t} x^\mu(\lambda(t))\Big|_{t=0} = X^\mu \left(\frac{\partial}{\partial x^\mu}\right) f.$$

Here, we set
$$X^\mu = \frac{\mathrm{d}}{\mathrm{d}t} x^\mu(\lambda(t)) \Big|_{t=0}.$$
These are the components of X_p with respect to the basis.

$$U \subset \mathbb{R}^2$$
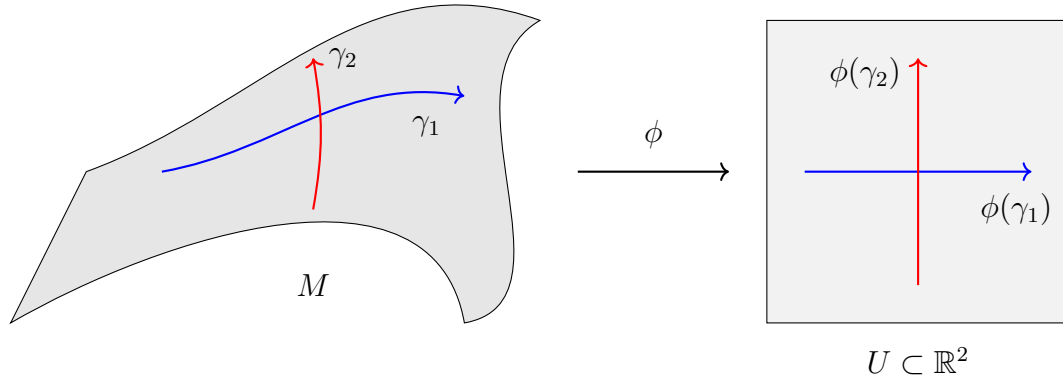
Notice that the basis $\{(\partial/\partial x^\mu)_p\}$ depends on the coordinate chart ϕ. Suppose we choose another chart (O', ϕ'), again centred at p. Write $\phi' = (x'^1, \ldots, x'^n)$.

Then if $F = f \circ (\phi')^{-1}$, then

$$F(x) = f \circ \phi^{-1}(x) = f \circ (\phi')^{-1} \circ \phi' \circ \phi^{-1}(x)$$
$$= F'(x'(x)).$$

Therefore,

$$\left(\frac{\partial}{\partial x^\mu}\right)_p f = \frac{\partial F}{\partial x^\mu}\Big|_{\phi(p)} = \left(\frac{\partial x'^\nu}{\partial x^\mu}\right)_{\phi(p)} \left(\frac{\partial F'}{\partial x'^\nu}\right)_{\phi'(p)} = \left(\frac{\partial x'^\nu}{\partial x^\mu}\right)_{\phi(p)} \cdot \left(\frac{\partial}{\partial x'^\nu}\right)_p f.$$

We deduce that

$$\left(\frac{\partial}{\partial x^\mu}\right)_p = \left(\frac{\partial x'^\nu}{\partial x^\mu}\right)_{\phi(p)} \left(\frac{\partial}{\partial x'^\nu}\right)_p.$$

So let X^μ be the components of X_p with respect to $\{(\partial/\partial x^\mu)_p\}$, and X'^μ be the components with respect to $\{(\partial/\partial x'^\mu)_p\}$. So,

$$X_p = X^\mu \left(\frac{\partial}{\partial x^\mu}\right)_p = X'^\mu \left(\frac{\partial}{\partial x'^\mu}\right)_p$$
$$= X^\mu \left(\frac{\partial x'^\sigma}{\partial x^\mu}\right)_{\phi(p)} \left(\frac{\partial}{\partial x'^\sigma}\right)_p.$$

Therefore,

$$X'^{\mu} = \left(\frac{\partial x'^{\mu}}{\partial x^{\nu}} \right)_{\phi(p)} X^{\nu}.$$

We do not have to choose a coordinate basis such as $\{(\partial/\partial x^{\mu})_p\}$. With respect to a general basis $\{e_{\mu}\}$ for T_pM, we write $X_p = X^{\mu}e_{\mu}$, for $X^{\mu} \in \mathbb{R}$ the components.

We always use summation conventions: we always contract one upstairs and one downstairs index.

1.3   Covectors

Recall that if V is a vector space over $\mathbb{R}$, the *dual space* V^* is the space of linear maps from V to $\mathbb{R}$. If V is n-dimensional, then so is V^*.

Given a basis $\{e_{\mu}\}$ for V, we define the *dual basis* $\{f^{\mu}\}$ for V^* by requiring that

$$f^{\mu}(e_{\nu}) = \delta^{\mu}{}_{\nu}.$$

If V is finite dimensional, then $V^{**} = (V^*)^*$ is isomorphic to V: to an element x of V, we assign a linear map $\Lambda_x : V^* \to \mathbb{R}$ by

$$\Lambda_x(\omega) = \omega(x).$$

Definition 1.3. The dual space of T_pM is denoted T_p^*M, and is called the *cotangent space* to M at p.

An element of this space is a covector at p. If $\{e_{\mu}\}$ is a basis for T_pM and $\{f^{\mu}\}$ the dual basis for T_p^*M, then we can expand a covector η as

$$\eta = \eta_{\mu}f^{\mu},$$

for $\eta_{\mu} \in \mathbb{R}$ the components of η.

Note that:

- $\eta(e_{\nu}) = \eta_{\mu}f^{\mu}(e_{\nu}) = \eta_{\mu}\delta^{\mu}{}_{\nu} = \eta_{\nu}$.
- $\eta(X) = \eta(X^{\mu}e_{\mu}) = X^{\mu}\eta(e_{\mu}) = X^{\mu}\eta_{\mu}$.

Definition 1.4. If $f : M \to \mathbb{R}$ is a smooth function, define $(\mathrm{d}f)_p \in T_p^*M$, the *differential* of f at p, by

$$(\mathrm{d}f)_p(X) = X(f),$$

for any $X \in T_pM$. $(\mathrm{d}f)_p$ is sometimes also called the *gradient* of f at p.

If f is a constant, then $X(f) = 0 \implies (\mathrm{d}f)_p = 0$.

If (O, ϕ) is a coordinate chart with $p \in O$ and $\phi = (x^1, \ldots, x^n)$, then we can set $f = x^\mu$ to find $(\mathrm{d}x^\mu)_p$. Now,

$$(\mathrm{d}x^\mu)_p \left(\frac{\partial}{\partial x^\nu} \right)_p = \left(\frac{\partial x^\mu}{\partial x^\nu} \right)_{\phi(p)} = \delta^\mu{}_\nu.$$

Hence, $\{(\mathrm{d}x^\mu)_p\}$ is the dual basis to $\{(\partial/\partial x^\mu)_p\}$. In this basis, we can compute

$$[(\mathrm{d}f)_p]_\mu = (\mathrm{d}f)_p \left(\frac{\partial}{\partial x^\mu} \right)_p = \left(\frac{\partial}{\partial x^\mu} \right)_p f = \left(\frac{\partial F}{\partial x^\mu} \right)_{\phi(p)},$$

justifying the language of the 'gradient'.

It can be shown that if (O', ϕ') is another chart with $p \in O'$, then

$$(\mathrm{d}x^\mu)_p = \left(\frac{\partial x^\mu}{\partial x'^\nu} \right)_{\phi'(p)} (\mathrm{d}x'^\nu)_p,$$

where $x(x') = \phi \circ \phi'^{-1}$, and hence if η_μ, η'_μ are components with respect to these bases, then

$$\eta'_\mu = \left(\frac{\partial x^\nu}{\partial x'^\mu} \right)_{\phi'(p)} \eta_\nu.$$

1.4   The (Co)tangent bundle

We can glue together the tangent spaces $T_p M$ as p varies to get a new $2n$ dimensional manifold, TM, the *tangent bundle*. We let

$$TM = \bigcup_{p \in M} \{p\} \times T_p M,$$

the set of ordered pairs (p, X), with $p \in M$, $X \in T_p M$. If $\{(O_\alpha, \phi_\alpha)\}$ is an atlas on M, we obtain an atlas for TM by setting

$$\tilde{O}_\alpha = \bigcup_{p \in O_\alpha} \{p\} \times T_p M,$$

and

$$\tilde{\phi}_\alpha((p, X)) = (\phi_\alpha(p), X^\mu) \in U_\alpha \times \mathbb{R}^n = \tilde{U}_\alpha,$$

where X^μ are the components of X with respect to the coordinate basis of ϕ_α.

It can be shown that if (O, ϕ) and (O', ϕ') are two charts on M, on $\tilde{U} \cap \tilde{U}'$, if we write $\phi' \circ \phi^{-1}(x) = x'(x)$, then

$$\tilde{\phi}' \circ \tilde{\phi}^{-1}(x, X^\mu) = \left(x'(x), \frac{\partial x'^\mu}{\partial x^\nu} X^\nu \right).$$

This lets us deduce that TM is a manifold.

A similar construction permits us to define the *cotangent bundle*

$$T^*M = \bigcup_{p \in M} \{p\} \times T_p^*M.$$

The map $\pi : TM \to M$ which takes $(p, X) \mapsto p$ is smooth (show this!)

1.5   Abstract Index Notation

We have used Greek letters μ, ν to label components of vectors (or covectors) with respect to the bases $\{e_\mu\}$. Equations involving these quantities refer to the specific basis, for example if we write $X^\mu = \delta^\mu$, this says X only has one non-zero component *in the current basis*, which will not be true in other bases.

However, we know that some equations hold in all bases, e.g.

$$\eta(X) = X^\mu \eta_\mu.$$

To capture this, we can use *abstract index notation*. We denote a vector by X^a, where the latin index a does not denote a component, rather it tells us that X^a is a vector. Similarly we denote a covector η by η_a.

If an equation is true in all bases, then we can replace Greek indices by latin indices:

$$\eta(X) = X^a \eta_a = \eta_a X^a,$$

or

$$X(f) = X^a (\mathrm{d}f)_a.$$

An equation in abstract index notation can always be turned into an equation for components, by picking a basis and changing $a \to \mu$, $b \to \nu$.

1.6   Tensors

In Newtonian physics, we know that some quantities are described by higher rank orders, e.g. the inertia tensor or the metric.

Definition 1.5. A *tensor* of type (r, s) is a multilinear map

$$T : (T_p^*M)^r \times (T_pM)^s \to \mathbb{R}.$$

> **Example 1.3.**
>
> 1. A tensor of type $(0, 1)$ is a linear map $T_pM \to \mathbb{R}$, i.e. just a covector.
>
> 2. A tensor of type $(1, 0)$ is a linear map $T^*pM \to \mathbb{R}$, i.e. an element of $(T^*pM)^* \cong T_pM$, a vector.
>
> 3. We can define a $(1, 1)$ tensor δ by
>
> $$\delta(\omega, X) = \omega(X),$$
>
> where $\omega \in T_p^*M$, $X \in T_pM$.

If $\{e_\mu\}$ is a basis for T_pM and $\{f^\mu\}$ its dual basis, then the components of an (r, s) tensor T are

$$T^{\mu_1 \cdots \mu_r}{}_{\nu_1 \cdots \nu_s} = T(f_1^\mu, \ldots, f^{\mu_r}, e_{\nu_1}, \ldots, e_{\nu_s}).$$

In abstract index notation, we denote T by $T^{a_1 \cdots a_r}{}_{b_1 \cdots b_s}$. Tensors at p form a vector space over $\mathbb{R}$ of dimension n^{r+s}.

> **Example 1.4.**
>
> 1. Consider δ above. Then,
>
> $$\delta^\mu{}_\nu = \delta(f^\mu, e_\nu) = f^\mu(e_\nu) = \begin{cases} 1 & \mu = \nu, \\ 0 & \mu \neq \nu. \end{cases}$$
>
> We can write δ as $\delta^a{}_b$ in AIN.
>
> 2. Consider a $(2, 1)$ tensor T, and let $\omega, \eta \in T_p^*M$, $X \in T_pM$. Then,
>
> $$\begin{aligned} T(\omega, \eta, X) &= T(\omega_\mu f^\mu, \eta_\nu f^\nu, X^\sigma e_\sigma) \\ &= \omega_\mu \eta_\nu X^\sigma T(f^\mu, f^\nu, e_\sigma) \\ &= \omega_\mu \eta_\nu X^\sigma T^{\mu\nu}{}_\sigma. \end{aligned}$$
>
> In AIN,
> $$T(\omega, \nu, X) = \omega_a \eta_b X^c T^{ab}{}_c.$$
>
> This can be generalized to higher ranks.

1.7   Change of Bases

We have already seen how the components of X or η with respect to a coordinate basis (X^μ, η_ν respectively) change under a change of coordinates.

But we do not only have to consider coordinate bases.

Suppose $\{e_\mu\}$ and $\{e'_\mu\}$ are two bases for $T_p M$ with dual bases $\{f^\mu\}$ and $\{f'^\mu\}$. As these are bases, we can expand

$$f'^\mu = A^\mu{}_\nu f^\nu, \qquad e'_\mu = B^\nu{}_\mu e_\nu,$$

for some $A^\mu{}_\nu, B^\mu{}_\nu \in \mathbb{R}$. But,

$$\begin{aligned}
\delta^\mu{}_\nu = f'^\mu(e'_\nu) &= A^\mu{}_\tau f^\tau (B^\sigma{}_\nu e_\sigma) \\
&= A^\mu{}_\tau B^\sigma{}_\nu f^\tau(e_\sigma) = A^\mu{}_\tau B^\sigma{}_\nu \delta^\tau{}_\sigma \\
&= A^\mu{}_\sigma B^\sigma{}_\nu.
\end{aligned}$$

Therefore, looking at these as matrices,

$$B^\mu{}_\nu = (A^{-1})^\mu{}_\nu.$$

If

$$e_\mu = \left(\frac{\partial}{\partial x^\mu}\right)_p, \qquad e'_\mu = \left(\frac{\partial}{\partial x'^\mu}\right)_p,$$

then we have already seen

$$A^\mu{}_\nu = \left(\frac{\partial x'^\mu}{\partial x^\nu}\right)_{\phi(p)}, \qquad B^\mu{}_\nu = \left(\frac{\partial x^\mu}{\partial x'^\nu}\right)_{\phi(p)},$$

which indeed satisfies $A^\mu{}_\sigma B^\sigma{}_\nu = \delta^\mu{}_\nu$, by the chain rule.

A chance of bases induces a transformation of tensor components, for example if T is a $(1,1)$-tensor, then

$$\begin{aligned}
T^\mu{}_\nu &= T(f^\mu, e_\nu) \\
T'^\mu{}_\nu = T(f'^\mu, e'_\nu) &= T(A^\mu{}_\sigma f^\sigma, (A^{-1})^\tau{}_\nu e_\tau) \\
&= A^\mu{}_\sigma (A^{-1})^\tau{}_\nu T(f^\sigma, e_\tau) \\
&= A^\mu{}_\sigma (A^{-1})^\tau{}_\nu T^\sigma{}_\tau.
\end{aligned}$$

1.8   Tensor Operations

Given an (r,s)-tensor, we can form a $(r-1, s-1)$-tensor by *contraction*.

For simplicity, assume T is a $(2,2)$-tensor. Define a $(1,1)$-tensor S by

$$S(\omega, X) = T(\omega, f^\mu, X, e_\mu).$$

To see that this is independent of the choice of basis, note

$$T(\omega, f'^\mu, X, e'_\mu) = T(\omega, A^\mu{}_\sigma f^\sigma, X, (A^{-1})^\tau{}_\mu e_\tau) = A^\mu{}_\sigma (A^{-1})^\tau{}_\mu T(\omega, f^\sigma, X, e_\tau)$$
$$= \delta^\tau{}_\sigma T(\omega, f^\sigma, X, e_\tau) = T(\omega, f^\tau, X, e_\tau) = S(\omega, X).$$

So this does not depend on the choice of basis. S and T have components related by

$$S^\mu{}_\nu = T^{\mu\sigma}{}_{\nu\sigma}.$$

In any basis in AIN we write

$$S^a{}_b = T^{ac}{}_{bc}.$$

We can generalise to contract any upstairs index with any downstairs index in a general (r, s)-tensor.

Another way to make new tensors from old tensors is to form the *tensor product*. If S is a (p, q)-tensor and T is a (r, s)-tensor, then $S \otimes T$ is an $(p + r, q + s)$-tensor:

$$S \otimes T(\omega^1, \ldots, \omega^p, \eta^1, \ldots, \eta^r, X_1, \ldots, X_q, Y_1, \ldots, Y_s)$$
$$= S(\omega^1, \ldots, \omega^p, X_1, \ldots, X_q) T(\eta^1, \ldots, \eta^r, Y_1, \ldots, Y_s).$$

This is independent of basis. In AIN,

$$(S \otimes T)^{a_1 \ldots a_p b_1 \ldots b_r}{}_{c_1 \ldots c_q d_1 \ldots, d_s} = S^{a_1 \ldots a_r}{}_{c_1 \ldots c_q} T^{b_1 \ldots b_r}{}_{d_1 \ldots b_r}.$$

We can show that for any $(1, 1)$-tensor T, in any basis we have

$$T = T^\mu{}_\nu e_\mu \otimes f^\nu.$$

The final tensor operations we require are *(anti)symmetrization*. If T is a $(0, 2)$-tensor, we can define two new tensors:

$$S(X, Y) = \frac{1}{2}(T(X, Y) + T(Y, X)),$$
$$A(X, Y) = \frac{1}{2}(T(X, Y) - T(Y, X)).$$

In AIN,

$$S_{ab} = \frac{1}{2}(T_{ab} + T_{ba}) = T_{(ab)},$$
$$A_{ab} = \frac{1}{2}(T_{ab} - T_{ba}) = T_{[ab]}.$$

These operations can be applied to any pair of matching symmetries in a more general tensor, for example:

$$T^{a(bc)}{}_{de} = \frac{1}{2}(T^{abc}{}_{de} + T^{acb}{}_{de}).$$

We can also (anti)symmetrize over more than two indices.

- To symmetrize over n indices, we sum over all permutations of the indices and divide by $n!$.

- To anti-symmetrize over n indices, we sum over all permutation weighted by their sign, and then divide by $n!$.

For example,

$$T^{(abc)} = \frac{1}{3!}(T^{abc} + T^{bca} + T^{cab} + T^{acb} + T^{cba} + T^{bac}),$$

$$T^{[abc]} = \frac{1}{3!}(T^{abc} + T^{bca} + T^{cab} - T^{acb} - T^{cba} - T^{bac}).$$

To exclude indices from (anti)symmetrization, we use vertical lines:

$$T^{(a|b|c)} = \frac{1}{2}(T^{abc} + T^{cba}).$$

1.9   Tensor Bundles

The space of (r,s)-tensors at a point p is the vector space $(T^r{}_s)_p M$. These can be glued together to form the bundle of (r,s)-tensors

$$T^r{}_s M = \bigcup_{p \in M} \{p\} \times (T^r{}_s)_p M.$$

If (O, ϕ) is a coordinate chart on M, set

$$\tilde{O} = \bigcup_{p \in O} \{p\} \times (T^r{}_s)_p M \subseteq T^r{}_s M,$$

and

$$\tilde{\phi}(p, S_p) = (\phi(p), S^{\mu_1 \dots \mu_r}{}_{\nu_1 \dots \nu_s}).$$

$T^r{}_s M$ is a manifold, with a natural smooth map $\Pi : T^r{}_s M \to M$ such that $\Pi(p, S_p) = p$.

A *tensor field* is a smooth map $T : M \to T^r{}_s M$ such that $\pi \circ T = \mathrm{id}$.

If (O, ϕ) is a coordinate chart on M, then

$$\tilde{\phi} \circ T \circ \phi(x) = (x, T^{\mu_1 \dots \mu_r}{}_{\nu_1 \dots \nu_s}(x)),$$

which is smooth provided the components $T^{\mu_1 \dots \mu_r}{}_{\nu_1 \dots \nu_s}(x)$ are smooth functions of x.

> **Example 1.5.**
>
> If $T^r_s M = T^1_0 M \cong TM$, the tensor field is called a *vector field*. In a local coordinate patch, if X is a vector field, we write $X(p) = (p, X_p)$, with
>
> $$X_p = X^\mu(x) \left(\frac{\partial}{\partial x^\mu} \right)_p.$$
>
> In particular, $\partial/\partial x^\mu$ are always smooth, but only defined locally.

