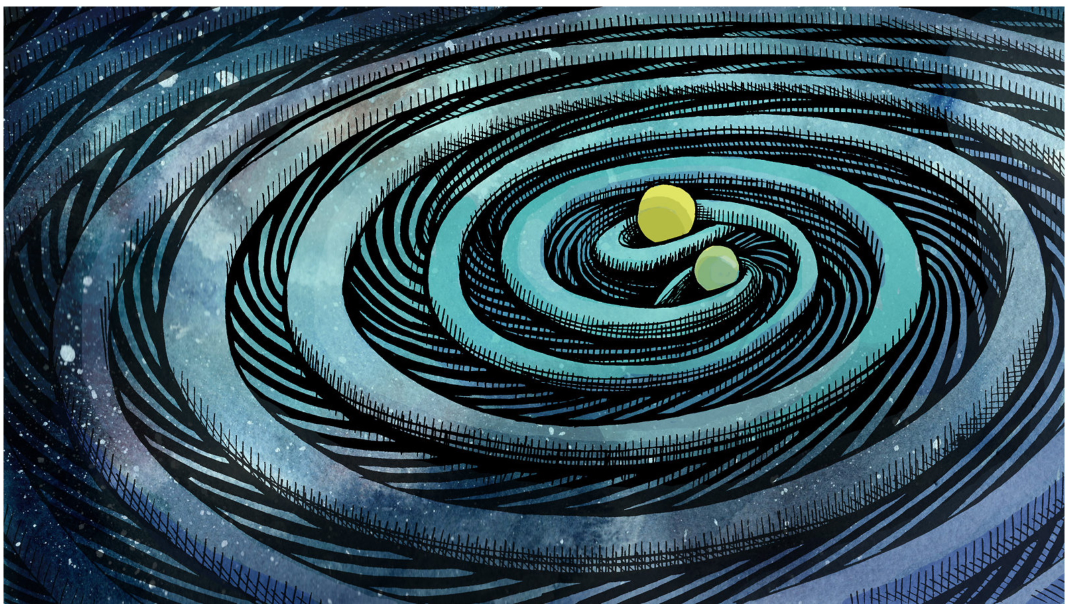
General Relativity - A Reading Course

Humberto Gilmer
Department of Physics
The Ohio State University



Preface

These are my reading course notes for general relativity, autumn 2017, with Dr. Samir Mathur as facilitator. The primary text used and referenced throughout is *Spacetime and Geometry* by Sean Carroll, with *A First Course in General Relativity* by Bernard Schutz, and *Einstein Gravity in a Nutshell* by Anthony Zee as secondary texts. This first pass is heavily coordinate-dependent. I hope to emphasize a geometric view on later passes, which will hopefully be added to these notes. Those sections denoted with $*$ may be omitted in first read-through. I am collecting them here for my purposes of understanding differential geometry.

Contents

Part I
Tensor calculus

1 Manifolds

The lynchpin of general relativity is the Einstein Equivalence Principle:

> In small enough regions of spacetime, the laws of physics reduce to those of special relativity; it is impossible to detect the existence of a gravitational field by means of local experiments.

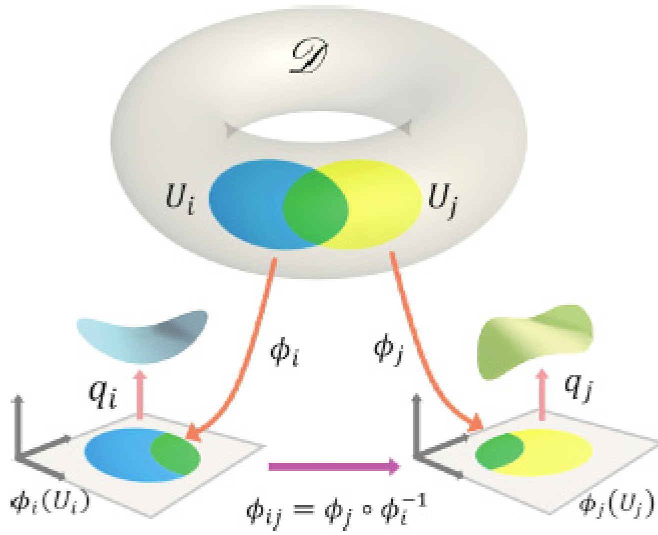
This means that the appropriate mathematical structure to describe curvature of which gravitation is a manifestation is a *differentiable manifold*: a set that looks locally like flat space but may have a different global geometry. A manifold of dimension n is constructed by smoothly "sewing together" different patches that each may be described by $\mathbb{R}^n$.

1.1 Definitions

We seek to formalize what is meant by "sewing together" and "described by $\mathbb{R}^n$". A *map* $\phi : M \to N$ is an operation that assigns to a value in M a corresponding value in N. A map may be *one-to-one* (or *injective*) meaning that each element in N that is mapped from M is mapped from at most one element in M or it may be *onto* (or *surjective*) meaning that every element of N is mapped from at least one element in M. (NB: onto and one-to-one (or one-from-one as Carroll suggests) are statements about the set N, not M). M is known as the *domain* of ϕ, while the set of points to which M is mapped in N is called the *image*. A smooth map is one which may be differentiated an infinite number of times, also known as a C^∞ map. Two sets M and N are *diffeomorphic* if there exists a C^∞ map with a C^∞ inverse; the map ϕ is called a diffeomorphism.

A *chart* or *coordinate system* consists of a subset U of a set M, along with a one-to-one map $\phi : U \to \mathbb{R}^n$, such that the image $\phi(U)$ is open in $\mathbb{R}^n$. An *atlas* is an indexed collection of charts $\{(U_\alpha, \phi_\alpha)\}$ that satisfies two conditions:

- The union of U_α is equal to M
- The charts are smoothly sown together, meaning if two charts overlap (meaning the sets U_α and U_β include some set of shared points, $U_\alpha \cap U_\beta \neq 0$), then the map $(\phi_\alpha \circ \phi_\beta^{-1})$ takes points in $\phi_\beta(U_\alpha \cap U_\beta) \subset \mathbb{R}^n$ onto an open set $\phi_\beta(U_\alpha \cap U_\beta) \subset \mathbb{R}^n$, and all of these maps must be C^∞.



Thus a C^∞ n-dimensional manifold is a set M along with a maximal atlas, one that contains every compatible chart.

1.2 Vectors

The whole point of the above song-and-dance was to establish a playing field for structures on manifolds. One such example is a vector, which lives in a tangent space. Usually when we think of vectors, we think of little arrows pointing from one point to another in space. Here we will dispense with this imagery. Instead, consider a tangent space, the space of all vectors associated with a single point on a manifold (spacetime, for example). Let's make this concrete.

Consider a point p on a manifold M. Let $\mathcal{F}$ be the space of all smooth functions on M (meaning the set of all C^∞ maps $f : M \to \mathbb{R}$). Now consider the set of all parameterized curves through p, meaning the set of all non-degenerate maps $\gamma : \mathbb{R} \to M$ such that p is in the image of γ. Therefore, the tangent space T_p may be identified with the space of directional derivative operators along curves through p. A directional derivative is a map $d : f \to \frac{df}{d\lambda}\big]_p$. It can be shown that the space of directional derivatives is a vector space and that it satisfies the requirements of the tangent space (dimensionality, etc.) [1].

So now we have the tangent space T_p. What is the basis for this space? Consider an n-manifold M with coordinate chart $\phi : M \to \mathbb{R}^n$, a curve $\gamma : \mathbb{R} \to M$ and a function

[1] see Carroll, pg. 63-65 for details

$f: M \to \mathbb{R}$ and let λ be the parameter for γ. Then the chain rule gives

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

Since f is arbitrary then

$$\frac{d}{d\lambda} = \frac{dx^{\mu}}{d\lambda}\partial_{\mu} \qquad \Rightarrow \qquad d = dx^{\mu}\partial_{\mu}$$

Thus the ∂_{μ} form a basis of T_p which point along the coordinate axes, known as a *coordinate basis*. A vector V therefore, may be written

$$V = V^{\mu}\partial_{\mu}$$

This feels…incomplete, somehow. Of course, if I slap a function f on that so we have

$$Vf = V^{\mu}\partial_{\mu} f$$

it's not so bad! But then, this indicates how a vector is to be viewed now. Instead of a little arrow connecting two points, it's an arrow that points in a direction that is a linear combination of directional derivatives (which lie tangent to curves passing through the point p), with the linear combination specified by V^{μ}! Thus, a vector is no longer an arrow but actually a map $\mathcal{F}: f \to df$ from the space of functions to the space of directional derivatives.

Of course, this is not the only basis for the space; it is always possible to execute a change of coordinates (perhaps by considering a different parameterization of the curves γ through p). With this language, a change of basis takes the form

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

A vector V should be left unchanged by a coordinate transformation; the same cannot be said for the components:

$$V = V^{\mu}\partial_{\mu} = V^{\mu}\frac{\partial x^{\mu'}}{\partial x^{\mu}}\partial_{\mu'} = V^{\mu'}\partial_{\mu'} = V$$

which is true if

$$V^{\mu'} = V^{\mu}\frac{\partial x^{\mu'}}{\partial x^{\mu}}$$

Thus a change in coordinates defines a change of basis. Now that we know what a vector is, we proceed to defining a vector field. Recall that a field is an operator that assigns to each point an object; so a vector field assigns to each point p in M a vector V.

1.3 Dual vectors

A dual vector (think a Bra in Bra-ket notation) is an object that accepts as input a vector and produces a scalar. Dual vectors, like vectors, exist in a linear space tangent to the manifold at a point p; in this case, however, they exist not in the tangent space T_p but in the cotangent space T_p^*. The cotangent space is nothing more than the collection of maps from the tangent space to the real numbers; thus a particular dual vector ω has the action $\omega : T_p \to \mathbb{R}$. This will be illustrated explicitly once we have a coordinate basis for the dual vectors.

The canonical example of a dual vector is the gradient operator d which maps from the space of functions to the space of dual vectors $\mathrm{d} : \mathcal{F} \to T_p^*$. It takes in a vector $\frac{d}{d\lambda}$ and produces a directional derivative (a scalar):

$$\mathrm{d}f\left(\frac{d}{d\lambda}\right) = \frac{df}{d\lambda}$$

(NB: d and d are deliberately different! d is the gradient or boundary operator which will be discussed at the end of this section). Recall from the previous subsection that a basis for the tangent space is provided by the partial derivatives ∂_μ along coordinate functions x^μ; similarly, gradients of these coordinate functions $\mathrm{d}x^\mu$ provide a basis for the cotangent space:

$$\mathrm{d}x^\mu(\partial_\nu) = \frac{\partial x^\mu}{\partial x^\nu} = \delta^\mu_\nu$$

A dual vector may be thought of as a set of oriented surfaces in space. The magnitude of the dual vector corresponds to the inverse spacing of the surfaces; the higher the magnitude the smaller the spacing and vice-versa. In this sense, a dual vector may be thought of as a wave frequency, while a vector corresponds to a wavelength. Their product produces a velocity. In like manner, the product of a vector and dual vector produces the directional derivative. An inner product may therefore be visualized as the number of times a vector 'pierces' the dual vector's surfaces.

Now that we have a basis for dual vectors, it is possible to decompose an arbitrary dual vector ω into components

$$\omega = \omega_\mu \mathrm{d}x^\mu$$

Note that this definition is just as open-ended as the one for vectors! Except here, the missing piece is the vector V that would be fed into the dual vector to get a number:

$$\omega(V) = V(\omega) = \omega_\mu \mathrm{d}x^\mu(V^\nu \partial_\nu) = \omega_\mu V^\nu \mathrm{d}x^\mu(\partial_\nu) = \omega_\mu V^\nu \delta^\mu_\nu = \omega_\mu V^\mu$$

This provides the explicit example of $\omega : T_p \to \mathbb{R}$ promised earlier. Returning to the example of the directional derivative, it is possible to show the explicit form using the coordinate expansions above

$$\mathrm{d}f\,(V) = \partial_\mu f \mathrm{d}x^\mu\,(V^\nu \partial_\nu) = V^\nu \partial_\mu f \mathrm{d}x^\mu(\partial_\nu) = V^\nu \partial_\mu f \delta^\mu_\nu = V^\mu \partial_\mu f = Vf$$

which is exactly the directional derivative of f in the V direction.