A vector field can act on a function $f : M \to \mathbb{R}$ to give a new function Xf by

$$Xf(p) = X_p(f).$$

In coordinates,

$$Xf(p) = X^\mu(\phi(p)) \left.\frac{\partial F}{\partial x^\mu}\right|_{\phi(p)},$$

where $F = f \circ \phi^{-1}$.

1.10   Integral Curves

Given a vector field X on M, we say a curve $\lambda : I \to M$ is an *integral curve* of X if its tangent at every point x, which we denote by $\mathrm{d}\lambda/\mathrm{d}t$, then

$$\frac{\mathrm{d}\lambda}{\mathrm{d}t}(t) = X_{\lambda(t)}.$$

Through each point p, there is a unique integral curve passing through it, unique up to extension or shift of the parameter.

To see this, pick a chart ϕ with $\phi = (x^1, \ldots, x^n)$, and $\phi(p) = 0$. In this chart, the equation becomes

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}t}(t) = X^\mu(x(t)),$$

where $x^\mu(t) = x^\mu(\lambda(t))$. Assuming without loss of generality that $\lambda(0) = p$, we get an initial condition $x^\mu(0) = 0$.

By standard ODE theory, this equation with initial conditions has a solution which is unique up to extension.

1.11   Commutators

Suppose X and Y are two vector fields, and $f : M \to \mathbb{R}$ is smooth. Then $X(Y(f))$ is a smooth function, however it is not of the form $K(f)$ for some vector field K, since

$$X(Y(fg)) = X(fY(g) + gY(f)) = X(f(Y(g)) + X(gY(f))$$
$$= fX(Y(g)) + gX(Y(f)) + X(f)Y(g) + X(g)Y(f).$$

So the Leibniz rule does not hold. But, if we consider $[X, Y](f) = X(Y(f)) - Y(X(f))$, then the Leibniz rule does hold. In fact, $[X, Y]$ defines a vector field.

To see this, we can use coordinates:

$$
\begin{aligned}
[X, Y](f) &= X\left(Y^\nu \frac{\partial f}{\partial x^\nu}\right) - Y\left(X^\mu \frac{\partial f}{\partial x^\mu}\right) \\
&= X^\mu \frac{\partial}{\partial x^\mu}\left(Y^\nu \frac{\partial f}{\partial x^\nu}\right) - Y^\nu \frac{\partial}{\partial x^\nu}\left(X^\mu \frac{\partial F}{\partial x^\mu}\right) \\
&= X^\mu Y^\nu \left(\frac{\partial^2 f}{\partial x^\mu \partial x^\nu} - \frac{\partial^2 f}{\partial x^\nu \partial x^\mu}\right) + X^\mu \frac{\partial Y^\nu}{\partial x^\mu}\frac{\partial F}{\partial x^\nu} - Y^\nu \frac{\partial X^\mu}{\partial x^\nu}\frac{\partial F}{\partial x^\mu} \\
&= \left(X^\mu \frac{\partial Y^\nu}{\partial x^\mu} - Y^\mu \frac{\partial X^\nu}{\partial x^\mu}\right)\frac{\partial F}{\partial x^\nu} \\
&= [X, Y]^\nu \frac{\partial F}{\partial x^\nu},
\end{aligned}
$$

where

$$[X, Y]^\nu = X^\mu \frac{\partial Y^\nu}{\partial x^\mu} - Y^\mu \frac{\partial X^\nu}{\partial x^\mu}$$

are the components of the commutator. Since f is arbitrary,

$$[X, Y] = [X, Y]^\nu \frac{\partial}{\partial x^\nu}$$

is valid only in a coordinate basis.

1.12   Metric Tensor

We are familiar from Euclidean geometry (and special relativity) with the fact that the fundamental object when talking about distances and angles (or time intervals and rapidity) is an inner product between vectors. For example,

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + x_3 y_3$$

in $\mathbb{R}^3$ with Euclidean geometry, and

$$X \cdot Y = -X^0 Y^0 + X^1 Y^1 + X^2 Y^2 + X^3 Y^3$$

for $\mathbb{R}^{3+1}$ with the Minkowski geometry (and $c = 1$).

Definition 1.6. A *metric tensor* at $p \in M$ is a $(0, 2)$-tensor satisfying:

  (i) g is symmetric, so $g(X, Y) = g(Y, X)$ for all $X, Y \in T_p M$. In coordinates, $g_{ab} = g_{ba}$.

  (ii) g is *non-degenerate*: $g(X, Y) = 0$ for all $Y \in T_p M$ if and only if $X = 0$.

We sometimes write

$$g(X, Y) = \langle X, Y \rangle = \langle X, Y \rangle_g = X \cdot Y.$$

By adapting the Gram-Schmidt algorithm, we can always find a basis $\{e_\mu\}$ for $T_p M$ such that

$$g(e_\mu, e_\nu) = \begin{cases} 0 & \mu \neq \nu, \\ +1 \text{ or } -1 & \mu = \nu. \end{cases}$$

The number of -1's and 1's does not depend on the choice of basis, by Sylvester's law of inertia, and is called the *signature*.

If g has signature that is entirely positive, we say it is *Riemannian*.

If g has signature that is positive apart from one component, we say it is *Lorentzian*.

Definition 1.7. A *Riemannian* (resp. *Lorentzian*) manifold is a pair (M, g) where M is a manifold, and g is a Riemannian (resp. Lorentzian) metric tensor field.

On a Riemannian manifold, the *norm* of a vector $X \in T_p M$ is

$$|X| = \sqrt{g(X, X)},$$

and the *angle* between $X, Y \in T_p M$ is given by

$$\cos \theta = \frac{g(X, Y)}{|X||Y|}.$$

The *length* of a curve $\lambda : (a, b) \to M$ is given by

$$\ell(\lambda) = \int_a^b \left| \frac{\mathrm{d}\lambda}{\mathrm{d}t}(t) \right| \mathrm{d}t.$$

It is an easy exercise to show that if $\tau : (c, d) \to (a, b)$ is a reparametrization with $d\tau / du > 0$, and $\tau(c) = a$, $\tau(d) = b$, then $\tilde{\lambda} = \lambda \circ \tau : (c, d) \to M$ is a reparametrization of λ, in which $\ell(\tilde{\lambda}) = \ell(\lambda)$.

In a coordinate basis, $g = g_{\mu\nu} \, dx^\mu \otimes dx^\nu$. We often write

$$dx^\mu \, dx^\nu = \frac{1}{2}(dx^\mu \otimes dx^\nu + dx^\nu \otimes dx^\mu).$$

And by convention we often write $g = ds^2$, so that

$$g = ds^2 = g_{\mu\nu} \, dx^\mu \, dx^\nu.$$

Example 1.6.

(i) Consider $\mathbb{R}^n$ with

$$g = ds^2 = (dx^1)^2 + (dx^2)^2 + (dx^n)^2 + \delta_{\mu\nu} \, dx^\mu \, dx^\nu.$$

This is called Euclidean space, and any chart covering $\mathbb{R}^n$ in which the metric takes this form is called *Cartesian*.

(ii) Consider $\mathbb{R}^{1+3}$ with

$$g = ds^2 = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2 = \eta_{\mu\nu} \, dx^\mu \, dx^\nu.$$

This is *Minkowski space*. A coordinate chart covering $\mathbb{R}^{1+3}$ in which the metric takes this form is an *intertial frame*.

(iii) On $S^2 = \{\mathbf{x} \in \mathbb{R}^3 \mid |\mathbf{x}| = 1\}$, define a chart by $\phi^{-1} : (0, \pi) \times (-\pi, \pi) \to S^2$ by

$$(\theta, \varphi) \mapsto (\sin\theta \cos\varphi, \sin\theta \sin\varphi, \cos\theta).$$

In this chart, the round metric is

$$g = ds^2 = d\theta^2 + \cos^2\theta \, d\varphi^2.$$

This covers $S^2 \setminus \{|\mathbf{x}| = 1, x^2 = 0, x^1 \leq 0\}$.

To cover the rest, let $\tilde{\phi}^{-1}$ be given by

$$(\theta', \varphi') \mapsto (-\sin\theta' \cos\varphi', \cos\theta', \sin\theta' \sin\varphi').$$

Setting $g = d\theta'^2 + \sin^2\theta' \, d\varphi'^2$ defines a metric on all of S^2.

Since g_{ab} is non-degenerate, it is invertible as a matrix in any basis. We can check

the inverse define a symmetric $(2, 0)$-tensor g^{ab} satisfying

$$g^{ab}g_{bc} = \delta^a{}_c.$$

> **Example 1.7.**
>
> In the ϕ coordinates of the S^2 example, note
>
> $$g^{\mu\nu} = \left(1, \frac{1}{\sin^2\theta}\right).$$

An important property of the metric is that it induces a canonical identification of T_pM and T_p^*M.

Given $X^a \in T_pM$, we define a covector $g_{ab}X^b = X_a$, and given $\eta_a \in T_p^*M$, we define a vector $g^{ab}\eta_b = \eta^a$.

In $(\mathbb{R}^3, \delta)$ we often do this without realizing. More generally, this allows us to raise tensor indices with g^{ab} and lower with g_{ab}. For example if $T^{ab}{}_c$ is $(2, 1)$-tensor, then $T_a{}^{bc}$ is the $(2, 1)$-tensor given by

$$T_a{}^{bc} = g_{ad}g^{ce}T^{db}{}_e.$$

1.13　Lorentzian Signature

In Lorentzian signature, with indices $0, 1, \ldots, n$. At any point in a Lorentzian manifold we can find a basis $\{e_\mu\}$ such that

$$g(e_\mu, e_\nu) = \eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, \ldots, 1).$$

This basis is not unique; if $e'_\mu = (A^{-1})^\nu{}_\mu e_\nu$ is another such basis, then

$$\eta_{\mu\nu} = g(e'_\mu, e'_\nu) = (A^{-1})^\sigma{}_\mu (A^{-1})^\tau{}_\nu g(e_\sigma, e_\tau) = (A^{-1})^\sigma{}_\mu (A^{-1})^\tau{}_\nu \eta_{\sigma\tau}.$$

This is the condition that A is a *Lorentz transformation*.

The tangent space at p has $\eta_{\mu\nu}$ as a metric tensor, so it has the structure of Minkowski space.

Definition 1.8. $X \in T_pM$ is:

- *spacelike* if $g(X, X) > 0$,
- *null* or *lightlike* if $g(X, X) = 0$,
- *timelike* if $g(X, X) < 0$.

A curve $\lambda : I \to M$ in a Lorentzian manifold is *spacelike/timelike/null* if the tangent vector is everywhere spacelike/timelike/null respectively.

A spacelike curve has a well-defined *length*, given by the same formula as in the Riemannian case. For a timelike curve $\lambda : (a, b) \to M$, the relevant quantity is the *proper time*

$$\tau(\lambda) = \int_a^b \sqrt{-g_{ab} \frac{\mathrm{d}x^a}{\mathrm{d}u} \frac{\mathrm{d}x^b}{\mathrm{d}u}} \, \mathrm{d}u.$$

If

$$g_{ab} \frac{\mathrm{d}\lambda^a}{\mathrm{d}u} \frac{\mathrm{d}\lambda^b}{\mathrm{d}u} = -1$$

for all u, then λ is parametrized by proper time. In this case, we call the tangent vector $U^a = \frac{\mathrm{d}\lambda^a}{\mathrm{d}x}$ the *four-velocity* of λ.

1.14   Curves of Extremal Proper Time

Suppose $\lambda : (0, 1) \to M$ is timelike, satisfies $\lambda(0) = p$, $\lambda(1) = q$, and extremizes the proper time along all such curves. This is a variational problem associated to

$$\tau[\lambda] = \int_0^1 G(x^\mu(u), \dot{x}^\mu(u)) \, \mathrm{d}u,$$

with

$$G(x^\mu(u), \dot{x}^\mu(u)) = \sqrt{-g_{\mu\nu}(x(u)) \dot{x}^\mu(u) \dot{x}^\nu(u)}$$

The Euler-Lagrange equation is

$$\frac{\mathrm{d}}{\mathrm{d}u} \left(\frac{\partial G}{\partial \dot{x}^\mu} \right) = \frac{\partial G}{\partial x^\mu}.$$

We can compute

$$\frac{\partial G}{\partial \dot{x}^\mu} = -\frac{1}{G} g_{\mu\nu} \dot{x}^\nu,$$
$$\frac{\partial G}{\partial x^\mu} = -\frac{1}{2G} \frac{\partial}{\partial x^\mu} (g_{\mu\nu}) \dot{x}^\sigma \dot{x}^\tau$$
$$= -\frac{1}{2G} g_{\sigma\tau,\mu} \dot{x}^\sigma \dot{x}^\tau.$$

Fix the parametrization so that the curve is parametrized by proper time τ. Doing this, we get

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \dot{x}^\mu \frac{\mathrm{d}u}{\mathrm{d}\tau}, \qquad -1 = g_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}.$$

We deduce that

$$-1 = g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu \left(\frac{\mathrm{d}u}{\mathrm{d}\tau}\right)^2 \implies \frac{\mathrm{d}u}{\mathrm{d}\tau} = \frac{1}{G}$$

$$\implies \frac{1}{G}\frac{\mathrm{d}}{\mathrm{d}u} = \frac{\mathrm{d}}{\mathrm{d}\tau}.$$

Returning to the Euler-Lagrange quations, we find

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(g_{\mu\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\right) = \frac{1}{2}g_{\nu\rho,\mu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau}.$$

Hence expanding the LHS,

$$g_{\mu\nu}\frac{\mathrm{d}^2 x^\nu}{\mathrm{d}\tau^2} + g_{\mu[\nu,\rho]}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} - \frac{1}{2}g_{\sigma\rho,\mu}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} = 0,$$

after symmetrizing. Thus,

$$\frac{\mathrm{d}^2 x^\nu}{\mathrm{d}\tau^2} + \Gamma_\mu{}^\nu{}_\rho \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} = 0, \qquad\qquad (*)$$

where

$$\Gamma_\mu{}^\nu{}_\rho = \frac{1}{2}g^{\nu\sigma}\left(g_{\mu\sigma,\rho} + g_{\sigma\rho,\mu} - g_{\mu\rho,\sigma}\right)$$

are the *Christoffel symbols* of g. Note that:

- $\Gamma_\nu{}^\mu{}_\rho = \Gamma_\rho{}^\mu{}_\nu$.

- $\Gamma_\nu{}^\mu{}_\rho$ are not tensor components.

- We can solve $(*)$ with standard ODE theory, to get *geodesics*.

- The same equation governs curves of extremal length in a Riemannian manifold, or spacelike curves in a Lorentzian manifold, parametrized by arc-length.

We can show that the geodesic equation can be obtained as the Euler-Lagrange equation for the Lagrangian

$$L = -g_{\mu\nu}(x(\tau))\dot{x}^\mu(\tau)\dot{x}^\nu(\tau).$$

Example 1.8.

1. In Minkowski space in an inertial frame, $g_{\mu\nu} = \eta_{\mu\nu}$, so $\Gamma_\mu{}^\nu{}_\rho = 0$, and

the geodesic equation is

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} = 0.$$

These are straight lines.

2. The Schwarzschild metric in the Schwarzchild coordinates is a metric on $M = \mathbb{R}_t \times (2m, \infty)_r \times S^2_{\theta,\varphi}$ given by

$$\mathrm{d}s^2 = -f\,\mathrm{d}t^2 + \frac{\mathrm{d}r^2}{f} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2),$$

where $f = 1 - 2m/r$. Then

$$L = f\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 - \frac{1}{f}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 - r^2\left(\frac{\mathrm{d}\theta}{\mathrm{d}\tau}\right)^2 - r^2\sin^2\theta\left(\frac{\mathrm{d}\varphi}{\mathrm{d}\tau}\right)^2.$$

The Euler-Lagrange equation for $t(\tau)$ is

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial L}{\partial t'}\right) = \frac{\partial L}{\partial t},$$

where $t' = \mathrm{d}t/\mathrm{d}\tau$. This gives

$$2\frac{\mathrm{d}}{\mathrm{d}\tau}\left(f\frac{\mathrm{d}t}{\mathrm{d}\tau}\right) = 0,$$

or

$$f\frac{\mathrm{d}^2 t}{\mathrm{d}\tau^2} + \frac{\mathrm{d}f}{\mathrm{d}r}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right) = 0.$$

We can compare this with the geodesic equation to see that

$$\Gamma_{1\ 0}^{\ 0} = \Gamma_{0\ 1}^{\ 0} = \frac{1}{2f}\frac{\mathrm{d}f}{\mathrm{d}r},$$

and $\Gamma_{\mu\ \nu}^{\ 0} = 0$ otherwise. The other symbols can be found from the other Euler-Lagrange equations.

1.15   Covariant Derivative

For a function $f : M \to \mathbb{R}$, we know that $\partial f/\partial x^\mu$ are components of a covector $(\mathrm{d}f)$. For a vector field, we cannot just differentiate the components.

Indeed, we can show that if V is a vector field, then

$$T^{\mu}{}_{\nu} = \frac{\partial V^{\mu}}{\partial x^{\nu}}$$

do not form the components of a $(1,1)$-tensor.

Definition 1.9. A *covariant derivative* ∇ on a manifold M is a map sending X, Y, smooth vector fields, to a vector field $\nabla_X Y$ satisfying, for X, Y, Z smooth vector fields and f, g functions:

   (i) $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$.

   (ii) $\nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z$.

   (iii) $\nabla_X(fY) = f\nabla_X Y + (\nabla_X f)Y$, where $\nabla_X f = X(f)$.

Here a *smooth vector field* is a map $X : C^{\infty}(M) \to C^{\infty}(M)$ that satisfies the Leibniz rule.

Note the first definition implies $\nabla Y : X \mapsto \nabla_X Y$ is a linear map of $T_p M$ to itself, so it defines a $(1,1)$-tensor, the *covariant derivative* of Y. In AIN,

$$(\nabla Y)^a{}_b = \nabla_b Y^a,$$

or $Y^a{}_{;b}$.

Definition 1.10. In a basis $\{e_{\mu}\}$, the *connection components* $\Gamma_{\nu}{}^{\mu}{}_{\rho}$ are defined by

$$\nabla_{e_{\rho}} e_{\nu} = \Gamma_{\nu}{}^{\mu}{}_{\rho} e_{\mu}.$$

These define ∇, as

$$\begin{aligned}
\nabla_X Y = \nabla_{X^{\mu} e_{\mu}}(Y^{\nu} e_{\nu}) &= X^{\mu} \nabla_{e_{\mu}}(Y^{\nu} e_{\nu}) \\
&= X^{\mu}(e_{\mu}(Y^{\nu})e_{\nu} + Y^{\sigma} \nabla_{e_{\mu}} e_{\sigma}) \\
&= (X^{\mu} e_{\mu}(Y^{\nu}) + \Gamma_{\sigma}{}^{\nu}{}_{\mu} Y^{\sigma} X^{\mu}) e_{\nu}.
\end{aligned}$$

Hence,

$$(\nabla_X Y)^{\nu} = X^{\mu}(e_{\mu}(Y^{\nu}) + \Gamma_{\sigma}{}^{\nu}{}_{\mu} Y^{\sigma}),$$

and so

$$Y^{\nu}{}_{;\mu} = e_{\mu}(Y^{\nu}) + \Gamma_{\sigma}{}^{\nu}{}_{\mu} Y^{\sigma}.$$

In the coordinate bases $e_{\mu} = \partial/\partial x^{\mu}$, then

$$Y^{\nu}{}_{;\mu} = \frac{\partial Y^{\nu}}{\partial x^{\mu}} + \Gamma_{\sigma}{}^{\nu}{}_{\mu} Y^{\sigma} = Y^{\nu}{}_{,\mu} + \Gamma_{\sigma}{}^{\nu}{}_{\mu} Y^{\sigma}.$$

Again note $\Gamma_\mu{}^\nu{}_\sigma$ are not components of a tensor.