Lastly, as with vectors, it is easy to express coordinate transformations in this language:

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

As with vectors, a dual vector should be independent of the coordinates used to express it; the same cannot be said for the components:

$$\omega = \omega_\mu \mathrm{d}x^\mu = \omega_\mu \frac{\partial x^\mu}{\partial x^{\mu'}}\mathrm{d}x^{\mu'} = \omega_{\mu'}\mathrm{d}x^{\mu'}$$

where

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

Lastly, as with vectors, a dual vector field assigns to each point p in M a dual vector ω.

1.4 Tensors

Vectors and one-forms can be thought of as special cases of a tensor. In general, a $\binom{m}{n}$ tensor is a multilinear map from a collection of m dual vectors and n vectors to $\mathbb{R}$. In other words

$$T : \underbrace{T_p^* \times \cdots \times T_p^*}_{(m \text{ times})} \times \underbrace{T_p \times \cdots \times T_p}_{(n \text{ times})} \to \mathbb{R}$$

All $\binom{m}{n}$ tensors form a vector space, meaning they can be added together and multiplied by real numbers. It is also possible to multiply them together using the tensor product. Suppose T is a $\binom{k}{l}$ tensor and S is a $\binom{m}{n}$ tensor; then $T \otimes S$ is a $\binom{m+k}{n+l}$ tensor defined by

$$\begin{aligned}&T \otimes S(\omega^{(1)}, \ldots, \omega^{(k)}, \ldots \omega^{(k+m)}, V^{(1)}, \ldots, V^{(l)}, \ldots, V^{(l+n)}) \\ &= T(\omega^{(1)}, \ldots, \omega^{(k)}, V^{(1)}, \ldots, V^{(l)}) \times S(\omega^{(k+1)}, \ldots, \omega^{(k+m)}, V^{(l+1)}, \ldots, V^{(l+n)})\end{aligned}$$

Tensor products in general do not commute, so $T \otimes S \neq S \otimes T$. As with vectors and dual vectors, it is possible to decompose a $\binom{m}{n}$ tensor using the coordinate bases defined previously:

$$T = T^{\mu_1 \ldots \mu_n}{}_{\nu_1 \ldots \nu_m} \partial_{\mu_1} \otimes \cdots \otimes \partial_{\mu_n} \otimes \mathrm{d}x^{\nu_1} \otimes \cdots \otimes \mathrm{d}x^{\nu_m}$$

The transformation properties of the components also carry over from vectors and dual vectors; namely, each μ_i coordinate transforms as the components of a vector and each ν_j component transforms as the components of a dual vector.

Recall that vectors and dual vectors are special cases of tensors; in particular, they are $\binom{0}{1}$ and $\binom{1}{0}$ tensors, respectively. That means they are restricted in both input and output, always producing a scalar. A general tensor has no such restriction. An $\binom{m}{n}$ tensor can be used to map to other types of tensor, depending on the number of inputs fed into the tensor. Suppose you have a $\binom{1}{1}$ tensor, into which you fed a single dual vector; the result would be a $\binom{0}{1}$ tensor, a vector!

Further such manipulations are possible, namely the operations of contraction, index raising/lowering and symmetrization/anti-symmetrization. Contraction turns a $\binom{m}{n}$ tensor into a $\binom{m-1}{n-1}$ tensor; in index notation, this is done by summing over one upper and one lower index:

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

Indices may be raised or lowered by means of the metric. The metric $\eta^{\mu\nu}$ is a $\binom{0}{2}$ tensor and will be discussed in the following section; for now its main purpose is to take in a $\binom{m}{n}$ tensor and change it into a $\binom{m-1}{n+1}$ tensor:

$$T^{\alpha\beta\mu}{}_{\delta} = \eta^{\mu\gamma} T^{\alpha\beta}{}_{\gamma\delta} \qquad T_{\mu}{}^{\beta}{}_{\gamma\delta} = \eta_{\mu\alpha} T^{\alpha\beta}{}_{\gamma\delta} \qquad T_{\mu\nu}{}^{\rho\sigma} = \eta_{\mu\alpha}\eta_{\nu\beta}\eta^{\rho\gamma}\eta^{\sigma\delta} T^{\alpha\beta}{}_{\gamma\delta}$$

Symmetrization (or anti-symmetrization) refers to the order in which arguments are fed into the tensor (equivalently, the order of the indices). A tensor symmetric in two given indices doesn't care when those indices are swapped (or which order those particular inputs are fed); a fully symmetric tensor is symmetric in all its indices. An antisymmetric tensor undergoes a sign change when two indices are swapped; similarly a fully antisymmetric tensor is antisymmetric in all pairs of its indices. A tensor may be symmetrized:

$$T_{(\mu_1 \dots \mu_n)}{}^{\rho}{}_{\sigma} = \frac{1}{n!}\left(T_{\mu_1 \dots \mu_n}{}^{\rho}{}_{\sigma} + \text{sum over permutations of } \mu_1 \dots \mu_n\right)$$

or antisymmetrized

$$T_{[\mu_1 \dots \mu_n]}{}^{\rho}{}_{\sigma} = \frac{1}{n!}\left(T_{\mu_1 \dots \mu_n}{}^{\rho}{}_{\sigma} + \text{alternating sum over permutations of } \mu_1 \dots \mu_n\right)$$

Lastly, a tensor may be decomposed into symmetric and antisymmetric parts in any two given indices

$$T_{\mu\nu\rho\sigma} = T_{(\mu\nu)\rho\sigma} + T_{[\mu\nu]\rho\sigma}$$

1.5 Tensor densities

It is possible to generalize the notion of a tensor to that of a tensor density. A tensor density is an object that transforms like a tensor, except for a factor proportional to the Jacobian determinant that arises due to the transformation; the power to which this factor is raised is known as the weight. Let T be a $\binom{m}{n}$ tensor density of weight w; then the transformation rule of the components of T is

$$T^{\alpha_1 \dots \alpha_m}{}_{\beta_1 \dots \beta_n} = \left|\det\left(\frac{\partial y^{\alpha}}{\partial x^{\mu}}\right)\right|^{w} \frac{\partial y^{\alpha_1}}{\partial x^{\mu_1}} \cdots \frac{\partial y^{\alpha_m}}{\partial x^{\mu_m}} \frac{\partial x^{\nu_1}}{\partial y^{\beta_1}} \cdots \frac{\partial x^{\nu_n}}{\partial y^{\beta_n}} T^{\mu_1 \dots \mu_m}{}_{\nu_1 \dots \nu_n}$$

The canonical example of a tensor density is the Levi-Civita symbol, $\tilde{\epsilon}_{\mu_1\mu_2 \dots \mu_n}$:

$$\tilde{\epsilon}_{\alpha_1 \dots \alpha_m} = \left|\det\left(\frac{\partial y^{\alpha}}{\partial x^{\mu}}\right)\right| \frac{\partial x^{\nu_1}}{\partial y^{\beta_1}} \cdots \frac{\partial x^{\nu_n}}{\partial y^{\beta_n}} \tilde{\epsilon}_{\nu_1 \dots \nu_n}$$

where, since $w = 1$, it is a density of weight 1. The determinant of the metric tensor g also transforms as a tensor density of weight -2

$$g = \left|\det\left(\frac{\partial y^{\alpha}}{\partial x^{\mu}}\right)\right|^{-2} g$$

This gives a simple way to transform a tensor density into a tensor, by multiplying the density by an appropriate factor of g to cancel the weighting factor. The general rule is multiplication by $g^{\frac{w}{2}}$ suffices to cancel the factor. Thus a Levi-Civita tensor can be made out of the Levi-Civita symbol as follows:

$$\epsilon^{\mu_1 \cdots \mu_m} = \sqrt{|g|}\tilde{\epsilon}_{\mu_1 \cdots \mu_m}$$

Note that indices may not be freely raised or lowered on a tensor density without consequence. The Levi-Civita symbol with raised indices ($w = -1$) is actually different from the Levi-Civita symbol with lowered indices ($w = 1$).

1.6 Differential forms

Forms constitute a special class of tensors with specific properties. A p-form is a $\binom{0}{p}$ tensor that is completely antisymmetric. Scalars are automatically 0-forms, and dual vectors are automatically 1-forms, in all dimensions. On an n-dimensional vector space, the number of p-forms is given by $\frac{n!}{p!(n-p)!}$; thus in 4 dimensions there is 1 0-form, 4 1-forms, six 2-forms, 4 3-forms and 1 4-form. Forms are nice to work with b/c they may be differentiated and integrated without reference to additional geometric information (completely coordinate-free and independent of metric, including the case that one does not exist).

Forms may be multiplied. This is known as the wedge product. Let A be a p-form and B be a q-form; then $A \wedge B$ is a $(p+q)$-form

$$(A \wedge B)_{\mu_1 \cdots \mu_{p+q}} = \frac{(p+q)!}{p!q!} A_{[\mu_1 \cdots \mu_p} B_{\mu_{p+1} \cdots \mu_{p+q}]}$$

Since we want the product of forms to itself be a form, the multiplication must be antisymmetric:

$$A \wedge B = (-1)^{pq} B \wedge A$$

Forms may also be differentiated. The exterior derivative d is an operator that takes p-forms to $(p+1)$-forms. It is defined as a normalized, antisymmetrized partial derivative. Let ω be a p-form; then $\mathrm{d}\omega$ is given by

$$(\mathrm{d}\omega)_{\mu_1 \cdots \mu_p \mu_{p+1}} = (p+1)\partial_{[\mu_{p+1}} \omega_{\mu_1 \cdots \mu_p]}$$

The exterior derivative satisfies a modified form of the Leibniz rule. Let ω be a p-form and η be a q-form. Then the exterior derivative of $\omega \wedge \eta$ is given by

$$\mathrm{d}(\omega \wedge \eta) = (\mathrm{d}\omega) \wedge \eta + (-1)^p \omega \wedge (\mathrm{d}\eta)$$

Unlike the partial derivative, the exterior derivative is always a tensor. Consider $\partial_\mu W_\nu$, the components of a partial derivative of a form in one coordinate set x^ν; transforming this to another coordinate system $x^{\nu'}$ gives the following transformation rule

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

Of these terms, the second is the typical tensorial transformation rule; the first is an exception that must be dealt with, if the object $\partial_\mu W_\nu$ is to transform tensorially. Recall that the exterior derivative is an antisymmetrized partial derivative. The offending term may be written as

$$\frac{\partial x^\mu}{\partial x^{\mu'}}\left(\frac{\partial}{\partial x^\mu}\frac{\partial x^\nu}{\partial x^{\nu'}}\right)W_\nu = \frac{\partial x^\mu}{\partial x^{\mu'}}\left(\frac{\partial}{\partial x^{\nu'}}\frac{\partial x^\nu}{\partial x^\mu}\right)W_\nu = \frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial}{\partial x^{\nu'}}\delta^\nu_\mu W_\nu = \frac{\partial^2 x^\nu}{\partial x^{\mu'}\partial x^{\nu'}}W_\nu$$

which is symmetric under $\mu' \leftrightarrow \nu'$ interchange. Thus when the partial derivative is properly antisymmetrized, the above term will vanish identically, leaving a tensorial object. The exterior derivative is nilpotent, meaning that given an arbitrary form A, $\mathrm{d}(\mathrm{d}A) = \mathrm{d}^2A = 0$ identically. This can be read as "the boundary of a boundary is 0". A p-form A is *closed* if $\mathrm{d}A = 0$; it is *exact* if $A = \mathrm{d}B$ where B is some $(p-1)$-form. Clearly, all exact forms are closed, but the converse is not true.