We can extend ∇ to arbitrary tensor field by requiring that the Leibniz property holds, for example for η a tensor field, we define $\nabla_X \eta$ by

$$(\nabla_X \eta)(Y) = \nabla_X(\eta(Y)) - \eta(\nabla_X Y).$$

In components,

$$\begin{aligned}
(\nabla_X \eta)Y &= X^\mu e_\mu(\eta_\sigma Y^\sigma) - \eta_\sigma(\nabla_X Y)^\sigma \\
&= X^\mu e_\mu(\eta_\sigma)Y^\sigma + X^\mu \eta_\sigma e_\mu(Y^\sigma) - \eta_\sigma(X^\nu e_\nu(Y^\sigma) + X^\nu \Gamma_\tau{}^\sigma{}_\nu Y^\tau) \\
&= (e_\mu(\eta_\sigma) - \Gamma_\sigma{}^\nu{}_\mu \eta_\nu)X^\mu Y^\sigma,
\end{aligned}$$

which is linear in X and Y. Hence $\nabla \eta$ is a $(0,2)$-tensor, with components

$$\nabla_\mu \eta_\sigma = e_\mu(\eta_\sigma) - \Gamma_\sigma{}^\nu{}_\mu \eta_\nu = \eta_{\sigma;\mu}.$$

In a coordinate basis,

$$n_{\sigma;\mu} = n_{\sigma,\mu} - \Gamma_\sigma{}^\nu{}_\mu \eta_\nu.$$

In general, we can show, due to the Leibniz rule, that in a coordinate basis,

$$\begin{aligned}
T^{\mu_1 \cdots \mu_r}{}_{\nu_1 \ldots \nu_s;\rho} &= T^{\mu_1 \cdots \mu_r}{}_{\nu_1 \ldots \nu_s,\rho} + \Gamma_\sigma{}^{\mu_1}{}_\rho T^{\sigma \mu_2 \cdots \mu_r}{}_{\nu_1 \ldots \nu_s} + \cdots \\
&\quad + \Gamma_\sigma{}^{\mu_r}{}_\rho T^{\mu_1 \cdots \mu_{r-1}\sigma}{}_{\nu_1 \ldots \nu_s} - \Gamma_{\nu_1}{}^\sigma{}_\rho T^{\mu_1 \cdots \mu_r}{}_{\sigma \nu_2 \ldots \nu_s} - \cdots \\
&\quad - \Gamma_{\nu_s}{}^\sigma{}_\rho T^{\mu_1 \cdots \mu_r}{}_{\nu_1 \ldots \nu_{s-1}\sigma}.
\end{aligned}$$

Remark. If $T^a{}_b$ is a $(1,1)$-tensor, then $T^a{}_{b;c}$ is a $(1,2)$-tensor, and we can take further derivatives:

$$(T^a{}_{b;c})_{;d} = T^a{}_{b;cd} = \nabla_d \nabla_c T^a{}_b.$$

In general, $T^a{}_{b;cd} \neq T^a{}_{b;dc}$.

If f is a function, then $f_{;a} = (\mathrm{d}f)_a$ is a covector. In a coordinate basis $f_{;\mu} = f_{,\mu}$, then

$$f_{;\mu\nu} = f_{,\mu\nu} - \Gamma_\mu{}^\sigma{}_\nu f_{,\sigma},$$

hence

$$f_{;[\mu\nu]} = -\Gamma_{[\mu}{}^\sigma{}_{\nu]} f_{,\sigma}.$$

Definition 1.11. A connection (or covariant derivative) is *torsion free* (or symmetric) if

$$\nabla_a \nabla_b f = \nabla_b \nabla_a f$$

for any function f. In a coordinate basis, this is equivalent to

$$\Gamma_{[\mu}{}^\rho{}_{\nu]} = 0 \implies \Gamma_\mu{}^\rho{}_\nu = \Gamma_\nu{}^\rho{}_\mu.$$

Lemma 1.1. *If ∇ is torsion free, then for X, Y vector fields,*

$$\nabla_X Y - \nabla_Y X = [X, Y].$$

> **Proof:**   We show this is true in one basis: then as the objects are defined in a basis-free way, this must be true in all bases.
>
> In a basis,
>
> $$\begin{aligned}
(\nabla_X Y - \nabla_Y X)^\mu &= X^\sigma Y^\mu{}_{;\sigma} - Y^\sigma X^\mu{}_{;\sigma} \\
&= X^\sigma (Y^\mu{}_{,\sigma} + \Gamma_\rho{}^\mu{}_\sigma Y^\rho) - Y^\sigma (X^\mu{}_{,\sigma} + \Gamma_\rho{}^\mu{}_\sigma Y^\rho) \\
&= [X, Y]^\mu + 2 X^\sigma Y^\rho \Gamma_{[\rho}{}^\mu{}_{\sigma]} = [X, Y]^\mu.
\end{aligned}$$

Even if ∇ is torsion free, we may still not have $\nabla_a \nabla_b X^c = \nabla_b \nabla_a X^c$.

1.16   The Levi-Civita Connection

For a manifold with a metric, there is a preferred connection.

Theorem 1.1 (Fundamental Theorem of Riemannian Geometry)**.** *If (M, g) is a manifold with a metric, then there is a unique, torsion free connection ∇ which satisfies*

$$\nabla g = 0.$$

This is called the *Levi-Civita connection*.

> **Proof:**   Suppose such a connection exists. Then by Leibniz, if X, Y, Z are smooth vector fields, then
>
> $$X(g(Y, Z)) = \nabla_X(g(Y, Z)) = (\nabla_X g)(Y, Z) + g(\nabla_X Y, Z) + g(Y, \nabla_X Z).$$
>
> If $\nabla_X g = 0$, then
>
> $$\begin{aligned}
X(g(Y, Z)) &= g(\nabla_X Y, Z) + g(Y, \nabla_X Z), \\
Y(g(Z, X)) &= g(\nabla_Y Z, X) + g(Z, \nabla_Y X), \\
Z(g(X, Y)) &= g(\nabla_Z X, Y) + g(X, \nabla_Z Y).
\end{aligned}$$
>
> Summing the first two equations and subtracting the third,
>
> $$\begin{aligned}
X(g(Y, Z)) + Y(g(Z, X)) - Z(g(X, Y)) = {}& g(\nabla_X Y + \nabla_Y X, Z) \\
& + g(\nabla_X Z - \nabla_Z X, Y) + g(\nabla_Y Z - \nabla_Z Y, X).
\end{aligned}$$

Using the fact that $\nabla_X Y - \nabla_Y X = [X, Y]$,

$$X(g(Y, Z)) + Y(g(Z, X)) - Z(g(X, Y)) = 2g(\nabla_X Y, Z) - g([X, Y], Z)$$
$$- g([Z, X], Y) + g([Y, Z], X).$$

Hence we can write

$$g(\nabla_X Y, Z) = \frac{1}{2}\Big(X(g(Y, Z)) + Y(g(Z, X)) - Z(g(X, Y))$$
$$+ g([X, Y], Z) + g([Z, X], Y) - g([Y, Z], X)\Big).$$

This determines $\nabla_X Y$ uniquely, since g is non-degenerate. We can thus use this to define $\nabla_X Y$. Then we need to check the properties of a symmetric connection hold. For example,

$$g(\nabla_{fX} Y, Z) = \frac{1}{2}\Big(fX(g(Y, Z)) + Y(fg(Z, X)) - Z(fg(X, Y))$$
$$+ g([fX, Y], Z) + g([Z, fX], Y) - g([Y, Z], fX)\Big)$$
$$= \frac{1}{2}\Big(fX(g(Y, Z)) + fY(g(Z, X)) - fZ(g(X, Y))$$
$$+ Y(f)g(Z, X) - Z(f)g(X, Y)$$
$$+ g(f[X, Y] - Y(f)X, Z) + g(f[Z, X]$$
$$+ Z(f)X, Y) - fg([Y, Z], X)\Big).$$

Hence by cancelling terms,

$$g(\nabla_{fX} Y, Z) = g(f\nabla_X Y, Z) \implies \nabla_{fX} Y = f\nabla_X Y.$$

We can check the other properties hold ourselves.

In a coordinate basis, we can compute

$$g(\nabla_{e_\mu} e_\nu, e_\sigma) = \frac{1}{2}\Big(e_\mu(g(e_\nu, e_\sigma)) + e_\nu(g(e_\sigma, e_\mu)) - e_\sigma(g(e_\mu, e_\nu))\Big).$$

This gives

$$g(\Gamma_\nu{}^\tau{}_\mu e_\tau, e_\sigma) = \Gamma_\nu{}^\tau{}_\mu g_{\tau\sigma} = \frac{1}{2}(g_{\sigma\nu,\mu} + g_{\sigma\mu,\nu} - g_{\mu\nu,\sigma}),$$

or,

$$\Gamma_\nu{}^\tau{}_\mu = \frac{1}{2}g^{\sigma\tau}(g_{\sigma\nu,\mu} + g_{\mu\sigma,\nu} - g_{\mu\nu,\sigma}).$$

This ∇ is the Levi-Civita connection, which is the same as the connection involved when extremizing proper time. We can raise or lower indices with ∇, and this commutes with covariant differentiation.

1.17   Geodesics

We found that a curve extremizing proper time satisfies

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} + \Gamma_\nu{}^\mu{}_\rho (x(\tau)) \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} = 0,$$

where τ is the proper time. The tangent vector X^a to the curve has components

$$X^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}.$$

Extending this off of the curve, we get a vector field, of which the geodesic is an integral curve. We note that

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} = \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\right) = \frac{\partial X^\mu}{\partial x^\nu} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} = X^\mu{}_{,\nu} X^\nu,$$

hence the geodesic equation gives

$$X^\mu{}_{,\nu} X^\nu + \Gamma_\nu{}^\mu{}_\rho X^\nu X^\rho = 0 \iff X^\nu X^\mu{}_{;\nu} = 0 \implies \nabla_X X = 0.$$

We can extend this to any connection.

Definition 1.12. Let M be a manifold with connection ∇. An *affinely parametrized geodesic* satisfies

$$\nabla_X X = 0,$$

where X is the tangent vector.

If we reparametrize $t \to t(u)$, then

$$\underbrace{\frac{\mathrm{d}x^\mu}{\mathrm{d}u}}_{Y} = \underbrace{\frac{\mathrm{d}x^\mu}{\mathrm{d}t}}_{X} \frac{\mathrm{d}t}{\mathrm{d}u},$$

so if we take $X \to Y = hX$ with $h > 0$, then

$$\nabla_Y Y = \nabla_{hX}(hX) = h \nabla_X (hX) = h^2 \underbrace{\nabla_X X}_{0} + hX \cdot X(h) = fY,$$

with

$$f = X(h) = \frac{\mathrm{d}}{\mathrm{d}t}(h) = \frac{1}{h}\frac{\mathrm{d}h}{\mathrm{d}u} = \frac{1}{h}\frac{\mathrm{d}^2 t}{\mathrm{d}u^2}.$$

So $\nabla_Y Y = 0 \iff t = \alpha u + \beta$, for $\alpha, \beta \in \mathbb{R}$, $\alpha > 0$.

Theorem 1.2. *Given $p \in M$, $X_p \in T_pM$, there exists a unique affinely parametrized geodesic $\lambda : I \to M$ satisfying*

$$\lambda(0) = p, \qquad \dot{\lambda}(0) = X_p.$$

> **Proof:** Choose coordinates with $\phi(p) = 0$. Then $x^\mu(t) = \phi(\lambda(t))$ satisfies $\nabla_X X = 0$ with $X = X^\mu \frac{\partial}{\partial x^\mu}$, $X^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}t}$. This becomes
>
> $$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}t^2} + \Gamma_\nu{}^\mu{}_\sigma \frac{\mathrm{d}x^\nu}{\mathrm{d}t} \frac{\mathrm{d}x^\sigma}{\mathrm{d}t} = 0.$$
>
> This has unique solution $x^\mu : (-\varepsilon : \varepsilon) \to \mathbb{R}^n$ for ε sufficiently small, given by standard ODE theory.

We just descend the condition $\nabla_X X = 0$ to $\mathbb{R}^n$.

1.18   Geodesic Postulate

In general relativity, free particles move along geodesics of the Levi-Civita connection. These are timelike for massive particles, and null/lightlike for massless particles.

If we fix $p \in M$, we can map T_pM into M by setting

$$\psi(X_p) = \lambda_{X_p}(1),$$

where λ_{X_p} is the unique affinely parametrized geodesic with $\lambda_{X_i}(0) = p$, $\dot{\lambda}_{X_i}(0) = X_p$. Notice that

$$\lambda_{\alpha X_p}(t) = \lambda_{X_p}(\alpha t)$$

for $\alpha \in \mathbb{R}$, since if $\tilde{\lambda}(t) = \lambda_{X_p}(\alpha t)$, this is an affine reparametrization, so is still a geodesic, and

$$\dot{\tilde{\lambda}}(0) = \alpha \dot{\lambda}_{X_p}(0) = \alpha X_p.$$

Moreover $\alpha \mapsto \psi(\alpha X_p)$ is an affinely parametrized geodesic, so must be $\lambda_{X_p}(\alpha)$.

We claim that if $U \subseteq T_pM$ is a sufficiently small neighbourhood of the origin, then $\psi : T_pM \to M$ is one-to-one and onto.

Definition 1.13. Suppose $\{e_\mu\}$ is a basis for T_pM. We construct *normal coordinates* at p as follows: for $q \in \psi(U) \subseteq M$, we define $\phi(q) = (X^1, \ldots, X^n)$, where X^μ are the components of the unique $X_p \in U$ with $\psi(X_p) = q$, with respect to $\{e_\mu\}$.

By our previous observation, the curve given in normal coordinates by $X^\mu(t) = tY^\mu$ for Y^μ a constant, is an affinely parametrized geodesic, so from the geodesic equation,

$$\Gamma_\nu{}^\mu{}_\sigma(tY)Y^\nu Y^\sigma = 0.$$

Set $t = 0$ to deduce that $\Gamma_{(\nu}{}^\mu{}_{\sigma)}|_p = 0$. So if ∇ is torsion free, then $\Gamma_\nu{}^\mu{}_\sigma|_p = 0$ in normal coordinates.

Hence if ∇ is the Levi-Civita connection of a metric, then

$$g_{\mu\nu,\rho}|_p = 0,$$

since

$$\begin{aligned}
g_{\mu\nu,\rho} &= \frac{1}{2}(g_{\mu\nu,\rho} + g_{\rho\nu,\mu} - g_{\mu\rho,\nu}) + \frac{1}{2}(g_{\mu\nu,\rho} + g_{\mu\rho,\nu} - g_{\rho\nu,\mu}) \\
&= \Gamma_\mu{}^\sigma{}_\rho g_{\sigma\nu} + \Gamma_\rho{}^\sigma{}_\nu g_{\sigma\mu} = 0
\end{aligned}$$

at p. We can always choose the basis $\{e_\mu\}$ for T_pM on which we base the normal coordinates to be orthonormal. We have:

Lemma 1.2. *On a Riemannian or Lorentzian field, we can choose normal coordinates at p such that $g_{\mu\nu,\rho}|_p = 0$ and*

$$g_{\mu\nu}|_p = \begin{cases} \delta_{\mu\nu} & \text{Riemannian}, \\ \eta_{\mu\nu} & \text{Lorentzian}. \end{cases}$$

Proof: The curve given in normal coordinates by $t \mapsto (t, 0, \ldots, 0)$ is the APG with $\lambda(0) = p$, and $\dot\lambda(0) = e_1$, by the previous argument. But by definition of the coordinate basis, this vector is

$$\left(\frac{\partial}{\partial x^1}\right)_p.$$

So if $\{e_\mu\}$ is orthonormal, then at p

$$\left(\frac{\partial}{\partial x^\mu}\right)_p$$

forms an orthonormal basis.

2   Curvature

2.1   Parallel Transport

Suppose $\lambda : I \to M$ is a curve with tangent vector $\dot{\lambda}(t)$. We say that a tensor field T is *parallelly transported* or *propagated* along λ if

$$\nabla_{\dot{\lambda}} T = 0$$

on λ. In particular, if λ is an APG, then $\dot{\lambda}$ is parallelly transported along λ.

A parallelly propagated tensor is determined everywhere on λ by its value at one point.

Example 2.1.

If T is a $(1,1)$ tensor, then in coordinates, parallel propagation becomes

$$0 = \frac{\mathrm{d}x^{\mu}}{\mathrm{d}t} T^{\nu}{}_{\sigma;\mu} = \frac{\mathrm{d}x^{\mu}}{\mathrm{d}t}(T^{\nu}{}_{\sigma,\mu} + \Gamma_{\rho}{}^{\nu}{}_{\mu} T^{\rho}{}_{\sigma} - \Gamma_{\sigma}{}^{\rho}{}_{\mu} T^{\nu}{}_{\rho}).$$

But

$$T^{\nu}{}_{\sigma,\mu} \frac{\mathrm{d}x^{\mu}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(T^{\nu}{}_{\sigma}),$$

so

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} T^{\nu}{}_{\sigma} + (\Gamma_{\rho}{}^{\nu}{}_{\mu} T^{\rho}{}_{\sigma} - \Gamma_{\sigma}{}^{\rho}{}_{\mu} T^{\nu}{}_{\rho}) \frac{\mathrm{d}x^{\mu}}{\mathrm{d}t}.$$

This is a first-order linear ODE for $T^{\nu}{}_{\sigma}(x(t))$, so ODE theory gives us a unique solution, once $T^{\nu}{}_{\sigma}(x(0))$ is specified.

Note that parallel transport along a curve from p to q gives an isomorphism between tensors at p and q, however this isomorphism depends on the choice of curve in general.

2.2   The Riemann Tensor

The Riemann tensor captures the extent to which parallel transport depends on the curve.

Lemma 2.1. *Given X, Y, Z smooth vector fields, and ∇ a connection, define*

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

Then we claim

$$(R(X,Y)Z)^a = R^a{}_{bcd} X^c Y^d Z^b,$$

for a $(1,3)$-tensor $R^a{}_{bcd}$, the Riemann tensor.

Proof: Suppose f is a smooth function. Then,

$$
\begin{aligned}
R(fX,Y)Z &= \nabla_{fX}\nabla_Y Z - \nabla_Y\nabla_{fX}Z - \nabla_{[fX,Y]}Z \\
&= f\nabla_X\nabla_Y Z - \nabla_Y(f\nabla_X Z) - \nabla_{f[X,Y]-Y(f)X}Z \\
&= f\nabla_X\nabla_Y Z - f\nabla_Y\nabla_X Z - Y(f)\nabla_X Z - f\nabla_{[X,Y]}Z + Y(f)\nabla_X Z \\
&= fR(X,Y)Z.
\end{aligned}
$$

Since $R(X,Y)Z = -R(Y,X)Z$, we have $R(X,fY)Z = fR(X,Y)Z$. Similarly, we can check $R(X,Y)(fZ) = fR(X,Y),Z$.

Now suppose we pick a basis $\{e_\mu\}$ with dual basis $\{f^\mu\}$. Then

$$
\begin{aligned}
R(X,Y)Z &= R(X^\rho e_\rho, Y^\sigma e_\sigma)(Z^\nu e_\nu) = X^\rho Y^\sigma Z^\nu R(e_\rho, e_\sigma)e_\nu \\
&= (R^\mu{}_{\nu\rho\sigma}X^\rho Y^\sigma Z^\nu)e_\mu,
\end{aligned}
$$

where $R^\mu{}_{\nu\rho\sigma} = f^\mu(R(e_\rho, e_\sigma)e_\nu)$ are the components of $R^a{}_{bcd}$ in this basis. Since this results holds in one basis, it hold in all, since the object was defined tensorially.

In a coordinate basis $e_\mu = \frac{\partial}{\partial x^\mu}$, with $[e_\mu, e_\nu] = 0$, we have

$$
\begin{aligned}
R(e_\rho, e_\sigma)e_\nu &= \nabla_{e_\rho}(\nabla_{e_\sigma}e_\nu) - \nabla_{e_\sigma}(\nabla_{e_\rho}e_\nu) - \nabla_{e_\rho}(\Gamma_\nu{}^\tau{}_\sigma e_\tau) - \nabla_{e_\sigma}(\Gamma_\nu{}^\tau{}_\rho e_\tau) \\
&= \partial_\rho(\Gamma_\nu{}^\tau{}_\sigma)e_\tau + \Gamma_\nu{}^\tau{}_\sigma\Gamma_\tau{}^\mu{}_\rho e_\mu - \partial_\sigma(\Gamma_\nu{}^\tau{}_\rho)e_\tau - \Gamma_\nu{}^\tau{}_\rho\Gamma_\tau{}^\mu{}_\sigma e_\mu.
\end{aligned}
$$

Hence the components are

$$
R^\mu{}_{\nu\rho\sigma} = \partial_\rho(\Gamma_\nu{}^\mu{}_\sigma) - \partial_\sigma(\Gamma_\nu{}^\mu{}_\rho) + \Gamma_\nu{}^\tau{}_\sigma\Gamma_\tau{}^\mu{}_\rho - \Gamma_\nu{}^\tau{}_\rho\Gamma_\tau{}^\mu{}_\sigma.
$$

In normal coordinates, we can drop the last two terms, since the connection is zero at p. However the derivative is not 0.