The final operation to be considered on forms is that of the *Hodge dual.* The Hodge star operator $*$ on an n-manifold M is a map

$$* : \Lambda^p(M) \to \Lambda^{n-p}(M)$$

In other words, a map from the space of p-forms $\Lambda^p(M)$ to the space of $(n-p)$-forms $\Lambda^{n-p}(M)$. The Hodge dual's action on components is

$$(*A)_{\mu_1\cdots\mu_{n-p}} = \frac{1}{p!}\epsilon^{\nu_1\ldots\nu_p}{}_{\mu_1\cdots\mu_{n-p}}A_{\nu_1\cdots\nu_p}$$

Notice the presence of the Levi-Civita tensor, with both raised and lowered indices. This indicates that the Hodge dual depends on a metric of some sort (both to generate the Levi-Civita tensor from the corresponding tensor density and to raise and lower the appropriate indices). The notion of duality merits some explanation. A duality is an operation that, when performed twice, brings the object upon back to itself, up to a sign. For the Hodge dual, this takes the form

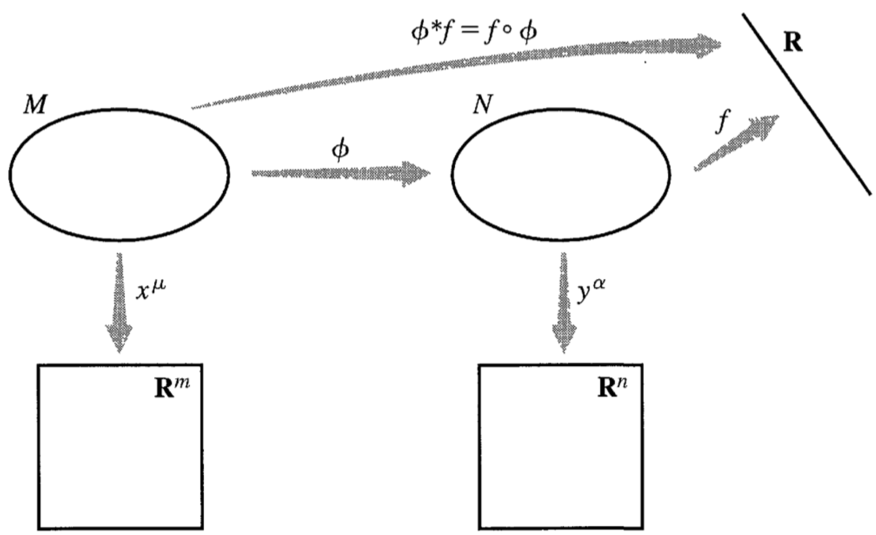
$$*(*A) = (-1)^{s+p(n-p)}A$$

where s is the number of -1 in the metric. Notice that here the Hodge dual's explicit dependence on the metric is clear. Dualities between vector spaces require that the original and transformed spaces have the same dimensionality.

2 Calculus on Manifolds

2.1 Maps between Manifolds*

Consider two manifolds M and N, possibly of different dimension, with coordinates $x^\mu : M \to \mathbb{R}^m$ and $y^\alpha : N \to \mathbb{R}^n$ respectively. Now suppose we have a map $\phi : M \to N$ and a function $f : N \to \mathbb{R}$. We can naturally construct a function $g : M \to \mathbb{R}$ by composing these maps. In other words $g = (f \circ \phi) : M \to \mathbb{R}$. This defines what is known as the *pullback* of f by ϕ, $\phi^* f = (f \circ \phi)$



Note that there is no way to pushforward a function $h : M \to \mathbb{R}$ to N. However, it is possible to pushforward vectors. How so? Suppose we have a vector V at a point p in M. Recall that a vector maps from the space of smooth functions to space of directional derivatives of f, the reals, $V : \mathcal{F}(M) \to \mathbb{R}$. Now p is mapped to $\phi(p) \in N$. Therefore ϕ "pushes" V to a map from $\mathcal{F}(N) \to \mathbb{R}$, $\phi_*(V) : \mathcal{F}(N) \to \mathbb{R}$. In our notation, this becomes

$$\phi_* V(f) = V(\phi^* f)$$

Quoting Carroll, "the action of $\phi_* V$ on any function is simply the action of V on the pullback of that function." With this in mind, recall that V is an element of the tangent space at p, $V \in T_p M$; it is a map from the manifold to the tangent space, $V : M \to T_p M$. So what is the action of the pushforward in this picture? In other words, where does $\phi_* V$ reside and what type of map is it? Quite simply, $\phi_* V$ must be a vector living in a tangent space of N, at the point $\phi(p) \in N$, $\phi_* V \in T_{\phi(p)} N$; furthermore, it must now be a map from N to the tangent space of N, $\phi_* V : N \to T_{\phi(p)} N$

There is a simple matrix representation of the above operation. The basis vectors on M are given by ∂_μ; on N they are given by ∂_α. Thus we'd like to find the relation

$$V = V^\mu \partial_\mu \stackrel{?}{\to} \phi_* V = (\phi_* V)^\alpha \partial_\alpha$$

In other words, what is the relation between V^μ and $(\phi_* V)^\alpha$? This can be done as follows:

$$\begin{aligned}
\phi_* V(f) &= V(\phi^* f)\\
(\phi_* V)^\alpha \partial_\alpha f &= V^\mu \partial_\mu (\phi^* f)\\
&= V^\mu \partial_\mu (f \circ \phi)\\
&= V^\mu \frac{\partial y^\alpha}{\partial x^\mu} \partial_\alpha f
\end{aligned}$$

This implies that $(\phi_* V)^\alpha = V^\mu \frac{\partial y^\alpha}{\partial x^\mu}$ and therefore the pushforward has the form $(\phi_*)^\alpha{}_\mu = \frac{\partial y^\alpha}{\partial x^\mu}$. Note that since $\{\mu\} \neq \{\alpha\}$, that the pushforward needn't be a square matrix (and thus may not be invertible). This looks just like a coordinate transformation, which it is; in fact it's a generalization, which will be explored further in the next section.

The logic above holds for dual vectors, but with slight modifications. Vectors are maps from a manifold to its tangent space; dual vectors are maps from the tangent space to the real numbers, similar to functions. Vectors can be pushed forward; dual vectors are pulled back, like functions. Let ω be a dual vector at a point $q = \phi(p)$ in N, $\omega \in T^*_{\phi(p)} N$; furthermore, ω is a map from the tangent space at $\phi(p)$ to the reals, $\omega : T_{\phi(p)} N \to \mathbb{R}$. The pullback of ω, $\phi^* \omega$ must therefore be an element of the cotangent space at p of M, $\phi^* \omega \in T^*_p M$, and be a map from the tangent space at p of M to the reals, $\phi^* \omega : T_p M \to \mathbb{R}$. As with the pushforward of vectors, the pullback of dual vectors may be identified with its action on vectors:

$$\phi^* \omega(V) = \omega(\phi_* V)$$

Similarly to the case with vectors, a matrix interpretation of the above expression yields a view of the pullback as a coordinate transformation. Let $\omega = \omega_\alpha \mathrm{d}y^\alpha$ and $\phi^* \omega = (\phi^* \omega)_\mu \mathrm{d}x^\mu$. Therefore, the above expression may be written as

$$\begin{aligned}
\phi^* \omega(V) &= \omega(\phi_* V)\\
(\phi^* \omega)_\mu \mathrm{d}x^\mu (V^\nu \partial_\nu) &= \omega_\alpha \mathrm{d}y^\alpha \left(V^\mu \frac{\partial y^\beta}{\partial x^\mu} \partial_\beta \right)\\
(\phi^* \omega)_\mu V^\nu \mathrm{d}x^\mu (\partial_\nu) &= \omega_\alpha V^\mu \frac{\partial y^\beta}{\partial x^\mu} \mathrm{d}y^\alpha (\partial_\beta)\\
(\phi^* \omega)_\mu V^\nu \delta^\mu_\nu &= \omega_\alpha V^\mu \frac{\partial y^\beta}{\partial x^\mu} \delta^\alpha_\beta\\
(\phi^* \omega)_\mu V^\mu &= \omega_\alpha \frac{\partial y^\alpha}{\partial x^\mu} V^\mu
\end{aligned}$$

Since the vector V (and thus its components V^μ) is arbitrary, equating the coefficients of both sides above yields $(\phi^* \omega)_\mu = \omega_\alpha \frac{\partial y^\alpha}{\partial x^\mu}$, meaning the pullback has components $(\phi^*)_\mu{}^\alpha$

Lastly, it is possible to generalize the above discussion to apply to $\binom{0}{n}$ and $\binom{m}{0}$ tensors. Recall that a $\binom{0}{n}$ tensor is the generalization of a dual vector (takes in n vectors rather than just 1), while a $\binom{m}{0}$ tensor is a generalization of a vector (takes in m dual vectors instead of just 1). The behavior of these tensors is exactly like that of their rank-1 counterparts. Let T be a $\binom{0}{n}$ tensor. Like a dual vector, it may be pulled back by ϕ^*

$$(\phi^* T)(V^{(1)}, \ldots, V^{(n)}) = T(\phi^* V^{(1)}, \ldots, \phi^* V^{(n)})$$

Thus the action of the pullback on T is equivalent to acting with T on the pushed forward vectors. Now let S be a $\binom{m}{0}$ tensor. Like a vector, it may be pushed forward by ϕ_*

$$(\phi_* S)(\omega^{(1)}, \ldots, \omega^{(m)}) = S(\phi^* \omega^{(1)}, \ldots, \phi^* \omega^{(m)})$$

Thus the action of the pushforward on S is equivalent to acting with S on the pulled back dual vectors. In component notation, the above rules become

$$(\phi^* T)_{\mu_1 \ldots \mu_n} = \frac{\partial y^{\alpha_1}}{\partial x^{\mu_1}} \cdots \frac{\partial y^{\alpha_n}}{\partial x^{\mu_n}} T_{\alpha_1 \ldots \alpha_n} \qquad\qquad (\phi_* S)_{\mu_1 \ldots \mu_m} = \frac{\partial y^{\alpha_1}}{\partial x^{\mu_1}} \cdots \frac{\partial y^{\alpha_m}}{\partial x^{\mu_m}} S_{\alpha_1 \ldots \alpha_n}$$

Note that the pushforward and pullback of arbitrary $\binom{m}{n}$ tensors is not in general possible.

2.2 Diffeomorphisms and Lie Derivatives

In the special case that M and N above are the same tensor, several nice properties emerge. In this case, if ϕ (and its inverse ϕ^{-1}) is smooth, are invertible, it is known as a diffeomorphism between M and N. In this case, both ϕ and ϕ^{-1} may be used to move tensors from M to N; furthermore, in this case it is possible to pushforward and pullback arbitrary $\binom{m}{n}$ tensors. The pushforward of a $\binom{m}{n}$ tensor field on M is defined by

$$(\phi_* T)(\omega^{(1)}, \ldots, \omega^{(m)}, V^{(1)}, \ldots, V^{(n)}) = T(\phi^* \omega^{(1)}, \ldots, \phi^* \omega^{(m)}, [\phi^{-1}]_* V^{(1)}, \ldots, [\phi^{-1}]_* V^{(n)})$$

In other words, the pushforward of tensor field M by ϕ, where it acts on dual vectors and vectors on N, is equivalent to acting with T on dual vectors pulled back to M by ϕ and on vectors pushed forward to M by ϕ^{-1}. In component notation, this takes the form

$$(\phi_* T)^{\alpha_1 \ldots \alpha_m}{}_{\beta_1 \ldots \beta_n} = \frac{\partial y^{\alpha_1}}{\partial x^{\mu_1}} \cdots \frac{\partial y^{\alpha_m}}{\partial x^{\mu_m}} \frac{\partial x^{\nu_1}}{\partial y^{\beta_1}} \cdots \frac{\partial x^{\nu_n}}{\partial y^{\beta_n}} T^{\mu_1 \ldots \mu_m}{}_{\nu_1 \ldots \nu_n}$$

Now recall that in this case $M = N$; y^α and x^μ are not necessarily the same coordinate functions, but their dimensionality is the same (meaning $\{\mu\} = \{\alpha\}$). Thus switching between x^μ and y^α amounts to a coordinate transformation; but wait, doesn't ϕ also amount to a coordinate transformation? The answer is yes; switching between x^μ and y^α is a passive coordinate transformation, while mapping under ϕ, $(\phi^* x)^\mu : M \to \mathbb{R}^n$ amounts to an active coordinate transformation.