Example 2.2.

For the Levi-Civita connection of Minkowski space in an intertial frame, $\Gamma_\mu{}^\nu{}_\sigma = 0$. Hence in these coordinates $R^\mu{}_{\nu\sigma\tau} = 0$. Thus

$$
R^a{}_{bcd} = 0.
$$

Such a connection is called *flat*.

Conversely, for a Lorentzian spacetime with flat Levi-Civita connection, we can locally find coordinates such that $g_{\mu\nu} = \mathrm{diag}(-1, 1, \ldots, 1)$.

Here is a note of caution:

$$(\nabla_X \nabla_Y Z)^c = X^a \nabla_a (Y^b \nabla_b Z^c) \neq X^a Y^b \nabla_a \nabla_b Z^c.$$

Hence

$$
\begin{aligned}
(R(X,Y)Z)^c &= X^a \nabla_a(Y^b \nabla_b Z^c) - Y^a \nabla_a(X^b \nabla_b Z^c) - [X,Y]^b \nabla_b Z^c \\
&= X^a Y^b \nabla_a \nabla_b Z^c - Y^a X^b \nabla_a \nabla_b Z^c + (\nabla_X Y - \nabla_Y X - [X,Y])^b \nabla_b Z^c \\
&= X^a Y^b R^c{}_{dab} Z^d.
\end{aligned}
$$

So if ∇ is torsion free,

$$\nabla_a \nabla_b Z^c - \nabla_b \nabla_a Z^c = R^c{}_{dab} Z^d.$$

This is the *Ricci identity*. We can generalise this to an expression for

$$\nabla_{[a} \nabla_{b]} T^{c_1 \ldots c_r}{}_{d_1 \ldots d_s}.$$

We can construct a new tensor from $R^a{}_{bcd}$ by contraction.

Definition 2.1. The *Ricci tensor* is the $(0,2)$-tensor defined by

$$R_{ab} = R^c{}_{acb}.$$

Suppose X, Y are vector fields satisfying $[X,Y] = 0$.

Insert picture of vector curves going from $A \to B \to C \to D \to A$.

Consider going from A to B flowing a distance ε along the integral curve of X, then B to C distance ε along the integral curve of Y, then C to D along X, and D to A along Y.

Since $[X,Y] = 0$, we return to the starting location. We claim the following of a propagated tensor:

Proposition 2.1. *If Z is parallelly transported around $ABCD$ to a vector Z', then*

$$(Z - Z')^\mu = \varepsilon^2 R^\mu{}_{\nu\rho\sigma} Z^\nu X^\rho Y^\sigma + \mathcal{O}(\varepsilon^3).$$

2.3   Geodesic Deviation

Let ∇ be a symmetric connection, and suppose $\lambda : I \to M$ is an affinely parametrized geodesic through p. We can pick normal coordinates centred at p, such at λ is given by $t \mapsto (t, 0, \ldots, 0)$. Suppose we start a geodesic with

$$x_s^\mu(0) = s x_0^\mu, \qquad \dot{x}_s^\mu(0) = s x_1^\mu + (1, 0, \ldots, 0).$$

Then we find that

$$x_s^\mu(t) = x^\mu(s, t) = (t, 0, \ldots, 0) + s Y^\mu(t) + \mathcal{O}(s^2),$$

where

$$Y^\mu(t) = \left. \frac{\partial x^\mu}{\partial s} \right|_{s=0}$$

are the components of a vector field along λ measuring the (infinitesimal) deviation of the geodesics. We have, by the geodesic equation,

$$\frac{\partial^2 x^\mu}{\partial t^2} + \Gamma_{\nu}{}^{\mu}{}_{\sigma}(x^\mu(s, t)) \frac{\partial x^\nu}{\partial t} \frac{\partial x^\mu}{\partial t} = 0.$$

Take the derivative with respect to s, at $s = 0$. Then,

$$\frac{\partial^2 Y^\mu}{\partial t^2} + \partial_\mu (\Gamma_{\nu}{}^{\mu}{}_{\sigma})_{s=0} T^\nu T^\sigma Y^\rho + Z \Gamma_{\rho}{}^{\mu}{}_{\sigma} \frac{\partial}{\partial t} Y^\rho T^\sigma = 0,$$

which gives

$$T^\nu (T^\sigma Y^\mu{}_{,\sigma})_{,\nu} + \partial_\rho (\Gamma_{\nu}{}^{\mu}{}_{\sigma})_{s=0} T^\nu T^\sigma Y^\rho + Z \Gamma_{\rho}{}^{\mu}{}_{\sigma} \frac{\partial Y^\rho}{\partial t} T^\sigma = 0.$$

At $p = 0$, $\Gamma = 0$ so

$$T^\nu (T^\sigma Y^\mu{}_{;\sigma} - \Gamma_{\rho}{}^{\mu}{}_{\sigma} T^\sigma Y^\rho)_{;\nu} + (\partial_\rho \Gamma_{\nu}{}^{\mu}{}_{\sigma})_p T^\nu T^\sigma Y^\rho = 0.$$

We can write this as

$$T^\nu (T^\sigma Y^\mu{}_{;\sigma})_{;\nu} + (\partial_\rho \Gamma_{\nu}{}^{\mu}{}_{\sigma} - \partial_\nu \Gamma_{\rho}{}^{\mu}{}_{\sigma}) T^\nu T^\sigma Y^\rho = 0$$
$$\implies (\nabla_T \nabla_T Y)^\mu + R^\mu{}_{\omega\rho\nu} T^\nu T^\sigma Y^\rho = 0$$
$$\implies \nabla_T \nabla_T Y + R(Y, T) T = 0.$$

This is the Jacobi equation for geodesic deviation.

2.4   Riemann Tensor Symmetries

From the definition, it is clear that

$$R^a{}_{bcd} = -R^a{}_{bdc} \iff R^a{}_{b(cd)} = 0.$$

Proposition 2.2. *If ∇ is torsion free, then*

$$R^a{}_{[bcd]} = 0.$$

Proof: Fix $p \in M$, and choose normal coordinates at p, and work in the coordinate basis. Then Γ vanishes, and $\Gamma_\mu{}^\sigma{}_\nu = \Gamma_\nu{}^\sigma{}_\mu$ everywhere. Then,

$$R^\mu{}_{\nu\rho\sigma}|_p = \partial_\rho(\Gamma_\nu{}^\mu{}_\sigma)|_p - \partial_\sigma(\Gamma_\nu{}^\mu{}_\rho)|_p$$
$$\implies R^\mu{}_{[\nu\rho\sigma]}|_p| = 0,$$

as by symmetry

$$\partial_\rho \Gamma_\nu{}^\mu{}_\sigma|_p = \partial_\rho \Gamma_\sigma{}^\mu{}_\nu|_p$$

Here p is arbitrary, so $R^\mu{}_{[\nu\rho\sigma]} = 0$ everywhere.

Proposition 2.3. *If ∇ is torsion free, then the Bianchi identity holds:*

$$R^a{}_{b[cd;e]} = 0.$$

Proof: Choose normal coordinates. Then

$$R^\mu{}_{\nu\rho\sigma;\tau}|_p = R^\mu{}_{\nu\rho\sigma,\tau}|_p.$$

Schematically, $R \sim \partial\Gamma + \Gamma^2$, so $\partial R \sim \partial^2\Gamma + \partial\Gamma \cdot \Gamma$, and since $\Gamma|_p = 0$, we deduce

$$R^\mu{}_{\nu\rho\sigma,\tau}|_p = \partial_\tau\partial_\rho\Gamma_\nu{}^\mu{}_\sigma|_p - \partial_\tau\partial_\sigma\Gamma_\nu{}^\mu{}_\rho|_p.$$

By symmetry of mixed partial derivatives, we see

$$R^\mu{}_{\nu[\rho\sigma;\tau]}|_p = 0.$$

Since p was arbitrary, the result follows.

Suppose ∇ is the Levi-Civita connection of a manifold with a metric g. We can lower an index with g_{ab}, and consider R_{abcd}.

Proposition 2.4. R_{abcd} *satisfies*

$$R_{abcd} = R_{cdab} \implies R_{(ab)cd} = 0.$$

Proof: Pick normal coordinates at p so that $\partial_\mu g_{\nu\rho}|_p = 0$. We notice that

$$0 = \partial_\mu \delta^\nu{}_\sigma|_p = \partial_\mu(g^{\nu\tau}g_{\tau\sigma}) = (\partial_\mu g^{\nu\tau})g_{\tau\sigma}|_p.$$

The derivative $\partial_\mu^{\nu\tau}|_p = 0$, hence

$$\partial_\rho(\Gamma_\nu{}^\mu{}_\sigma)|_p = \partial_\rho \left(\frac{1}{2} g^{\mu\tau}(g_{\tau\sigma,\nu} + g_{\nu\tau,\sigma} - g_{\nu\sigma,\tau}) \right)\Big|_p$$

$$= \frac{1}{2} g^{\mu\tau}(g_{\tau\sigma,\nu\rho} + g_{\nu\tau,\sigma\rho} - g_{\nu\sigma,\tau\rho})|_p.$$

We have then that

$$R_{\mu\nu\rho\sigma}|_p = g_{\mu\kappa}(\partial_\rho\Gamma_\nu{}^\kappa{}_\sigma - \partial_\sigma\Gamma_\nu{}^\kappa{}_\rho)|_p$$

$$= \frac{1}{2}(g_{\mu\sigma,\nu\rho} + g_{\nu\rho,\mu\sigma} - g_{\nu\sigma,\mu\rho} - g_{\mu\rho,\nu\sigma})|_p.$$

This satisfies

$$R_{\mu\nu\rho\sigma}|_p = R_{\rho\sigma\mu\nu}|_p,$$

hence this is true everywhere.

Corollary 2.1. *The Ricci tensor is symmetric:*

$$R_{ab} = R_{ba}.$$

Proof: We have

$$R_{ab} = R^c{}_{acb} = g^{cd}R_{cadb} = g^{cd}R_{dbca} = R_{ba}.$$

Definition 2.2. The *Ricci scalar* (scalar curvature) is

$$R = R^a{}_a = g^{ab}R_{ab}.$$

The *Einstein tensor* is

$$G_{ab} = R_{ab} - \frac{1}{2}g_{ab}R.$$

We can show that the Bianchi identity implies

$$\nabla_a G^a_{\ b} = 0.$$

This is the contracted Bianchi identity.

3   Diffeomorphisms and the Lie Derivative

Suppose $\varphi : M \to N$ is a smooth map. Then φ induces maps between corresponding vector or covector bundles.

Definition 3.1. Given $f : N \to \mathbb{R}$, the *pullback* of f by φ is the map $\varphi^* f : M \to \mathbb{R}$ given by

$$\varphi^* f(p) = f(\varphi(p)).$$

Definition 3.2. Given $X \in T_p M$, we define the *push forward* of X by φ, say $\varphi_* X \in T_{\varphi(p)} N$ as follows:

Let $\lambda : I \to M$ be a curve with $\lambda(0) = p$, $\dot\lambda(0) = X$. Then set $\tilde\lambda = \varphi \circ \lambda$. $\tilde\lambda : I \to N$ gives a curve in N with $\tilde\lambda(0) = \varphi(p)$.

We set $\varphi_* X = \dot{\tilde\lambda}(0)$.

Note that if $f : N \to \mathbb{R}$, then

$$\varphi_* X(f) = \frac{\mathrm{d}}{\mathrm{d}t}(f \circ \tilde\lambda(t))|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}(f \circ \varphi \circ \lambda(t))|_{t=0}$$
$$= \frac{\mathrm{d}}{\mathrm{d}t}(\varphi^* f \circ \lambda(t))|_{t=0} = X(\varphi^* f).$$

We can show that if x^μ are coordinates on M near p, and y^α are coordinates on N near $\varphi(p)$, then φ gives a map $y^\alpha(x^\mu)$, and we can show in a coordinate basis that

$$(\varphi_* X)^\alpha = \left(\frac{\partial y^\alpha}{\partial x^\mu} \right)_p X^\mu.$$

On the contangent bundle, we go backwards.

Definition 3.3. If $\eta \in T^*_{\varphi(p)} N$, then the *pullback* of η, $\varphi^* \eta \in T^*_p M$ is defined by

$$\varphi^* \eta(X) = \eta(\varphi_* X),$$

for all $X \in T_p M$.

Note that if $f : N \to \mathbb{R}$, then

$$\varphi^*(\mathrm{d}f)[X] = \mathrm{d}f[\varphi_* X] = \varphi_* X(f) = X(\varphi^* f)$$
$$= \mathrm{d}(\varphi^* f)[X],$$

hence $\varphi^* \, \mathrm{d}f = \mathrm{d}(\varphi^* f)$. We can also show that in a basis,

$$(\varphi^* \eta)_\mu = \left(\frac{\partial y^\alpha}{\partial x^\mu} \right)_p \eta_\alpha.$$

We can extend the pullback to map a $(0, s)$-tensor T at $\varphi(p) \in N$, to a $(0, s)$-tensor $\varphi^* T$ at $p \in M$, by

$$\varphi^* T(X_1, \ldots, X_s) = T(\varphi_* X_1, \ldots, \varphi_* X_s)$$

for all $X_i \in T_p M$. Similarly, we can push forward a $(s, 0)$-tensor S at $p \in M$ to a $(S, 0)$-tensor $\varphi_* S$ at $\varphi(p)$ by

$$\varphi_* S(\eta_1, \ldots, \eta_s) = S(\varphi^* \eta_1, \ldots, \varphi^* \eta_s),$$

for all $\eta_1, \ldots, \eta_s \in T^*_{\varphi(p)} N$.

If $\varphi : M \to N$ has the property that $\varphi_* : T_p M \to T_{\varphi(p)} N$ is injective, we say that φ is an *immersion*. In particular $\dim N \geq \dim M$.

If N is a manifold with metric g and $\varphi : M \to N$ is an immersion, we can consider $\varphi^* g$. If g is Riemannian, then $\varphi^* g$ is non-degenerate and positive definite, so it defines a metric on M, the *induced metric*.

> **Example 3.1.**
>
> Let $(N, g) = (\mathbb{R}^3, \delta)$, and $M = S^2$. Let φ be the map taking a point on S^2 with spherical coordinates (θ, ϕ) to
>
> $$(x^1, x^2, x^3) = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta).$$
>
> Then
> $$\varphi^*((\mathrm{d}x^1)^2 + (\mathrm{d}x^2)^2 + (\mathrm{d}x^3)^2) = \mathrm{d}\theta^2 + \sin^2 \theta \, \mathrm{d}\phi^2.$$

If φ is an immersion and (N, g) is Lorentzian, then $\varphi^* g$ is not in general a metric on M. There are three important cases:

- $\varphi^* g$ is a Riemannian metric $\implies \varphi(M)$ is spacelike.
- $\varphi^* g$ is a Lorentian metric $\implies \varphi(M)$ is timelike.
- $\varphi^* g$ is everywhere degenerate $\implies \varphi(M)$ is null.

Recall that $\varphi : M \to N$ is a *diffeomorphism* if it is bijective with a smooth inverse. If we have a diffeomorphism, we can push forward a general (r, s)-tensor at p to an (r, s)-tensor at $\varphi(p)$ by'

$$\varphi_* T(\eta^1, \ldots, \eta^r, X_1, \ldots, X_s) = T(\varphi^* \eta^1, \ldots, \varphi^* \eta^r, \varphi_*^{-1} X_1, \ldots, \varphi_*^{-1} X_s),$$

for all $\eta_i \in T^*_{\varphi(p)} N$, and $X_i \in T_{\varphi(p)} N$. We can similarly define a pullback by $\psi_*^{-1} = \varphi^*$.

If M, N are diffeomorphic, we often do not distinguish between them, and can think of $\varphi : M \to M$.

We say that a diffeomorphism $\varphi : M \to M$ is a symmetry of T if $\varphi_* T = T$. If T is the metric, we say φ is an *isometry*.

Example 3.2.

In a Minkowski space with an interial frame,

$$\varphi(x^0, x^1, \ldots, x^n) = (x^0 + 1, x^1, \ldots, x^n)$$

is a symmetry of g.

An important class of diffeomorphisms are those that are generated by a vector field. If X is a smooth vector field, we associate to each point $p \in M$ the point $\varphi_t^X(p) \in M$ given by flowing a parameter distance t along the integral curve of X starting at p.

Suppose $\varphi_t^X(p)$ is well defined for all $t \in I \subseteq \mathbb{R}$ for each $p \in M$. Then $\varphi_t^X : M \to M$ is a diffeomorphism for all $t \in I$. Further,

- If $t, s, t + s \in I$, then $\varphi_t^X \circ \varphi_s^X = \varphi_{t+s}^X$, and $\varphi_0^X = \text{id}$.

- If $I = \mathbb{R}$, this gives $\{\varphi_t^X\}$ the structure of a one-parameter abelian group.

- If φ_t is any smooth family of diffeomorphisms satisfying the group property, we can define a vector field by

$$X_p = \frac{\mathrm{d}}{\mathrm{d}t}(\varphi_t(p))\bigg|_{t=0}.$$

  Then $\varphi_t = \varphi_t^X$.

We can use φ_t^X to compare tensors at different points. As $t \to 0$, this gives a new notion of derivative.

3.1   The Lie Derivative

Suppose $\varphi_t^X : M \to M$ is the smooth one-parameter family of diffeomorphisms generated by a vector field X.

Definition 3.4. For a tensor field T, the *Lie derivative* of T with respect to X is

$$(\mathcal{L}_X T)_p = \lim_{t \to 0} \frac{((\varphi_t^X)^* T)_p - T_p}{t}.$$

This is the pullback of the tensor field minus the original tensor field.

It is easy to see that for constants α, β and (r, s)-tensors S, T,

$$\mathcal{L}_X(\alpha S + \beta T) = \alpha \mathcal{L}_X S + \beta \mathcal{L}_X T.$$

To see how $\mathcal{L}_X$ acts in components, it is useful to construct coordinates adapted to X. Near p, we can construct an $(n-1)$-surface Σ which is transverse to X. Pick coordinates x^i on Σ and assign the coordinate (t, x^i) to the point a parameter distance t along the integral curve of X starting at x^i on Σ.

In these coordinates $X = \frac{\partial}{\partial t}$, and $\varphi_t^X(\tau, x^i) = (\tau + t, x^i)$. So if $y^\mu = \varphi_t^X(x^\mu)$, then

$$\frac{\partial y^\mu}{\partial x^\nu} = \delta^\mu{}_\nu.$$

Hence we find

$$[(\varphi_t^X)^* T]^{\mu_1 \dots \mu_r}{}_{\nu_1 \dots \nu_s}|_{(t, x^i)} = T^{\mu_1 \dots \mu_r}{}_{\nu_1 \dots \nu_s}|_{(\tau + t, x^i)}.$$

Thus we find

$$(\mathcal{L}_X T)^{\mu_1 \dots \mu_r}{}_{\nu_1 \dots \nu_s}|_p = \frac{\partial T^{\mu_1 \dots \mu_r}}{\partial t}{}_{\nu_1 \dots \nu_s}|_p.$$

So in these coordinates, $\mathcal{L}_X$ acts on components by $\frac{\partial}{\partial t}$. In particular, we immediately see that

$$\mathcal{L}_X(S \otimes T) = (\mathcal{L}_X S) \otimes T + S \otimes \mathcal{L}_X T.$$

So $\mathcal{L}_X$ commutes with contraction. To write $\mathcal{L}_X$ in a coordinate free fashion, we can simply seek a basis independent expression that agrees with $\mathcal{L}_X$ in these coordinates.

For example, for a function

$$\mathcal{L}_X f = \frac{\partial f}{\partial t} = X(f)$$

in these coordinates. For a vector field Y, we observe that

$$\frac{\partial Y^\mu}{\partial t} = X^\sigma \frac{\partial}{\partial x^\sigma}(Y^\mu) - Y^\sigma \frac{\partial}{\partial x^\sigma} X^\mu = [X, Y]^\mu.$$

This is because in the basis (t, x^i), $X = \partial/\partial t$, hence $X^t = 1$ is constant so has derivative 0. So $\mathcal{L}_X Y = [X, Y]$. For a covector field, we can show

$$(\mathcal{L}_X \omega)_\mu = X^\sigma \partial_\sigma \omega_\mu + \omega_\sigma \partial_\mu X^\sigma.$$

If ∇ is torsion-free, then

$$(\mathcal{L}_X \omega)_a = X^b \nabla_b \omega_a + \omega_b \nabla_a X^b.$$

If g_{ab} is a metric tensor, and ∇ is Levi-Civita, then

$$(\mathcal{L}_X g)_{ab} = \nabla_a X_b + \nabla_b X_a.$$

If φ_t^X is a one-parameter family of isometries for a manifold with metric g, then $\mathcal{L}_X g = 0$. Conversely if $\mathcal{L}_X g = 0$, then X generates a one-parameter family of isometries.