Let's suppose we wish to compare the value of a tensor T at p and the value of the same tensor, evaluated at $\phi(p)$ and pulled back to $\phi^*[T(\phi(p))]$. In other words, what is the relation

$$\Delta T(p) = \phi^*[T(\phi(p))] - T(p)$$

What is the rate of change of the tensor under the diffeomorphism ϕ? Instead of a finite diffeomorphism, consider instead an infinitesimal diffeomorphism ϕ_t, parameterized by t. When $t = 0$, $\phi_{t=0} = I$, the identity map. These infinitesimal diffeomorphisms may be composed; thus $\phi_t \circ \phi_s = \phi_{s+t}$; thus the finite diffeomorphism ϕ can simply be considered the sum of a series of infinitesimal diffeomorphisms. Infinitesimal diffeomorphisms may be thought of as arising from a vector field. For a vector field $V^\mu(x)$, there is a family of curves that are tangent to V at each x; these *integral curves* satisfy the relation

$$V^\mu = \frac{dx^\mu}{dt}$$

Each ϕ_t represents a diffeomorphism "flowing down the integral curves" of the vector field V^μ; this vector field V^μ is known as the generator of the diffeomorphism. In other words, ϕ_t is a diffeomorphism, flowing along V^μ and ϕ_s is a different diffeomorphism, flowing along a different vector W^μ. Thus the rate of change defined earlier depends on which diffeomorphism is used:

$$\Delta_t T(p) = \phi_t^*[T(\phi_t(p))] - T(p)$$

The above hints at a new derivative: the Lie derivative defined as

$$\mathcal{L}_V T = \lim_{t\to 0}\left(\frac{\Delta_t T(p)}{t}\right)$$

The Lie derivative is linear

$$\mathcal{L}_V(aT + bS) = a\mathcal{L}_V T + b\mathcal{L}_V S$$

and obeys the Leibniz rule

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

Let's examine the action of the Lie derivative on functions, vectors and one forms. The Lie derivative on a function reduces to the ordinary direction derivative:

$$\mathcal{L}_V f = V(f) = V^\mu \partial_\mu f$$

The Lie derivative on vectors is simply the Lie bracket

$$\mathcal{L}_V U^\mu = [V,U]^\mu = V^\nu \partial_\nu U^\mu - U^\nu \partial_\nu V^\mu$$

which manifests the antisymmetry under $U \leftrightarrow V$ exchange[2].

The action on one forms is relatively easy to derive; consider the scalar $\omega_\mu U^\mu$ and act upon it using the Lie derivative:

$$\begin{aligned}\mathcal{L}_V(\omega_\mu U^\mu) &= V(\omega_\mu U^\mu)\\ &= V^\nu\partial_\nu(\omega_\mu U^\mu)\\ &= V^\nu\left[\left(\partial_\nu\omega_\mu\right)U^\mu + \omega_\mu\left(\partial_\nu U^\mu\right)\right]\end{aligned}$$

Now act upon the scalar again, employing the Leibniz rule:

$$\begin{aligned}\mathcal{L}_V(\omega_\mu U^\mu) &= \mathcal{L}_V(\omega_\mu)U^\mu + \omega_\mu\mathcal{L}_V(U^\mu)\\ &= \mathcal{L}_V(\omega_\mu)U^\mu + \omega_\mu V^\nu\partial_\nu U^\mu - \omega_\mu U^\nu\partial_\nu V^\mu\end{aligned}$$

[2]See Carroll, pg. 432-433 for details

Setting these equal gives

$$\begin{aligned}
\mathcal{L}_V(\omega_\mu)U^\mu + \omega_\mu V^\nu \partial_\nu U^\mu - \omega_\mu U^\nu \partial_\nu V^\mu &= V^\nu \left[\left(\partial_\nu \omega_\mu\right) U^\mu + \omega_\mu \left(\partial_\nu U^\mu\right)\right] \\
\mathcal{L}_V(\omega_\mu)U^\mu - \omega_\mu U^\nu \partial_\nu V^\mu &= V^\nu \left(\partial_\nu \omega_\mu\right) U^\mu \\
\mathcal{L}_V(\omega_\mu)U^\mu &= V^\nu \left(\partial_\nu \omega_\mu\right) U^\mu + \omega_\mu U^\nu \partial_\nu V^\mu \\
\mathcal{L}_V(\omega_\mu)U^\mu &= \left(V^\nu \left(\partial_\nu \omega_\mu\right) + \omega_\nu \partial_\mu V^\nu\right) U^\mu
\end{aligned}$$

Since this must hold for arbitrary U^μ then

$$\mathcal{L}_V(\omega_\mu) = V^\nu \left(\partial_\nu \omega_\mu\right) + \omega_\nu \partial_\mu V^\nu$$

It is worth noting that for coordinate (or holonomic) bases $U_\mu = \partial_\mu$, the Lie derivative vanishes identically (as a result of the equality of mixed partial derivatives):

$$\mathcal{L}_{\partial_\mu}(\partial_\nu) = [\partial_\mu, \partial_\nu] = 0$$

Using these relations, it is possible to find the Lie derivative of a $\binom{m}{n}$ tensor:

$$\begin{aligned}
\mathcal{L}_V T^{\mu_1 \cdots \mu_m}{}_{\nu_1 \dots \nu_n} &= V^\sigma \partial_\sigma T^{\mu_1 \cdots \mu_m}{}_{\nu_1 \dots \nu_n} \\
&- \partial_\sigma V^{\mu_1} T^{\sigma \mu_2 \cdots \mu_m}{}_{\nu_1 \dots \nu_n} - \cdots - \partial_\sigma V^{\mu_m} T^{\mu_1 \cdots \mu_{m-1}\sigma}{}_{\nu_1 \dots \nu_n} \\
&+ \partial_{\nu_1} V^\sigma T^{\mu_1 \cdots \mu_m}{}_{\sigma \nu_2 \dots \nu_n} + \cdots + \partial_{\nu_n} V^\sigma T^{\mu_1 \cdots \mu_m}{}_{\nu_1 \dots \nu_{n-1}\sigma}
\end{aligned}$$