Definition 3.5. A vector field K satisfies $\mathcal{L}_K g = 0$ is called a *Killing vector*. It satisfies Killing's equation:

$$\nabla_a K_b + \nabla_b K_a = 0.$$

Lemma 3.1. *Suppose K is Killing and $\lambda : T \to M$ is a geodesic of the Levi-Civita connection then $g_{ab} \dot{\lambda}^a K^b$ is constant along λ.*

Proof: We have

$$\frac{\mathrm{d}}{\mathrm{d}t}(K_b \dot{\lambda}^b) = \dot{\lambda}^a \nabla_a (K_b \dot{\lambda}^b) = (\nabla_{(a} K_{b)}) \dot{\lambda}^a \dot{\lambda}^b + K_b \dot{\lambda}^a \nabla_a \dot{\lambda}^b$$

$$= 0,$$

for the first part as K is Killing and the latter as $\nabla_\lambda \lambda = 0$ for a geodesic.

4   Physics in Curved Spacetimes

4.1   Minkowski Space

We review physical theories in Minkowski $\mathbb{R}^{1+3}$ equipped with

$$\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1).$$

We set $c = 1$.

1. The first is the Klein-Gordon equation, where

$$\partial_\mu \partial^\mu \phi - m^2 \phi = 0.$$

Note that in inertial coordinates, $\partial_\mu = \nabla_\mu$, so we can write this in a covariant manner as

$$\nabla_a \nabla^a \phi - m^2 \phi = 0.$$

Associated to this is the energy-momentum tensor

$$T_{\mu\nu} = \partial_\mu \phi \partial_\nu \phi - \frac{1}{2} \eta_{\mu\nu} (\partial_\sigma \phi \partial^\sigma \phi + m^2 \phi^2).$$

Covariantly, this becomes

$$T_{ab} = \nabla_a \phi \nabla_b \phi - \frac{1}{2} \eta_{ab} (\nabla_c \phi \nabla^c \phi + m^2 \phi^2).$$

We can check that $T_{ab} = T_{ba}$, and $\nabla_a T^a{}_b = 0$, giving a conserved tensor.

2. Maxwell's equations. The Maxwell field is an anti-symmetric 2-tensor $F_{\mu\nu} = -F_{\nu\mu}$, where $F_{0i} = E_i$ and $F_{ij} = \varepsilon_{ijk} B_k$ for E and B the electric and magnetic fields.

If j_μ is the charge current density, then Maxwell's equations are

$$\partial_\mu F^\mu{}_\nu = 4\pi j_\nu, \qquad \partial_{[\mu} F_{\nu\sigma]} = 0.$$

Making this covariant,

$$\nabla_a F^a{}_c = 4\pi j_c, \qquad \nabla_{[a} F_{bc]} = 0.$$

Associated to this is the energy-momentum tensor

$$T_{\mu\nu} = F_\mu{}^\delta F_{\nu\delta} - \frac{1}{4} \eta_{\mu\nu} F_{\sigma\tau} F^{\sigma\tau}.$$

Once again, promoting the indices,

$$T_{ab} = F_a{}^c F_{bc} - \frac{1}{4} \eta_{ab} F_{cd} F^{cd}.$$

We can check that $T_{ab} = T_{ba}$ and $\nabla_a T^a{}_b = 0$ for the source-free Maxwell equation.

3. A perfect fluid is described by a local velocity field U^μ satisfying $U^\mu U_\mu = -1$, together with a pressure p and density ρ. They satisfy the first law of thermodynamics,

$$U^\mu \partial_\mu \rho + (\rho + p)\partial_\mu U^\mu = 0 \implies U^a \nabla_a \rho + (\rho + p)\nabla_a U^a = 0,$$

and Euler's equations

$$(\rho+p)U^\nu \partial_\nu U^\mu + \partial^\mu p + U^\mu U^\nu \partial_\nu p = 0 \implies (\rho+p)U^a \nabla_a U^b + \nabla^b p + U^b U^a \nabla_a p = 0.$$

Associated to this is the energy-momentum tensor

$$T_{\mu\nu} = (\rho + p)U_\mu U_\nu + p\eta_{\mu\nu} \implies T_{ab} = (\rho + p)U_a U_b + p\eta_{ab}.$$

Again we claim $T_{ab} = T_{ba}$, $\nabla_a T^a{}_b = 0$.

Noteice that in all cases we can promote the Minkowski metric η to a general Lorentzian metric g, and we can take ∇ to be the Levi-Civita connection.

Consider normal coordinates. We see that coordinates exist near any point $q \in M$, such that the physics described is approximately Minkowskian, with correction of the order of the curvature.

4.2   General Relativity

In Einstein's theory of general relativity, we postulate that spacetime is a four-dimensional Lorentzian manifold (M, g). We also require that any matter model consists of:

- some matter fields ϕ^A,
- equations of motion for ϕ^A,

which are expressed geometrically in terms of g and its derivatives, such as ∇, R etc.

An *energy-momentum tensor* T_{ab} depending on ϕ^A should satisfy

$$T_{ab} = T_{ba}, \qquad \nabla_a T^a{}_b = 0.$$

The matter should reduce to a non-gravitational theory when (M, g) is fixed to be Minkowskian, and the metric g should satisfy the *Einstein equation*:

$$R_{ab} - \frac{R}{2}g_{ab} + \Lambda g_{ab} = 8\pi G T_{ab}.$$

Moreover Λ is the cosmological constant, which is experimentally positive but small, and G is Newton's constant.

The Einstein equations, together with the equations of motion for ϕ^A, constitute a *coupled system* which must be solved simultaneously.

We need to fix the following *geodesic postulate*:

Free test particles move along timelike or null geodesics if they have nonzero or zero rest mass, respectively.

4.3   Gauge Freedom

Consider the Maxwell equations with no sources, so

$$\partial_\mu F^\mu{}_\nu = 0, \qquad \partial_{[\mu} F_{\nu\sigma]} = 0.$$

A standard approach to solve the second equation is to introduce the *gauge potential* A_μ such that
$$F_{\mu\nu} = \partial_{[\mu} A_{\nu]}.$$

Then the first equation becomes

$$\partial_\mu \partial^\mu A_\nu - \partial_\mu \partial_\nu A^\mu = 0.$$

We want to solve this given some data at $\{x^0 = 0\} = \Sigma$. However this equation doesn't lead to a nice evolution problem. If $\chi \in C^\infty(\mathbb{R}^{1+3})$ which vanishes near Σ, then
$$\tilde{A}_\mu = A_\mu + \partial_\mu \chi$$

will also solve the equation and satisfy

$$\partial_{[\mu} \tilde{A}_{\nu]} = \partial_{[\mu} A_{\nu]}.$$

To resolve this, we can fix a gauge g if we assume that $\partial_\mu A^\mu = 0$, then the equation for A becomes $\partial_\mu \partial^\mu A_\nu = 0$. This is a wave equation for each component of A_ν, with a unique solution given $A_\nu|_{t=0}$, and $\partial_0 A_\nu|_{t=0}$. There is no nontrivial χ such that $\tilde{A}_\nu$ also satisfies $\partial_\mu \tilde{A}^\mu = 0$, and the same initial conditions.

If we solve the wave equation $\Box A_\nu = 0$ such that $\partial_\mu A^\mu$ and $\partial_0(\partial_\mu A^\mu)$ at $t = 0$ is 0, then we find $\partial_\mu A^\mu = 0$ and so our solution solves Maxwell's equations.

4.4   Gauge Freedom for Einstein

If (M, g) solves the Einstein equation with EM tensor T, and $\varphi : M \to M$ is a diffeomorphism, then $\varphi^* g$ also solves the Einstein equation with EM tensor $\varphi^* T$. At the local level, this arises as the coordinate independence of Einstein's equations.

In order to solve the Einstein equations, we need to find a way to fix the coordinates. There are several approaches, we consider wave or harmonic gauge.

Lemma 4.1. *In any local coordinate system,*

$$R_{\rho\delta\mu\nu} = \frac{1}{2}(g_{\rho\nu,\mu\delta} + g_{\delta\mu,\nu\rho} - g_{\rho\mu,\nu\delta} - g_{\delta\nu,\mu\gamma}) - \Gamma_{\mu\lambda\rho}\Gamma_{\nu}{}^{\lambda}{}_{\delta} + \Gamma_{\nu\lambda\rho} + \Gamma_{\mu}{}^{\lambda}{}_{\delta},$$

and also

$$R_{\delta\nu} = -\frac{1}{2}g^{\mu\rho}g_{\delta\nu,\mu\rho} + \frac{1}{2}\partial_\delta \Gamma_{\mu\nu}{}^{\mu} + \frac{1}{2}\partial_\nu \Gamma_{\mu\delta}{}^{\mu} - \Gamma_{\mu\lambda}{}^{\lambda}\Gamma_{\nu}{}^{\lambda}{}_{\delta}$$
$$+ \Gamma_{\lambda\tau\nu}\Gamma^{\lambda\tau}{}_{\delta} + \Gamma_{\lambda\tau\nu}\Gamma^{\tau}{}_{\delta}{}^{\lambda} + \Gamma_{\lambda\tau\delta}\Gamma^{\tau}{}_{\nu}{}^{\lambda}.$$

The proof of this is omitted; it is on the Moodle page.

This form of the Ricci tensor is well-adapted to wave or harmonic coordinates. Suppose we choose a system of coordinates $\{x^\nu\}$ for $\nu = 0, 1, 2, 3$ which satisfies the wave equation

$$\nabla_\mu \nabla^\mu x^\nu = 0.$$

Hence we get

$$0 = \nabla^\mu(\partial_\mu x^\nu) = \partial^\mu \partial_\mu x^\nu + \Gamma^{\mu\delta}{}_\mu \partial_\delta x^\nu.$$

Hence

$$\Gamma^{\mu\nu}{}_\mu = 0 \iff \frac{1}{2}g^{\mu\delta}\left(g_{\mu\kappa,\delta} - \frac{1}{2}g_{\mu\delta,\kappa}\right) = 0.$$

For such coordinates,

$$R_{\delta\nu} = -\frac{1}{2}g^{\mu\rho}g_{\delta\nu,\mu\rho} + (\Gamma_{\lambda\tau\nu}\Gamma^{\lambda\tau}{}_{\delta} + \Gamma_{\lambda\tau\nu}\Gamma^{\tau}{}_{\delta}{}^{\lambda} + \Gamma_{\lambda\tau\delta}\Gamma^{\tau}{}_{\nu}{}^{\lambda}).$$

The latter part is quadratic. So $R_{\delta\nu} = 0$ reduces to a system of a nonlinear wave equation in their gauge. We can solve this (locally) given initial data. Further, we can show that if the gauge condition is initially satisfied, then it will remain true in spacetime.

4.5   Linearized Theory

Suppose we are in a situation where the gravitational field is weak. We expect to be able to describe the metric as a perturbation of Minkowski, so

$$g_{\mu\nu} = \eta_{\mu\nu} + \varepsilon h_{\mu\nu},$$

where $\eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$, and $\varepsilon \ll 1$ is a small parameter. If the metric has this form, then we say that we are in *almost inertial* coordinates. One can check that

$$g^{\mu\nu} = \eta^{\mu\nu} - \varepsilon h^{\mu\nu},$$

where $h^{\mu\nu} = \eta^{\mu\sigma} h_{\sigma\tau} \eta^{\tau\nu}$. Suppose our metric is also in the wave gauge. Then

$$0 = g^{\mu\delta}\left(g_{\mu\kappa,\delta} - \frac{1}{2}g_{\mu\delta,\kappa}\right) = \varepsilon\partial^{\mu}\left(h_{\mu\kappa} - \frac{1}{2}\eta_{\mu\kappa}h\right),$$

where $h = \eta^{\mu\nu}h_{\mu\nu}$. Using our expression for the Ricci tensor,

$$R_{\mu\nu} = -\frac{\varepsilon}{2}\eta^{\delta\tau}\partial_{\delta}\partial_{\tau}h_{\mu\nu}.$$

In order for the Einstein equation to holds, we should take

$$\mathbf{T}_{\mu\nu} = \varepsilon T_{\mu\nu}.$$

Then to order ε in the Einstein equation,

$$-\frac{1}{2}\Box h_{\mu\nu} + \frac{1}{4}\eta_{\mu\nu}\eta^{\delta\tau}\partial_{\delta}\partial_{\tau}h = 8\pi G T_{\mu\nu}$$
$$\implies \Box \hbar_{\mu\nu} = -16\pi G T_{\mu\nu},$$
$$\partial_{\mu}\hbar^{\mu}{}_{\nu} = 0,$$

where $\hbar_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}h\eta_{\mu\nu}$, so $h_{\mu\nu} = \hbar_{\mu\nu} - \frac{1}{2}\hbar_{\mu\nu}$.

These are the linearised Einstein equations in wave (harmonic, Lorenz, de Donder) gauge. We can solve these given initial data at $\{t = 0\} = \Sigma$. If the data satisfies

$$\partial_{\mu}\hbar^{\mu}{}_{\nu}|_{\Sigma} = \partial_{0}(\partial_{\mu}\hbar^{\mu}{}_{\nu})|_{\Sigma} = 0,$$

then since $\partial_{\mu}T^{\mu}{}_{\nu} = 0$, we have $\Box(\partial_{\mu}\hbar^{\mu}{}_{\nu}) = 0$, so the gauge holds for all times.

Suppose we have physically equivalent solutions not necessarily in wave gauge. We have seen that two such equivalent solutions are related by a diffeomorphism, in order that the diffeomorphism reduces smoothly to the identity as $\varepsilon \to 0$, should take the form Φ_{ε}^{ξ}, for ξ a vector field.

If S is any tensor field, then from the definition of the Lie derivatives,

$$(\Phi_\varepsilon^\xi)^* S = S + \varepsilon \mathcal{L}_\xi S + \mathcal{O}(\varepsilon^2).$$

Note that if $S = \mathcal{O}(\varepsilon)$, then it will be invariant under gauge transformations to order $\mathcal{O}(\varepsilon)$. Tensors vanishing on Minkowski are hence gauge invariant in linear perturbation theory. In particular, $T_{\mu\nu}$ is gauge invariant. However,

$$((\Phi_\varepsilon^\xi)^* \eta)_{\mu\nu} = \eta_{\mu\nu} + \varepsilon(\partial_\mu \xi_\nu + \partial_\nu \xi_\mu) + \mathcal{O}(\varepsilon^2),$$

where we can take $h_{\mu\nu} = \partial_\mu \xi_\nu + \partial_\nu \xi_\mu$. In particular, we change

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu,$$

which represents a linear gauge transformation. If we linearize $R_{\mu\nu}$ without fixing a gauge, we find

$$R_{\mu\nu} = \varepsilon \left(\partial^\rho \partial_{(\mu} h_{\nu)\rho} - \frac{1}{2} \partial^\rho \partial_\rho h_{\mu\nu} - \frac{1}{2} \partial_\mu \partial_\nu h \right).$$

Substituting in $h_{\mu\nu} = \partial_\mu \xi_\nu - \partial_\nu \xi_\mu$, then $R_{\mu\nu}$ vanishes, so $h_{\mu\nu}$ solves the vacuum Einstein equations. We call such solutions a *pure gauge solution*.

We can show that (exercise!) if $h'_{\mu\nu} = h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu$, then

$$\partial_\mu \bar{h}'^\mu_{\ \nu} = \partial_\mu \bar{h}^\mu_{\ \nu} + \partial_\mu \partial^\mu \xi_\nu.$$

We deduce that:

- Any linearized perturbation can be put into wave gauge by a suitable gauge transformation.

- Any pure gauge solution of the wave gauge fixed equations with $\xi_\mu = 0$ and $\partial_t \xi_\mu = 0$ at $t = 0$, vanishes everywhere.

4.6   The Newtonian Limit

We expect that if GR is a good theory of gravity, we should be able to recover Newton's theory of gravity in the limit where fields are weak, and matter is slowly moving in comparison with the speed of light.

Let us suppose that matter is modelled by a perfect fluid, with velocity field U^a, density ρ and pressure p. In all but the most extreme situations, we have $\rho \gg p$. In standard units, $p/\rho \sim v_s^2 \ll c^2$.

We choose coordinates such that $U^a = \frac{\partial}{\partial t}$ (can think of this as Lagrangian coordinates, see Mach's principle).

Note that this does not imply the fluid is at rest; distances are measured with the metric. The condition that the fluid moves non-relativistically becomes the assumption that we are in the weak field limit:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}.$$

We also want relative velocities and accelerations to not be erratic, so

$$h_{\mu\nu} = \mathcal{O}(\varepsilon), \qquad \partial_0 h_{\mu\nu} \sim \varepsilon^{1/2} h_{\mu\nu}, \qquad \partial_0 \partial_0 h_{\mu\nu} \sim \varepsilon h_{\mu\nu}$$

For consistency, we require that $\rho = \mathcal{O}(\varepsilon)$ and $p = \mathcal{O}(\varepsilon^2)$. The linearized Einstein equations becomes

$$\partial^\mu \partial_\mu \bar{h}_{\sigma\tau} = -16\pi T_{\sigma\tau},$$

where

$$T_{00} = \rho, \qquad T_{0i} = T_{i0} = 0.$$

We deduce that

$$\Delta \bar{h}_{00} = -16\pi\rho,$$

and $\bar{h}_{0i}, \bar{h}_{ii}$ at $\mathcal{O}(\varepsilon)$. Since $h_{00} = \bar{h}_{00} - \frac{1}{2}\eta_{\mu\nu}\bar{h}$, we get $h_{00} = \frac{1}{2}\bar{h}_{00}$. Hence

$$\Delta\left(-\frac{h_{00}}{2}\right) = 4\pi\rho.$$

This suggest we identify $-\frac{h_{00}}{2}$ with the Newtonian potential ϕ. Consider a test particle in this background. By the geodesic postulate, its motion is determined by the Lagrangian

$$\begin{aligned}\mathcal{L} &= g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = \eta_{\mu\nu}\dot{x}^\mu \dot{x}^\nu + h_{\mu\nu}\dot{x}^\mu \dot{x}^\nu \\ &= -\dot{t}^2 + |\dot{\mathbf{x}}|^2 + h_{00}\dot{t}^2 + \mathcal{O}(\varepsilon^2).\end{aligned}$$

Suppose the motion is non-relativistic, so $|\dot{\mathbf{x}}|^2 = \mathcal{O}(\varepsilon)$. Conservation of $\mathcal{L}$ gives

$$-\dot{t}^2 = -1 + \mathcal{O}(\varepsilon) \implies \dot{t} = 1 + \mathcal{O}(\varepsilon).$$

The Euler-Lagrange equations give

$$2\frac{\mathrm{d}}{\mathrm{d}\tau}(\dot{x}_i) = 2\ddot{x}_i = h_{00,i}\dot{t}^2.$$

Since $\dot{t} \approx 1$, we can sway τ with t to get

$$\frac{\mathrm{d}^2 x_i}{\mathrm{d}t^2} = \frac{1}{2}h_{00,i} + \mathcal{O}(\varepsilon^2) = -\partial_i\phi + \mathcal{O}(\varepsilon^2).$$

This recovers Newton's laws of gravitation.

4.7   Gravitational Waves

One of the most spectacular recent results in gravitational physics was the measurement in 2015 of gravitational waves, sourced by two colliding black holes.

Near the source, the field is not weak, but by the time we detect the waves the weak field approximation is relevant.

We seek a propagating wave solution to the vacuum linearized Einstein equation in the wave gauge:

$$\Box \bar{h}_{\mu\nu} = 0, \qquad \partial_\mu \bar{h}^\mu{}_\nu = 0.$$

Making the ansatz

$$\bar{h}_{\mu\nu} = \Re\left(H_{\mu\nu} e^{ik^\mu x_\mu} \right),$$

where H is symmetric and R are constants. We see that this is a solution if:

$$k^\mu k_\mu = 0, \qquad k_\mu H^\mu{}_\nu = 0.$$

The first equation says that k^μ is null, so waves travel at the speed of light. By gauge freedom, we can take $k^\mu = \omega(1,0,0,1)$.

There is still some residual gauge freedom. Recall that the diffeomorphism generated by ξ acts on $\partial^\mu \bar{h}_{\mu\nu}$ by

$$\partial^\mu \bar{h}_{\mu\nu} \to \partial^\mu \bar{h}_{\mu\nu} + \partial_\mu \partial^\mu \xi_\nu.$$

If $\partial_\mu \partial^\mu \xi_\nu = 0$, then the gauge condition is preserved. Let

$$\xi_\nu = \Re(-iX_\nu e^{ik^\mu x_\mu}).$$

Then $\partial_\mu \partial^\mu \xi_\nu = 0$, and we can check that

$$H_{\mu\nu} \to H_{\mu\nu} + k_\mu X_\nu + k_\nu X_\mu - \eta_{\mu\nu} k_\sigma X^\sigma.$$

We can show that if we take

$$X_0 = 0, \qquad X_i = -\frac{H_{0i}}{k_0},$$

then we get $H_{0\mu} = 0$. Moreover, making a further transformation of the form

$$X_0 = \alpha k_0, \qquad X_i = \alpha k_i,$$

we can further additionally suppose that $H_\mu{}^\mu = 0$, hence $h_{\mu\nu} = \bar{h}_{\mu\nu}$.

Since $H_{0\mu} = 0$ and $H_{\nu\mu}k^\nu = \omega(H_{0\mu} + H_{3\mu}) = 0$, we deduce $H_{3\mu} = 0$. With symmetry and tracelessness of $H_{\mu\nu}$, we have

$$H_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & H_+ & H_\times & 0 \\ 0 & H_\times & -H_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Here H_+ and $H_\times$ are constants corresponding to two independent polarizations of the wave.

To understand the consequences of such a gravitational wave, we consider the geodesic deviation between test particles in the path of this wave. Recall the geodesic deviation equation: if λ is a geodesic with tangent T, then a vector Y joining λ to a nearby geodesic satisfies

$$T^a \nabla_a (T^b \nabla_b Y^c) = R^c{}_{dab} T^d T^a Y^b.$$

Suppose a freely falling observer sets up a frame consisting of $e_0 = T$, tangent to the observer's worldline, together with three spacelike vectors e_1, e_2, e_3 which are parallelly transported along the world-line, and initially are orthonormal.

Since parallel transport preserves inner products, we can assume that $\{e_\mu\}$ is orthonormal for all time. Since $T^a \nabla_a(e_\alpha) = 0$, the geodesic equation implies

$$T^a \nabla_a (T^b \nabla_c (Y_c e^c_\alpha)) = R_{abcd} e^a_\alpha e^b_0 e^c_0 Y^d.$$

Now $Y_c e^c_\alpha = Y_\alpha$ is a scalar, so the equation becomes

$$\frac{\mathrm{d}^2 Y_\alpha}{\mathrm{d}t^2} = R_{abcd} e^a_\alpha e^b_0 e^c_0 Y^\beta e^d_\beta.$$

For our problem of a gravitational wave spacetime, the Riemann curvature is $\mathcal{O}(\varepsilon)$, so we only need e_α to leading order. We can assume that $e_0 = \partial_t$, $e_i = \partial_i$, and λ is $\mapsto (t, 0, 0, 0)$, so

$$\frac{\mathrm{d}^2 Y_\alpha}{\mathrm{d}\tau^2} = R_{\alpha 0 0 \beta} Y^\beta + \mathcal{O}(\varepsilon^2).$$

To order ε, we have

$$R_{\rho\sigma\mu\nu} = \frac{1}{2} \left(h_{\rho\nu,\mu\sigma} + h_{\sigma\mu,\nu\rho} - h_{\rho\mu,\nu\sigma} - h_{\sigma\nu,\mu\rho} \right).$$

But since $h_{0\mu} = 0$, we find

$$R_{\alpha 0 0 \beta} = \frac{1}{2} h_{\alpha\beta,00}.$$

So,
$$\frac{\mathrm{d}^2 Y_\alpha}{\mathrm{d}t^2} = \frac{1}{2}\frac{\partial^2 h_{\alpha\beta}}{\partial t^2}Y^\beta.$$

Considering the plus polarization,
$$h_{\mu\nu} = \Re(H_{\mu\nu}e^{ik^\mu x_\mu}) = |H_+|(\delta_{12} + \delta_{21})\cos\omega(t - zt_0).$$

Along λ, the 3 and 0 equations are
$$\frac{\mathrm{d}^2 Y_0}{\mathrm{d}t^2} = \frac{\mathrm{d}^2 Y_3}{\mathrm{d}t^2} = 0,$$

so no relative acceleration in the direction in the wave. The other equations are
$$\frac{\mathrm{d}^2 Y^1}{\mathrm{d}t^2} = -\frac{1}{2}\omega^2|H_+|\cos\omega(t - t_0)Y^1,$$
$$\frac{\mathrm{d}^2 Y^2}{\mathrm{d}t^2} = \frac{1}{2}\omega^2|H_+|\cos\omega(t - t_0)Y^2.$$

H_+ is small, so we can solve perturbatively with first derivatives 0. This gives
$$Y^1 = Y_0^1\left(1 + \frac{1}{2}|H_+|\cos\omega(t - t_0)\right),$$
$$Y^2 = Y_0^2\left(1 - \frac{1}{2}|H_+|\cos\omega(t - t_0)\right).$$

(include image on the effect on test particles.)

4.8   Field Far from a Source

We return to the linearized Einstein equation with matter:
$$\partial_\rho\partial^\rho \bar{h}_{\mu\nu} = -16\pi T_{\mu\nu}\ (*),\qquad\qquad \partial_\mu \bar{h}^\mu{}_\nu = 0.$$

As is the case for electromagnetism, we can solve the first equation $(*)$ explicitly using a retarded Greens function:
$$\bar{h}_{\mu\nu}(t, \mathbf{x}) = 4\int \mathrm{d}^3 x' \frac{T_{\mu\nu}(t - |\mathbf{x} - \mathbf{x}'|, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|},$$

where $|\mathbf{x} - \mathbf{x}'|$ is the distance computed in the Euclidean metric.

If the matter is concentrated within a distance d of the origin, so $T_{\mu\nu}(t', \mathbf{x}') = 0$ for $|\mathbf{x}'| > l$, then we can expand in the far field region where $r = |\mathbf{x}| \gg |\mathbf{x}'| \sim d$. Then,
$$|\mathbf{x} - \mathbf{x}'|^2 = r^2 - 2\mathbf{x}\cdot\mathbf{x}' + |\mathbf{x}'|^2 = r^2\left(1 - \frac{2}{r}\hat{\mathbf{x}}\cdot\mathbf{x}' + \mathcal{O}\left(\frac{d^2}{r^2}\right)\right),$$

with $\hat{\mathbf{x}} = \mathbf{x}/r$. Hence,

$$|\mathbf{x} - \mathbf{x}'| = r\left(1 - \frac{1}{r}\hat{\mathbf{x}} \cdot \mathbf{x}' + \mathcal{O}\left(\frac{d^2}{r^2}\right)\right) = r - \hat{\mathbf{x}} \cdot \mathbf{x} + \mathcal{O}\left(\frac{d^2}{r}\right),$$

and also

$$T_{\mu\nu}(t - |\mathbf{x} - \mathbf{x}'|, \mathbf{x}') = T_{\mu\nu}(t', \mathbf{x}') + \hat{\mathbf{x}} \cdot \mathbf{x}' \partial_0 T_{\mu\nu}(t', \mathbf{x}') + \cdots,$$

where $t' = t - r$. If $T_{\mu\nu}$ varies on a timescale τ so that $\partial_0 T_{\mu\nu} \sim T_{\mu\nu}/\tau$, then the second term above is of order $\mathcal{O}(1/\tau)$, which we can neglect if matter moves non-relativistically. We thus have

$$\hbar_{ij} = \frac{4}{r}\int \mathrm{d}^3x' \, T_{ij}(t', x'), \qquad (**)$$

where $t' = t - r$. This gives the spatial components of $\hbar_{ij}$. To find the remaining components we use the gauge condition: $\partial_0 \hbar_{0i} = \partial_j \hbar_{ji}$ and $\partial_0 \hbar_{00} = \partial_i \hbar_{i0}$. We first solve for $\hbar_{0i}$, then $\hbar_{00}$.

We can simplify the integral in $(**)$ by recalling that $\partial_\mu T^\mu{}_\nu = 0$ and that $T_{\mu\nu}(t', \mathbf{x}')$ vanishes for $|\mathbf{x}| > l$. Dropping the primes in the integral,

$$\int \mathrm{d}^3x \, T^{ij}(t, \mathbf{x}) = \int \mathrm{d}^3x \, \partial_k(T^{ik}x^j) - (\partial_k T^{ik})x^j$$

$$= \underbrace{\int_{|\mathbf{x}|=2d} T^{ik}x^j n^k \, \mathrm{d}S}_{0 \text{ as } T_{ij}=0} - \int \mathrm{d}^3x \, (\partial_k T^{ik})x^j$$

$$= \int \mathrm{d}^3x \, \partial_0 T^{0i}x^j = \partial_0 \int \mathrm{d}^3x \, T^{0i}x^j$$

Symmetrizing on i and j,

$$\int \mathrm{d}^3x \, T^{ij}(t, \mathbf{x}) = \partial_0 \int \mathrm{d}^3x \left(\frac{1}{2}T^{0i}x^j + \frac{1}{2}T^{0j}x^i\right)$$

$$= \partial_0 \int \mathrm{d}^3x \left(\frac{1}{2}\left(\partial_k(T^{k0}x^ix^j) - \frac{1}{2}\partial_k T^{0k}x^ix^j\right)\right)$$

$$= \partial_0 \int \mathrm{d}^3x \, \frac{1}{2}\partial_0 T^{00}x^ix^j$$

$$= \frac{1}{2}\partial_0^2 \int \mathrm{d}^3x \, T^{00}x^ix^j = \frac{1}{2}\ddot{I}^{ij}(t),$$

where

$$I^{ij}(t) = \int \mathrm{d}^3x \, T^{00}(t, \mathbf{x})x^ix^j.$$

Noting that $T_{00} = T^{00}$ and $T_{ij} = T^{ij}$, we deduce that

$$\hbar_{ij} = \frac{2}{r}\ddot{I}_{ij}(t - r).$$

Now we need to reconstruct the remaining components using the gauge condition:

$$\partial_0 \hbar_{0i} = \partial_j \hbar_{ji} = \partial_j \left(\frac{2}{r}\ddot{I}_{ij}(t - r) \right),$$

so integrating

$$\hbar_{0i} = \partial_j \left(\frac{2}{r}\dot{I}_{ij}(t - r) \right) + k_i(\mathbf{x}).$$

Expanding this all out,

$$\hbar_{0i} = -\frac{2\hat{x}_j}{r^2}\dot{I}_{ij}(t - r) - 2\frac{\hat{x}_j}{r^2}\ddot{I}_{ij}(t - r) + k_i(\mathbf{x}).$$

If we now assume that $r \gg \tau$, so we are in the radiation zone, we can drop the first term as it is order (τ/r) relative to the second. Hence

$$\hbar_{0i} = -2\frac{\hat{x}_j}{r}\ddot{I}_{ij}(t - r) + k_i(\mathbf{x}).$$

Now we use

$$\partial_0 \hbar_{00} = \partial_i \hbar_{0i} = \partial_i \left(-2\frac{\hat{x}_j}{r}\ddot{I}_{ij}(t - r) + k_i(\mathbf{x}) \right),$$

which gives

$$\hbar_{00} = -2\partial_i \left(\frac{\hat{x}_j}{r}\dot{I}_{ij}(t - r) \right) + t\partial_i k_i(\mathbf{x}) + f(\mathbf{x})$$

$$= 2\frac{\hat{x}_i \hat{x}_j}{r}\ddot{I}_{ij}(t - r) + t\partial_i k_i + f(\mathbf{x}),$$

plus terms in τ/r. To fix the constants of integration, return to

$$\hbar_{\mu\nu} = 4 \int \mathrm{d}^3 x' \frac{T_{\mu\nu}(t - |\mathbf{x} - \mathbf{x}'|, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|},$$

and observe that to leading order in d/τ,

$$\hbar_{00} \approx \frac{4E}{r}, \qquad \hbar_{0i} = -\frac{4P_i}{r},$$

where

$$E = \int \mathrm{d}^3 x' \, T_{00}(t', \mathbf{x}'), \qquad P_i = \int \mathrm{d}^3 x' \, T_{0i}(t', \mathbf{x}').$$

Observing that

$$\partial_0 \int \mathrm{d}^3 x'\, T_{0\mu}(t', \mathbf{x}') = \int \mathrm{d}^3 x'\, \partial_0 T_{0\mu}(t', x')$$
$$= \int \mathrm{d}^3 x'\, \partial_i T_{i\mu}(t', x') = 0,$$

by divergence, E and $\mathbf{P}$ are constant in time.

Exercise: By a gauge transformation generated by a multiple of

$$\xi^\mu = (\mathbf{P} \cdot \mathbf{x}, -\mathbf{P}t),$$

we can set $\mathbf{P} = 0$. This is the *center of momentum frame*.

We have shown that

$$\hbar_{00}(t, \mathbf{x}) = \frac{4M}{r} + 2\frac{\hat{x}_i \hat{x}_j}{r} \ddot{I}_{ij}(t - r),$$
$$\hbar_{0i}(t, \mathbf{x}) = -2\frac{\hat{x}_j}{r} \ddot{I}_{ij}(t, r),$$
$$\hbar_{ij}(t, \mathbf{x}) = \frac{2}{r} \ddot{I}_{ij}(t - r),$$

where $r \gg \tau \gg d$, in the center of momentum frame where $E = M$.

4.9   Energy in Gravitational Waves

Defining the local energy or local energy flux for a gravitational field is hard in general, because we can always choose coordinates such that the first derivatives $\partial_\mu g_{\sigma\tau}|_p = 0$.

There is no hope of an energy density quadratic in the first derivatives.

In the context of perturbation theory, there are various ways to define an energy. To do this, we consider how to continue a perturbative solution beyond linear order.

We consider the ungauged vacuum Einstein equations, and suppose that

$$g_{\mu\nu} = \eta_{\mu\nu} + \varepsilon h^{(1)}_{\mu\nu} + \varepsilon^2 h^{(2)}_{\mu\nu},$$

where we word to order $\mathcal{O}(\varepsilon^2)$. We observe that

$$R_{\mu\nu}(\eta_{\mu\nu} + \varepsilon h_{\mu\nu}) = \varepsilon R^{(1)}_{\mu\nu}[h] + \varepsilon^2 R^{(2)}_{\mu\nu}[h],$$

where

$$R^{(1)}_{\mu\nu}[h] = \partial^\rho \partial_{(\mu} h_{\nu)\rho} - \frac{1}{2} \partial^\rho \partial_\rho h_{\mu\nu} - \frac{1}{2} \partial_\mu \partial_\nu h,$$

and the quadratic terms are

$$R_{\mu\nu}^{(2)}[h] = \frac{1}{2}h^{\rho\sigma}\partial_\mu\partial_\nu h_{\rho\sigma} - h^{\rho\sigma}\partial_\rho\partial_{(\mu}h_{\nu)\sigma} + \frac{1}{4}\partial_\mu h_{\rho\sigma}\partial_\nu h^{\rho\sigma}$$
$$+ \partial^\sigma h^\rho{}_\nu\partial_{[\sigma}h_{\rho]\mu} + \frac{1}{2}\partial_\sigma(h^{\sigma\rho}\partial_\rho h_{\mu\nu}) - \frac{1}{4}\partial^\rho h\partial_\rho h_{\mu\nu}$$
$$- \left(\partial_\sigma h^{\rho\sigma} - \frac{1}{2}\partial^\rho h\right)\partial_{(\mu}h_{\nu)\rho}.$$

This implies that

$$R_{\mu\nu}(\eta_{\mu\nu} + \varepsilon h_{\mu\nu}^{(1)} + \varepsilon^2 h_{\mu\nu}^{(2)}) = \varepsilon R_{\mu\nu}^{(1)}[h^{(1)}] + \varepsilon^2(R_{\mu\nu}^{(2)}[h^{(1)}] + R_{\mu\nu}^{(1)}[h^{(2)}]).$$

Thus,

$$G_{\mu\nu}[\eta + \varepsilon h^{(1)} + \varepsilon^2 h^{(2)}] = \varepsilon G_{\mu\nu}^{(1)}[h^{(1)}] + \varepsilon^2\Big(G_{\mu\nu}^{(1)}[h^{(2)}] + R_{\mu\nu}^{(2)}[h^{(1)}] - \frac{1}{2}\eta_{\mu\nu}\eta^{\sigma\tau}R_{\sigma\tau}^{(2)}[h^{(1)}]$$
$$+ \frac{1}{2}\eta_{\mu\nu}h^{(1)\sigma\tau}R_{\sigma\tau}^{(1)}[h^{(2)}] - \frac{1}{2}h_{\mu\nu}\eta^{\sigma\tau}R_{\sigma\tau}^{(1)}[h^{(2)}]\Big).$$

Here,

$$G_{\mu\nu}^{(1)}[h] = R_{\mu\nu}^{(1)}[h] - \frac{1}{2}\eta_{\mu\nu}\eta^{\sigma\tau}R_{\sigma\tau}^{(1)}[h].$$

We now consider the contracted Bianchi identity:

$$g^{\mu\rho}\nabla_\rho G_{\mu\nu} = 0,$$

which holds for any metric. Using our expression for $G_{\mu\nu}$ and expanding gives

$$0 = \varepsilon\eta^{\mu\sigma}\partial_\sigma G_{\mu\nu}^{(1)}[h^{(1)}] + \varepsilon^2(\eta^{\mu\sigma}\partial_\sigma G_{\mu\nu}^{(1)}[h^{(2)}] - 8\pi\partial_\mu t_{\sigma\nu}\eta^{\mu\sigma}[h^{(1)}] + h^{(1)} \cdot R^{(1)}[h^{(1)}]),$$
$$\text{(†)}$$

where t is what we had in the previous equation. Considering

$$G_{\mu\nu}[\eta + \varepsilon h^{(1)} + \varepsilon^2 h^{(2)}] = 0,$$

looking at the equation order by order we deduce that

$$G_{\mu\nu}^{(1)}[h^{(1)}] = 0,$$
$$G_{\mu\nu}^{(1)}[h^{(2)}] = -R_{\mu\nu}^{(2)}[h^{(1)}] + \frac{1}{2}\eta_{\mu\nu}\eta^{\sigma\tau}R_{\sigma\tau}^{(2)}[h^{(1)}] = 8\pi t_{\mu\nu}[h^{(1)}].$$

Thus $h^{(2)}$ solves the linearized Einstein equations sourced by an 'energy momentum tensor':

$$t_{\mu\nu} = -\frac{1}{8\pi}\left(R_{\mu\nu}^{(2)}[h^{(1)}] - \frac{1}{2}\eta_{\mu\nu}\eta^{\rho\sigma}R_{\rho\sigma}^{(2)}[h^{(1)}]\right).$$

From (†), we deduce that

$$\eta^{\mu\sigma}\partial_\mu G^{(1)}_{\sigma\nu}[h] = 0,$$

which holds for any perturbation. Hence,

$$\eta^{\mu\sigma}\partial_\mu t_{\sigma\nu}[h^{(1)}] = 0,$$

when $h^{(1)}$ satisfies the linearized Einstein equations. We can identify $t_{\mu\nu}$ with the energy-momentum tensor of the gravitational field, however it is not gauge invariant.

If $h^{(1)}$ decays sufficiently at ∞, then we can show that

$$\int t_{00}\, \mathrm{d}^3 x$$

is invariant, so this gives the total energy of the field, but there is no gauge invariant local conservation.

We can get approximate gauge invariance by averaging. Let W be smooth, and vanish for $|\mathbf{x}|^2 + t^2 > a$ and satisfies

$$\int_{\mathbb{R}^4} W(\mathbf{x}, t)\, \mathrm{d}^3 x\, \mathrm{d}t = 1.$$

We define the *average* of a tensor in almost-inertial coordinates by setting

$$\langle x_{\mu\nu}(\mathbf{x}) \rangle = \int W(y - x) X_{\mu\nu}(y)\, \mathrm{d}^4 y.$$

Suppose we are in the far-field regime, with radiation of wavelength λ, and we average over a region of size $a \gg \lambda$. Since W is supported on the ball of size a, roughly speaking $\partial_\mu W \sim W/a$, and so if we compute the average of a derivative of a tensor field, by IBP,

$$\langle \partial_\mu X_{\rho\sigma} \rangle = \int_{\mathbb{R}^4} \partial_\mu W(y - x) X_{\rho\sigma}(y)\, \mathrm{d}^4 y \sim \frac{\langle X_{\rho\sigma} \rangle}{a} \sim \frac{\lambda}{a}\langle \partial_\mu X_{\rho\sigma} \rangle,$$

so we can ignore total derivatives inside averages, and hence

$$\langle A\partial B \rangle = \langle \partial(AB) \rangle - \langle (\partial A)B \rangle \approx -\langle (\partial A)B \rangle.$$

With this, we can show that:

- If h solves the vacuum linearized Einstein equations,

$$\langle \eta^{\mu\nu} R^{(2)}_{\mu\nu}[h] \rangle = 0.$$

- We have

$$\langle t_{\mu\nu} \rangle = \frac{1}{32\pi} \langle \partial_\mu \hbar_{\rho\sigma} \partial_\nu \hbar^{\rho\sigma} - \frac{1}{2} \partial_\mu \hbar \partial_\nu \hbar - 2\partial_\sigma \hbar^{\rho\sigma} \partial_{(\mu} \hbar_{\nu)\rho} \rangle.$$

- $\langle t_{\mu\nu} \rangle$ is gauge invariant.

Using this formula and the results obtained previously, we can find the energy lost by a system producing gravitational waves. The average spatial energy flux is

$$S_i = \langle -t_{0i} \rangle.$$

We calculate the average energy flux across a sphere of radius r centred on the source:

$$\langle P \rangle = - \int_{S_r} r^2 \, \mathrm{d}\Omega \, \langle t_{0i} \rangle \hat{x}_i.$$

In wave gauge,

$$\langle t_{0i} \rangle = \frac{1}{32\pi} \langle \partial_0 \hbar_{\rho\sigma} \partial_\nu \hbar^{\rho\sigma} - \frac{1}{2} \partial_0 \hbar \partial_i \hbar \rangle$$

$$= \frac{1}{32\pi} \langle \partial_0 \hbar_{jk} \partial_i \hbar_{jk} - 2\partial_0 \hbar_{0j} \partial_i \hbar_{0j} + \partial_0 \hbar_{00} \partial_i \hbar_{00} - \frac{1}{2} \partial_0 \hbar \partial_i \hbar \rangle.$$

Using the fact that

$$\hbar_{ij} = \frac{2}{r} \ddot{I}_{ij}(t - r),$$

$$\partial_0 \hbar_{jk} = \frac{2}{r} \dddot{I}_{jk}(t - r),$$

$$\partial_i \hbar_{jk} = \left(-\frac{2}{r} \dddot{I}_{jk}(t - r) - \frac{2}{r^2} \ddot{I}_{jk}(t - r) \right) \hat{x}_i,$$

hence we get

$$-\frac{1}{32\pi} \int r^2 \, \mathrm{d}\Omega \, \langle \partial_0 \hbar_{jk} \partial_i \hbar_{jk} \rangle \hat{x}_i = \frac{1}{2} \langle \dddot{I}_{ij} \dddot{I}_{ij} \rangle_{t-r},$$

the average over the window centred at $t - r$.

Inserting the expressions from the radiation zone solution derived previously, we can compute the flux through a large sphere as

$$\langle P \rangle_t = \int_{|\mathbf{x}| \leq r} r^2 \, \mathrm{d}\Omega \, \langle t_{0i} \rangle \hat{x}_i = \frac{1}{5} \langle \dddot{Q}_{ij} \dddot{Q}_{ij} \rangle_{t-r},$$

where $Q_{ij} = I_{ij} - \frac{1}{3} I_{kk} \delta_{ij}$ is the *quadrupole tensor*.

5   Differential Forms

We want to derive Einstein's equations from an action principle; for this, we need to discuss integration on manifolds, and to do so we first discuss differential forms.

Definition 5.1. A *p-form* is a totally anti-symmetric $(0, p)$-tensor field on M. The space of p-forms in $\Omega^p M$.

Note that if $p > \dim M$, then a p-form must vanish, and a 1-form is a covector field.

The reason p-forms are nice is that we have a natural product on forms: if X is a p-form and Y is a q-form, then $(X \wedge Y)$ is the $(p + q)$-form given by

$$(X \wedge Y)_{a_1 \ldots a_p b_1 \ldots b_q} = \frac{(p+q)!}{p!q!} X_{[a_1 \ldots a_p} Y_{b_1 \ldots b_q]}.$$

It has the following properties:

- $X \wedge Y = (-1)^{pq} Y \wedge X$, so $X \wedge X = 0$ if p is odd.

- $(X \wedge Y) \wedge Z = X \wedge (Y \wedge Z)$.

If $\{f^\mu\}$ is a dual basis, then $\{f^{\mu_1} \wedge \cdots \wedge f^{\mu_p}\}$ is a basis for the p-forms, and we can write

$$X = \frac{1}{p!} X_{\mu_1 \ldots \mu_p} f^{\mu_1} \wedge \cdots \wedge f^{\mu_p}.$$

Another important feature of forms is that we can define the derivative $\mathrm{d} : \Omega^p M \to \Omega^{p+1} M$ by, if X is a p-form, then in a coordinate basis

$$(\mathrm{d}X)_{\mu_1 \ldots \mu_{p+1}} = (p + 1)\partial_{[\mu_1} X_{\mu_2 \ldots \mu_{p+1}]}.$$

The first thing to do is check that this is a tensor. Suppose that ∇ is any symmetric connection. Then

$$\nabla_{\mu_1} X_{\mu_2 \ldots \mu_{p+1}} = \partial_{\mu_1} X_{\mu_2 \ldots \mu_{p+1}} - \Gamma_{\mu_2}{}^\sigma{}_{\mu_1} X_{\sigma \mu_3 \ldots \mu_{p+1}} - \cdots - \Gamma_{\mu_{p+1}}{}^\sigma{}_{\mu_1} X_{\mu_2 \ldots \mu_p \sigma}.$$

The connection terms are symmetric in μ_1 and μ_i, so antisymmetrizing these terms drop out, hence

$$\nabla_{[\mu_1} X_{\mu_2 \ldots \mu_{p+1}]} = \partial_{[\mu_1} X_{\mu_2 \ldots \mu_{p+1}]}.$$

Hence we can write, in AIN

$$(\mathrm{d}X)_{a_1 \ldots a_{p+1}} = (p + 1)\nabla_{[a_1} X_{a_2 \ldots a_{p+1}]}$$

is well-defined independently of coordinates. The reason we do not define this directly is that the first representation shows that the derivative does not depend on a metric or the connection. We can show that:

- $d(dX) = 0$.

- $d(X \wedge Y) = dX \wedge Y + (-1)^p X \wedge dY$, where X is a p-form and Y is a q-form.

- $\phi^* \, dX = d(\phi^* X)$, where $\phi : N \to M$.

The last property implies that d commutes with Lie derivatives, do

$$\mathcal{L}_V(dX) = d(\mathcal{L}_V X).$$

d is called the *exterior derivative*. We say that X is *closed* if $dX = 0$, and X is *exact* if $X = dY$ for some Y.

Note that being exact implies being closed, but the converse is only true locally, due to the following:

Lemma 5.1 (Poincaré Lemma)**.** *If X is a closed p-form, then for any $r \in M$ there is an open neighbourhood $N \subseteq M$ with $r \in N$ and a $(p-1)$ form Y defined on N such that*

$$X = dY.$$

The extent to which closed does not equal exact captures topological properties of M.

> **Example 5.1.**
>
> On S^1, the form $d\theta$ is closed, but not exact, despite confusing notation.

5.1   The Tetrad Formalism

In GR, it is often useful to work with an orthonormal basis of vector fields, called a *tetrad* $\{e_\mu^a\}$ satisfying

$$g_{ab} e_\mu^a e_\nu^b = \eta_{\mu\nu},$$

or $\delta_{\mu\nu}$ if M is Riemannian. Recall that the dual basis $\{f_a^\mu\}$ is defined by

$$\delta^\mu{}_\nu = f^\mu(e_\nu) = f_a^\mu e_\nu^a.$$

Note that Greek indices denote which of the vector fields we are taking, and Roman indices denote the tensor components.

We claim that

$$f_a^\mu = \eta^{\mu\sigma} g_{ab} e_\sigma^b.$$

Proof: Indeed,

$$(\eta^{\mu\sigma} g_{ab} e^b{}_\sigma) e^a{}_\nu = \eta^{\mu\sigma} \eta_{\sigma\nu} = \delta^\mu{}_\nu.$$

Recalling that g_{ab} raises and lowers Roman indices, and introducing the convention that $\eta_{\mu\nu}$ raises and lowers Greek indices, we have

$$f^\mu{}_a = e^\mu{}_a = \eta^{\mu\sigma} g_{ab} e^b{}_\sigma.$$

We will thus denote basis vectors e_μ and dual basis vectors are e^μ.

Recall that two orthonormal bases are related by

$$e'^a{}_\mu = (A^{-1})^\nu{}_\mu e^a{}_\nu,$$

where

$$\eta_{\mu\nu} A^\mu{}_\rho A^\nu{}_\sigma = \eta_{\rho\sigma}.$$

This looks like special relativity, but $A^\mu{}_\nu$ need not be constant. Here general relativity arises by gauging the Lorentz symmetry of special relativity.

We claim that

$$\eta_{\mu\nu} e_a{}^\mu e_b{}^\nu = g_{ab}, \qquad e_a{}^\mu e^b{}_\mu = \delta_a{}^b.$$

Proof: For the first, contract with $e^b{}_\rho$, to get

$$\eta_{\mu\nu} e_a{}^\mu e_b{}^\nu e^b{}_\rho = \eta_{\mu\nu} e_a{}^\mu \delta^\nu{}_\rho = \eta_{\mu\rho} e_a{}^\mu = (e_a)_\rho = g_{ab} e^b{}_\rho.$$

Since this equation holds contracted with any basis vector, it holds in generality. The second equation follows from the first by raising b.

5.2   Connection 1-Forms

Let ∇ be the Levi-Civita connection. The *connection 1-forms* are defined to be

$$(\omega^\mu{}_\nu)_a = e_b{}^\mu \nabla_a e^b{}_\nu.$$

Recalling

$$e^a{}_\nu \nabla_a e^b{}_\mu = \Gamma_\mu{}^\sigma{}_\nu e^b{}_\sigma,$$

multiplying by $e_c{}^\nu$,

$$\nabla_c e^b{}_\mu = \Gamma_\mu{}^\sigma{}_\nu e^b{}_\sigma e_c{}^\nu,$$

hence we find

$$(\omega^\mu{}_\nu)_a = \Gamma_\rho{}^\mu{}_\nu e^\rho{}_a.$$

So $(\omega^\mu{}_\nu)_a$ encodes the connection components.

Lemma 5.2.
$$(\omega_{\mu\nu})_a = -(\omega_{\nu\mu})_a.$$

> **Proof:** We have
> $$(\omega_{\mu\nu})_a = (e_b)_\mu \nabla_a e^b{}_\nu = \nabla_a((e_b)_\mu e^b{}_\nu) - e^b{}_\nu \nabla_a(e_b)_\mu$$
> $$= -(e_b)_\nu \nabla_a e^b{}_\mu = -(\omega_{\nu\mu})_a.$$

Now consider the exterior derivative of a basis 1-form:

Lemma 5.3. *The one-form e^μ satisfies Cartan's first structure equation:*
$$\mathrm{d}e^\mu + \omega^\mu{}_\nu \wedge e^\nu = 0.$$

> **Proof:** Note
> $$(\omega^\nu{}_\mu)_a e^b{}_\nu = (e_c{}^\nu \nabla_a e^c{}_\mu) e^b{}_\nu = \nabla_a e^b{}_\mu.$$
> Thus,
> $$\nabla_a e_b{}^\mu = (\omega_\nu{}^\mu)_a e_b{}^\nu = -(\omega^\mu{}_\nu)_a e_b{}^\nu,$$
> which gives
> $$(\mathrm{d}e^\mu)_{ab} = 2\nabla_{[a} e^\mu{}_{b]} = -2(\omega^\mu{}_\nu)_{[a} e_{b]}{}^\nu = -(\omega^\mu{}_\nu \wedge e^\nu)_{ab}.$$

Note that in the orthonormal basis,
$$(\mathrm{d}e^\mu)_{\nu\sigma} = 2(\omega^\mu{}_{[\nu})_{\sigma]},$$

so computing $\mathrm{d}e^\mu$ leads to $(\omega^\mu{}_{[\nu})_{\sigma]}$. Since $(\omega_{\mu\nu})_\rho = -(\omega_{\nu\mu})_\rho$, we can check that
$$(\omega_{\mu\nu})_\rho = (\omega_{\mu[\nu})_{\rho]} + (\omega_{\nu[\rho})_{\mu]} - (\omega_{\rho[\mu})_{\nu]}.$$

We can determine $\omega^\mu{}_\nu$ and hence Γ by computing $\mathrm{d}e^\mu$.

> **Example 5.2.**
>
> The Schwarzschild metric has an obvious tetrad:
> $$e^0 = f\,\mathrm{d}t, \qquad\qquad e^1 = \frac{1}{f}\,\mathrm{d}r,$$
> $$e^2 = r\,\mathrm{d}\theta \qquad\qquad e^3 = r\sin\theta\,\mathrm{d}\phi,$$

where $f = \sqrt{1 - 2M/r}$. Then

$$\mathrm{d}e^0 = \mathrm{d}f \wedge \mathrm{d}t + f \, \mathrm{d}(\mathrm{d}t) = f' \, \mathrm{d}r \wedge \mathrm{d}t = f' e^1 \wedge e^0,$$

$$\mathrm{d}e^1 = \mathrm{d}\left(\frac{1}{f}\right) \wedge \mathrm{d}r + \frac{1}{f} \, \mathrm{d}(\mathrm{d}r) = -\frac{f'}{f^2} \, \mathrm{d}r \wedge \mathrm{d}r = 0,$$

$$\mathrm{d}e^2 = \mathrm{d}r \wedge \mathrm{d}\theta = \frac{f}{r} e^1 \wedge e^2,$$

$$\mathrm{d}e^3 = \sin\theta \, \mathrm{d}r \wedge \mathrm{d}\phi + r \cos\theta \, \mathrm{d}\theta \wedge \mathrm{d}\phi = \frac{f}{r} e^1 \wedge e^3 + \frac{\cot\theta}{r} e^2 \wedge e^3.$$

Note $\mathrm{d}e^i = -\omega^i{}_\mu \wedge e^\mu$, so we can work out the components of ω.

This lets us compute what ω is. Note $\omega^0{}_0 = 0$ by antisymmetry, so the first equation implies

$$\omega^0{}_1 = f' e^0, \qquad \omega^0{}_2 \propto e^2, \qquad \omega^0{}_3 \propto e^3.$$

Similarly we get

$$\omega^2{}_1 = \frac{f}{r} e^2, \qquad \omega^3{}_1 = \frac{f}{r} e^3, \qquad \omega^3{}_2 = \frac{\cot\theta}{r} e^3.$$

This gives non-zero terms

$$\omega_{01} = -\omega_{10} = -f' e^0,$$

$$\omega_{21} = -\omega_{12} = \frac{f}{r} e^2,$$

$$\omega_{31} = -\omega_{13} = \frac{f}{r} e^3,$$

$$\omega_{32} = -\omega_{23} = \frac{\cot\theta}{r} e^3,$$

and all other components vanish.

5.3   Curvature 2-forms

We can compute $\mathrm{d}\omega^\mu{}_\nu$:

$$
\begin{aligned}
(\mathrm{d}\omega^\mu{}_\nu)_{ab} &= \nabla_a(\omega^\mu{}_\nu)_b - \nabla_b(\omega^\mu{}_\nu)_a \\
&= \nabla_a(e_c{}^\mu \nabla_b e^c{}_\nu) - \nabla_b(e_c{}^\mu \nabla_a e^c{}_\nu) \\
&= e_c{}^\mu(\nabla_a\nabla_b e^c{}_\nu - \nabla_b\nabla_a e^c{}_\nu) + \nabla_a e_c{}^\mu \nabla_b e^c{}_\nu - \nabla_b e_c{}^\mu \nabla_a e^c{}_\nu \\
&= e_c{}^\mu(R^c{}_{dab})e^d{}_\nu + e^d{}_\sigma(\nabla_a e_d{}^\mu)e_f{}^\sigma(\nabla_b e^f{}_\nu) - e^d{}_\sigma(\nabla_b e_d{}^\mu)e_f{}^\sigma(\nabla_a e^f{}_\nu) \\
&= (\Theta^\mu{}_\nu)_{ab} + (\omega_\sigma{}^\mu \wedge \omega^\sigma{}_\nu)_{ab},
\end{aligned}
$$

where

$$
\Theta^\mu{}_\nu = \frac{1}{2}R^\mu{}_{\nu\sigma\tau}e^\sigma \wedge e^\tau
$$

are the *curvature 2-forms*. We have shown Cartan's second structure equation:

$$
\mathrm{d}\omega^\mu{}_\nu + \omega^\mu{}_\sigma \wedge \omega^\sigma{}_\nu = \Theta^\mu{}_\nu,
$$

which gives an efficient way to compute $R^\mu{}_{\nu\sigma\tau}$ in an orthonormal basis.

Example 5.3.

Returning to our example,

$$
\begin{aligned}
\mathrm{d}\omega^0{}_1 = \mathrm{d}(f'e^0) = \mathrm{d}(f'f\,\mathrm{d}t) &= (f''f + f'^2)\,\mathrm{d}r \wedge \mathrm{d}t \\
&= (f''f + f'^2)e^1 \wedge e^0,
\end{aligned}
$$

We also have

$$
\omega^0{}_\mu \wedge \omega^\mu{}_1 = \omega^0{}_1 \wedge \omega^1{}_1 + \cdots + \omega^0{}_0 \wedge \omega^0{}_1 = 0,
$$

hence we find that

$$
\Theta^0{}_1 = -(f''f + f'^2)e^0 \wedge e^1,
$$

giving

$$
R^0{}_{101} = -(f''f + f'^2) = -\frac{1}{2}(f^2)'' = \frac{2M}{r^3},
$$

and $R^\mu{}_{\nu01} = 0$ otherwise.

5.4   Volume Form and Hodge Dual

We say a manifold is *orientable* if it admits a nowhere vanishing n-form, where $n = \dim M$, say $\varepsilon_{a_1\ldots a_n}$. This is called an *orientation form*.

Two such forms are equivalent if $\varepsilon' = f\varepsilon$ for some smooth, everywhere positive f. $[\varepsilon]_\sigma$ is an orientation.

A basis of vectors $\{e_\mu\}$ is *right-handed* if

$$\varepsilon(e_1, \ldots, e_n) > 0.$$

A coordinate system is *right-handed* if the corresponding basis vectors $\{\partial/\partial x^\mu\}$ are right-handed.

An oriented manifold with metric has a preferred normalization for ε for a right-handed orthonormal basis. We define the *volume form* ε by

$$\varepsilon(e_1, \ldots, e_n) = 1.$$

This is independent of the choice of right-handed orthonormal basis.

If we work in a right-handed coordinate system $\{x^\alpha\}$, then

$$\frac{\partial}{\partial x^\alpha} = e_\alpha{}^\mu e^\beta{}_\mu \frac{\partial}{\partial x^\beta} = e_\alpha{}^\mu e_\mu,$$

then we have that

$$\varepsilon\left(\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}\right) = \varepsilon(e_1{}^{\mu_1} e_{\mu_1}, \ldots, e_n{}^{\mu_n} e_{\mu_n})$$

$$= \sum_{\pi \in \mathrm{Sym}(n)} \sigma(\pi) e_1{}^{\pi(1)} \cdots e_n{}^{\pi(n)} = \det(e_\alpha{}^\mu).$$

But also,

$$e_\alpha{}^\mu e_\beta{}^\nu \eta_{\mu\nu} = g_{\alpha\beta} \implies \det(e_\alpha{}^\mu) = \sqrt{|g|},$$

where $g = \det(g_{\alpha\beta})$. Hence,

$$\varepsilon = \sqrt{|g|}\, \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \cdots \wedge \mathrm{d}x^n.$$

Equivalently, $\varepsilon_{123\ldots n} = \sqrt{|g|}$ in a coordinate basis. In the same coordinate basis, we can show

$$\varepsilon^{123\ldots n} = \pm\frac{1}{\sqrt{|g|}},$$

where the sign is $+$ in Riemannian, and $-$ in Lorentzian.

Lemma 5.4. $\nabla\varepsilon = 0$.

> **Proof:** In normal coordinates at p, $\partial_\mu g_{\nu\sigma}|_p = 0$, so $\Gamma_\mu{}^\sigma{}_\nu|_p = 0$, hence
>
> $$\nabla_{\mu_1}\varepsilon_{\mu_2\ldots\mu_{n+1}} = \partial_{\mu_1}\varepsilon_{\mu_2\ldots\mu_{n+1}} + (\Gamma \cdot \varepsilon) = 0$$
>
> at p. This is a tensor equation, so it holds everywhere.

Lemma 5.5.

$$\varepsilon^{a_1\ldots a_p c_{p+1}\ldots c^n}\varepsilon_{b_1\ldots b_p c_{p+1}\ldots c_n} = \pm p!(n-p)!\delta^{a_1}{}_{[b_1}\delta^{a_2}{}_{b_2}\cdots\delta^{a_p}{}_{b_p]},$$

where again the sign is $+$ in the Riemannian case, and $-$ in the Lorentzian case.

The proof is an exercise.

We can use ε to relate $\Omega^p M$ and $\Omega^{n-p} M$.

Definition 5.2. On an oriented manifold with a metric, the *Hodge dual* of a p-form X is

$$(*X)_{a_1\ldots a_{n-p}} = \frac{1}{p!}\varepsilon_{a_1\ldots a_{n-p}b_1\ldots b_p}X^{b_1\ldots b_p}.$$

From the previous results, we can show:

Lemma 5.6.

- $*(*X) = \pm(-1)^{p(n-p)}X$.

- $(*\mathrm{d}*X)_{a_1\ldots a_{p-1}} = \pm(-1)^{p(n-p)}\nabla^b X_{a_1\ldots a_{p-1}b}$.

> ### Example 5.4.
>
> 1. In Euclidean space, identifying a vector field X^a with the one form X_a, the usual operations of vector calculus become
>
> $$\nabla f = \mathrm{d}f, \qquad \mathrm{div}X = *\mathrm{d}*X, \qquad \mathrm{curl}X = *\mathrm{d}X.$$
>
>   Note that $\mathrm{d}^2 = 0$ means
>
> $$\mathrm{curl}\,\nabla f = 0, \qquad \mathrm{div}\,\mathrm{curl}\,X = 0.$$
>
> 2. Maxwell's equations are
>
> $$\nabla^a F_{ab} = -4\pi j_b, \qquad \nabla_{[a}F_{bc]} = 0.$$
>
>   This can be written as
>
> $$\mathrm{d}*F = -4\pi *j, \qquad \mathrm{d}F = 0.$$
>
>   From Poincaré's lemma, we can locally write $F = \mathrm{d}A$ for some one-form A.

5.5   Integration on Manifolds

Suppose on a manifold M we have a right-handed coordinate chart $\phi : \mathcal{O} \to U$ with x^μ. If X is an n-form which vanishes outside $\mathcal{O}$, we can write

$$X = X_{1\ldots n}\, dx^1 \wedge \ldots \wedge dx^n.$$

If $\psi : \mathcal{O} \to U$ is another right-handed coordinate chart with coordinates $\{y^\mu\}$ then

$$X = \tilde{X}_{1\ldots n}\, dy^1 \wedge \ldots \wedge dy^n = \tilde{X}_{1\ldots n} \frac{\partial y^1}{\partial x^{\mu_1}} \cdots \frac{\partial y^n}{\partial x^{\mu_n}}\, dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_n}$$

$$= \tilde{X}_{1\ldots n} \det\left(\frac{\partial y^\mu}{\partial x^\nu}\right) dx^1 \wedge \ldots \wedge dx^n.$$

Hence we find that

$$X_{1\ldots n} = \tilde{X}_{1\ldots n} \det\left(\frac{\partial y^\mu}{\partial x^\nu}\right).$$

As a result,

$$\int_U X_{1\ldots n}\, dx^1 \ldots dx^n = \int_U \tilde{X}_{1\ldots n}\, dy^1 \ldots dy^n.$$

We can define

$$\int_M X = \int_{\mathcal{O}} X = \int_U X_{1\ldots n}\, dx^1 \ldots dx^n.$$

On any (second countable) manifold, we can find a countable atlas of charts $(\phi_i, \mathcal{O}_i)$ and smooth functions $\chi_i : M \to [0,1]$ such that χ_i vanishes outside $\mathcal{O}_i$, and

$$\sum_{i=1}^{\infty} \chi_i(p) = 1$$

for all $p \in M$, and the sum is locally finite. Then for any n-form X, we efine

$$\int_M X = \sum_{i=1}^{\infty} \int_M \chi_i X.$$

This does not depend on a choice of χ_i.

Remark.

- The computation showing coordinate invariance implies that for a diffeomorphism $\phi : M \to M$,

$$\int_M X = \int_M \phi^* X.$$

- If X is a manifold with metric and volume form ε, then if $f : M \to \mathbb{R}$ is a scalar, then $f\varepsilon$ is an n-form, and we can define

$$\int_M f = \int_M f\varepsilon.$$

In local coordinates, if f vanishes outside of $\mathcal{O}$, then

$$\int_M f = \int_U f(x)\sqrt{|g|}\,\mathrm{d}x^1 \ldots \mathrm{d}x^n = \int_M f\,\mathrm{dvol}_g.$$

5.6   Submanifolds and Stokes Theorem

Definition 5.3. Suppose S, M are manifolds and $\dim S = m < n = \dim M$. A smooth map $\iota : S \to M$ is an *embedding* if it is an injection, i.e. $\iota_* : T_pS \to T_{\iota(p)}M$ is injective, and if ι is injective.

If ι is an embedding, then $\iota(S)$ is an embedded *submanifold*. If $m = n - 1$, we call it a hypersurface.

If S, M are orientable, and $\iota(S)$ is an embedded submanifold of M, we define the integral of an m-form X over $\iota(S)$ by

$$\int_{\iota(S)} X = \int_S \iota^*(X).$$

Note that if $X = \mathrm{d}Y$, then

$$\int_{\iota(S)} \mathrm{d}Y = \int_S \mathrm{d}(i^*Y).$$

Definition 5.4. A *manifold with boundary*, M, is defined just as for a manifold, but charts are maps

$$\phi_a : \mathcal{O}_a \to U_a,$$

where U_a is an open subset of

$$\mathbb{R}^n_{\leq 0} = \{(x^1, \ldots, x^n) \mid x^1 \leq 0\}.$$

The *boundary* of M, ∂M, is the set of points mapped in any chart to $\{x^1 = 0\}$. It is naturally an $n - 1$ dimensional manifold with embedding $\iota : \partial M \to M$.

If M is oriented, then ∂M inherits an orientation by requiring that $(x^2, \ldots, x^n)$ is a right-handed chart on ∂M when $(x^1, \ldots, x^n)$ is right-handed on M.

5.7   Stokes' Theorem

Theorem 5.1 (Stokes' Theorem). *If N is an oriented n-dimensional manifold with boundary, and X is an $(n-1)$-form, then*

$$\int_N \mathrm{d}X = \int_{\partial N} X.$$

Stokes' theorem is the basis of all 'integration by parts' arguments. If N carries a metric, we can reformulate Stokes' theorem as the divergence theorem.

If V is a vector field on N, then we define the *interior product* or *interior derivative* or *hook product*

$$(V \lrcorner \varepsilon)_{a_2 \ldots a_n} = V^a \varepsilon_{a a_2 \ldots a_n}.$$

We can check that

$$\mathrm{d}(V \lrcorner \varepsilon) = (\nabla_a V^a)\varepsilon.$$

If we define the flux of V through an embedded hypersurface S by

$$\int_S V \cdot \mathrm{d}S = \int_S V \lrcorner \varepsilon,$$

then Stokes' theorem implies that

$$\int_N \nabla_a V^a \, \mathrm{dvol}_g = \int_{\partial N} V \cdot \mathrm{d}S.$$

Recall that a hypersurface $\iota(S)$ is:

- spacelike if $h = \iota^* g$ is Riemannian.

- timelike if $h = \iota^* g$ is Lorentzian.

In these cases we can relate $\iota^*(V \lrcorner \varepsilon_g)$ to the volume form on (S, h).

Pick $b_2, \ldots, b_n$ a right-handed orthonormal basis on S with respect to h. Then $\iota_* b_2, \ldots, \iota_* b_n$ are orthonormal with respect to N.

The *unit normal* to S is the unique unit vector $\vec{n}$ orthogonal to $\iota_* b_2, \ldots, \iota_* b_n$ with

$$\varepsilon(\vec{n}, \iota_* b_2, \ldots, \iota_* b_n) = g(\vec{n}, \vec{n}).$$

Then if $\iota(S)$ is spacelike, then $\vec{n}$ is timelike, and if $\iota(S)$ is timelike, $\vec{n}$ is spacelike.

With this definition,

$$V \lrcorner \varepsilon(\iota_* b_2, \ldots, \iota_* b_n) = V^a \vec{n}_a.$$

Thus,

$$\iota^*(V \lrcorner \varepsilon_g) = \iota^*(V^a \vec{n}_a) \varepsilon_h,$$

which means

$$\int_S V \cdot \mathrm{d}S = \int_S V^a \vec{n}_a \, \mathrm{dvol}_h.$$

We have shown that

$$\int_{\partial N} V^a \vec{n}_a \, \mathrm{dvol}_h = \int_N \nabla_a V^a \, \mathrm{dvol}_g.$$

Checking the definition, $\vec{n}$ points out of N for ∂N timelike, and into N for ∂N spacelike.

5.8   The Einstein Hilbert Action

We want to derive Einstein's equations from an action principle. We expect an action of the form

$$S[g, \mathrm{matter}] = \int_M L(g, \mathrm{matter}) \, \mathrm{dvol}_g,$$

where L is a scalar Lagrangian. We ignore matter for now. An obvious guess for L is $L = R$, the scalar curvature, in which case we get the *Einstein-Hilbert action*:

$$S_{\mathrm{EH}}[g] = \frac{1}{16\pi} \int_M R_g \, \mathrm{dvol}_g.$$

In order to derive the equations of motion from an action, we consider $g + \delta g$, where δg vanishes outside a compact set in M. Expanding to first order in δg, we require the variation of S_{EH} to vanish at this order, hence we want to compute

$$\delta S_{\mathrm{EH}} = S_{\mathrm{EH}}[g + \delta g] - S_{\mathrm{EH}}[g].$$

First we consider

$$\mathrm{dvol}_g = \sqrt{|g|} \, \mathrm{d}x^1 \dots \mathrm{d}x^n.$$

To compute $\delta|g|$, recall that we can write the determinant as

$$g = g_{\bar{\mu}\nu} \Delta^{\bar{\mu}\nu}.$$

Here $\bar{\mu}$ is a fixed value with no sum, and we sum over ν. $\Delta^{\mu\nu}$ is the cofactor matrix of g, where we delete the ν'th column and the μ'th row.

Hence $\Delta^{\mu\nu}$ is independent of $g_{\mu\nu}$, and moreover

$$\Delta^{\mu\nu} = g g^{\mu\nu}.$$

Hence under $g_{\mu\nu} \to g_{\mu\nu} + \delta g_{\mu\nu}$,

$$\delta g = \frac{\partial g}{\partial g_{\mu\nu}} \delta g_{\mu\nu} = \Delta^{\mu\nu} \delta g_{\mu\nu} = g g^{\mu\nu} \delta g_{\mu\nu},$$

so applying this to the square root,

$$\delta \sqrt{-g} = \delta \sqrt{|g|} = \frac{1}{2} \frac{1}{\sqrt{-g}} (-\delta g) = \frac{1}{2} \sqrt{-g} g^{\mu\nu} \delta t_{\mu\nu},$$

and so in AIN,

$$\delta(\mathrm{dvol}_g) = \frac{1}{2} g^{ab} \delta g_{ab} \, \mathrm{dvol}_g.$$

To compute δR_g, we first consider $\delta \Gamma_\nu{}^\mu{}_\rho$. The difference of two connections is a tensor, so this quantity is a tensor $\delta \Gamma_a{}^b{}_c$.

To compute this, we can consider normal coordinates for g_{ab} at p. Then since $\partial_\mu g_{\nu\sigma}|_p = 0$,

$$\delta \Gamma_\nu{}^\mu{}_\rho = \frac{1}{2} g^{\mu\sigma} (\delta g_{\sigma\nu,\rho} + \delta g_{\rho\sigma,\nu} - \delta g_{\nu\rho,\sigma}) \Big|_p$$

$$= \frac{1}{2} g^{\mu\sigma} (\delta g_{\sigma\nu;\rho} + \delta g_{\rho\sigma;\nu} - \delta g_{\nu\rho;\sigma}) \Big|_p$$

$$\delta \Gamma_b{}^a{}_c = \frac{1}{2} g^{ad} (\delta g_{db;c} + \delta g_{cd;b} - \delta g_{bc;d}),$$

as this is a tensor. Next we consider $\delta R^\mu{}_{\nu\rho\sigma}$. Again, working in normal coordinates at p,

$$R^\mu{}_{\nu\rho\sigma} = \partial_\rho(\Gamma_\nu{}^\mu{}_\sigma) - \partial_\sigma(\Gamma_\nu{}^\mu{}_\rho) + \Gamma^2$$

$$\delta R^\mu{}_{\nu\rho\sigma} = [(\partial_\rho(\delta \Gamma_\nu{}^\mu{}_\sigma) - \partial_\sigma(\delta \Gamma_\nu{}^\mu{}_\rho)] \Big|_p$$

$$= [\nabla_p \delta \Gamma_\nu{}^\mu{}_\sigma - \nabla_\sigma \delta \Gamma_\nu{}^\mu{}_\rho] \Big|_p,$$

$$\delta R^a{}_{bcd} = \nabla_c \delta \Gamma_b{}^a{}_d - \nabla_d \delta \Gamma_b{}^a{}_c,$$

$$\implies \delta R_{ab} = \nabla_c \delta \Gamma_a{}^c{}_b - \nabla_b \delta \Gamma_a{}^c{}_c,$$

Observe that

$$(\delta g^{ab} g_{bc}) = 0 \implies (\delta g^{ab}) = -g^{ac} g^{bd} \delta g_{cd}.$$

We finally have that

$$\delta R_g = \delta(g^{ab} R_{ab}) = (\delta g^{ab}) R_{ab} = (\delta g^{ab}) R_{ab} + g^{ab} \delta R_{ab}$$

$$= R^{ab} \delta g_{ab} + g^{ab} (\nabla_c \delta \Gamma_a{}^c{}_b - \nabla_b \delta \Gamma_a{}^c{}_c)$$

$$= -R^{ab} \delta g_{ab} + \nabla_c X^c,$$

where
$$X^c = g^{ab}\delta\Gamma_{a\ b}^{\ c} - g^{cb}\delta\Gamma_{b\ a}^{\ a}.$$

We deduce from the formulae for δR_g and $\delta\,\mathrm{dvol}_g$ that

$$\delta S_{\mathrm{EH}} = \frac{1}{16\pi}\int_M \left(\frac{1}{2}g^{ab}R - R^{ab}\right)\delta g_{ab} + \nabla_c X^c\,\mathrm{dvol}_g.$$

Since g vanishes outside some region, using Stokes' theorem on the last part, it vanishes. Hence this is

$$\frac{1}{16\pi}\int_M -G^{ab}\delta g_{ab}\,\mathrm{dvol}_g,$$

where we have used the fact that δg, and hence X, vanish on ∂M. We immediately see that $\delta S_{\mathrm{EH}} = 0$ for all variation δg_{ab} if g_{ab} solves the vacuum Einstein equations, i.e. $G^{ab} = 0$.

Suppose we also have a contribution from matter fields:

$$S_{\mathrm{tot}} = S_{\mathrm{EH}} + S_{\mathrm{MATTER}},$$

where

$$S_{\mathrm{MATTER}} = \int_M L[\Phi, g]\,\mathrm{dvol}_g.$$

Under a variation $g \to g + \delta g$, we must have that

$$\delta S_{\mathrm{MATTER}} = \frac{1}{2}\int_M T^{ab}\delta g_{ab}\,\mathrm{dvol}_g,$$

where T is some symmetric two-tensor (if T is not symmetric, symmetrize it). Varying g, we see that

$$G^{ab} = 8\pi T^{ab},$$

which in other words are Einstein's equations.

Example 5.5.

If ψ is a scalar and
$$L_{\mathrm{MATTER}} = -\frac{1}{2}g^{ab}\nabla_a\psi\nabla_b\psi,$$

then under $g \to g + \delta g$,

$$\delta S_{\text{MATTER}} = -\int_M \left[\frac{1}{2} \delta(g^{ab}) \nabla_a \psi \nabla_b \psi \, \text{dvol} - g + \frac{1}{2} g^{ab} \nabla_a \psi \nabla_b \psi \delta(\text{dvol}_g) \right]$$

$$= \frac{1}{2} \int_M \left(g^{ac} g^{bd} \nabla_c \psi \nabla_d \psi \delta g_{ab} - \frac{1}{2} (g^{cd} \nabla_c \psi \nabla_d \psi) g^{ab} \delta g_{ab} \right) \text{dvol}_g$$

$$= \frac{1}{2} \int_M T^{ab} \delta g_{ab} \, \text{dvol}_g,$$

where

$$T^{ab} = \nabla^a \psi \nabla^b \psi - \frac{1}{2} g^{ab} \nabla_c \psi \nabla^c \psi.$$

We can show that varying $\psi \to \psi + \delta \psi$ in the above that

$$\nabla_c \nabla^c \psi = 0.$$

It can be shown that diffeomorphism of the matter action implies that $\nabla_a T^{ab} = 0$.

5.9   Cauchy Problem for Einstein's Equations

We expect that Einstein's equation can be solved given data on a spacelike hypersurface. What is the right data?

Suppose that $\iota : \Sigma \to M$ is an embedding, such that $\iota(\Sigma)$ is spacelike. This means that $h = \iota^*(g)$ is Riemannian.

Let n be a choice of unit normal to $\iota(\Sigma)$. We define, for X, Y vector fields on Σ,

$$k(X, Y) = \iota^*(g(n, \nabla_{\tilde{X}} \tilde{Y})),$$

where $\iota_* X = \tilde{X}$, and $\iota_* Y = \tilde{Y}$. This is sort of the curvature.

We pick local coordinates $\{y^i\}$ on Σ, and $\{x^\mu\}$ on M such that $\iota : (y^1, y^2, y^3) \mapsto (0, y^1, y^2, y^3)$. Then $n_a \propto dx^0$, so say $n_\mu = \alpha \delta^0{}_\mu$.

If X, Y are vector fields on Σ, say that

$$X = X^i \frac{\partial}{\partial y^i}, \qquad Y = Y^i \frac{\partial}{\partial y^i}.$$

Take

$$\tilde{X} = X^i \frac{\partial}{\partial x^i}, \qquad \tilde{Y} = Y^i \frac{\partial}{\partial x^i}.$$

Then,

$$k(X, Y) = g_{\mu\nu} n^\mu \tilde{X}^\sigma \nabla_\sigma \tilde{Y}^\nu$$
$$= \alpha \delta^0{}_\nu \tilde{X}^\sigma (\partial_\sigma \tilde{Y}^\nu + \Gamma_\sigma{}^\nu{}_\tau \tilde{Y}^\tau) = \alpha \Gamma_i{}^0{}_j X^i Y^j.$$

Hence k is a symmetric 2-tensor on Σ. We can show that if (M, g) solves the vacuum Einstein equations, then the Einstein constraint equations hold:

$$\nabla^h_i k^i{}_j - \nabla^h_j k^i{}_i = 0, \qquad R_h - k^i{}_j k^j{}_i + k^i{}_i k^j{}_j = 0. \tag{\dagger}$$

Conversely, if we are given:

- Σ a 3-manifold,

- h a Riemannian metric on Σ,

- k a symmetric 2-tensor on Σ,

such that ($\dagger$) holds, then there exists a solution (M, g) of the vacuum Einstein equations, and an embedding $\iota : \Sigma \to M$ such that h is the induced metric on Σ, and k is the second fundamental form of $\iota(\Sigma)$.

There may be more than one solution to this problem. However it turns out on the overlap of two solutions, there is a diffeomorphism from one solution to another, and we can define a unique maximal solution in this way.

Index