Lastly, consider the explicit example of the metric tensor. Since it is a $\binom{0}{2}$ tensor, the Lie derivative is given by

$$\mathcal{L}_V g_{\mu\nu} = V^\rho \partial_\rho g_{\mu\nu} + \partial_\mu V^\sigma g_{\sigma\nu} + \partial_\nu V^\sigma g_{\mu\sigma}$$

This may be rewritten into a cleaner, manifestly covariant form by promoting the partial derivatives to covariant derivatives and employing metric compatibility:

$$\begin{aligned}
\mathcal{L}_V g_{\mu\nu} &= V^\rho \partial_\rho g_{\mu\nu} + \partial_\mu V^\sigma g_{\sigma\nu} + \partial_\nu V^\sigma g_{\mu\sigma} \\
&\to V^\rho \nabla_\rho g_{\mu\nu} + \nabla_\mu V^\sigma g_{\sigma\nu} + \nabla_\nu V^\sigma g_{\mu\sigma} \\
&= \nabla_\mu V_\nu + \nabla_\nu V_\mu \\
&= \nabla_{(\mu} V_{\nu)}
\end{aligned}$$

This allows us to explicitly define what a symmetry is. Recall that a diffeomorphism is a map $\phi : M \to M$ under which a tensor T can be pulled back. Then that mapping ϕ is considered a symmetry of T if

$$\phi^* T = T$$

If the diffeomorphism is generated by a vector field V, then this amounts to the Lie derivative along that vector field vanishing:

$$\mathcal{L}_V T = 0$$

Stated in words, this amounts to stating that if T is symmetric under some diffeomorphism, then there exists a coordinate system in which the components of T are independent of one of the coordinates. The converse is also true.

2.3 Integration

Integration on manifolds is a generalization of integration from ordinary calculus on $\mathbb{R}^n$. Rather than a simple sum over infinitesimal values, integration is instead understood as a map from an n-form field ω to the real numbers:

$$\int_\Sigma : \omega \to \mathbb{R}$$

where $\Sigma \subset M$ is an n-dimensional subset of M. This certainly looks strange; where is the infinitesimal dx? Well, consider ω as a 1-form; then $\omega = \omega_\mu \mathrm{d}x^\mu$ and let's suppose further that ω is one-dimensional. Then

$$\int_\Sigma \omega = \int_\Sigma \omega_\mu(x)\mathrm{d}x^\mu = \int_a^b \omega(x)\mathrm{d}x$$

which is a good, old-fashioned integral! Of course, there are complications when ω is a general n-form. There is the question as to why the integrand must be a form. Recall that a form is simply a $\binom{0}{n}$ tensor that is antisymmetric. The antisymmetry allows a notion of orientation to be defined and this is important b/c it allows a "positive" volume to be distinguished from a "negative" volume. A volume element defined by two vectors should vanish when the vectors are parallel, which only occurs when the vectors are antisymmetrically multiplied. Thus a volume element should be an n-form, defined by

$$dx^n = \mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^{n-1}$$

This, however, is not a tensor. It looks like one, but let's see if it transforms properly:

$$\begin{aligned} dx^n = \mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^{n-1} &= \frac{1}{n!}\tilde{\epsilon}_{\mu_1 \cdots \mu_n} \mathrm{d}x^{\mu_1} \wedge \cdots \wedge \mathrm{d}x^{\mu_n} \\ &= \frac{1}{n!}\tilde{\epsilon}_{\mu_1 \cdots \mu_n} \frac{\partial x^{\mu_1}}{\partial x^{\mu_1'}} \cdots \frac{\partial x^{\mu_n}}{\partial x^{\mu_n'}} \mathrm{d}x^{\mu_1'} \wedge \cdots \wedge \mathrm{d}x^{\mu_n'} \\ &= \frac{1}{n!}\left|\det\left(\frac{\partial x^\mu}{\partial x^{\mu'}}\right)\right| \tilde{\epsilon}_{\mu_1' \cdots \mu_n'} \mathrm{d}x^{\mu_1'} \wedge \cdots \wedge \mathrm{d}x^{\mu_n'} \\ &= \left|\det\left(\frac{\partial x^\mu}{\partial x^{\mu'}}\right)\right| \mathrm{d}^n x' \end{aligned}$$

The factor of the Jacobian determinant means that the volume element is a tensor density of weight $w = 1$. This can be easily rectified by multiplying by a factor of $g^{\frac{1}{2}}$ which cancels the factor and leaves a tensor. Thus an invariant volume element is

$$\epsilon = \sqrt{|g|}\mathrm{d}^n x$$

Therefore an integral I over a scalar function $\phi(x)$ is given by

$$I = \int \phi(x)\sqrt{|g|}d^n x$$

2.4 The Stokes theorem

3 The Metric

The metric $g_{\mu\nu}$ is a symmetric $\binom{0}{2}$ tensor. Its role in general relativity cannot be understated as it encodes the notion of length and all information about the curvature of a manifold. Since GR can be summed up as "curvature = gravity", it's worth taking a moment to study the metric explicitly.

3.1 Flat spacetime metric

Let's briefly review the role the metric plays in flat spacetime. Here, $g_{\mu\nu} = \eta_{\mu\nu}$ where $\eta_{\mu\nu}$ is a 4×4 symmetric matrix of the form[3]

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

As a $\binom{0}{2}$ tensor, the metric may be used to raise and lower indices; this means it can take vectors to dual vectors and vice-versa, as well as much more. In addition, the metric when acting on two vectors provides the inner product

$$\eta(V, W) = \eta_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu \left(\partial_\alpha \partial_\beta\right) V^\alpha W^\beta = \eta_{\mu\nu} V^\alpha W^\beta \delta^\mu_\alpha \delta^\nu_\beta = \eta_{\mu\nu} V^\mu W^\nu = V^\mu W_\mu = V \cdot W$$

Two vectors with a vanishing inner product are, of course, orthogonal. The inner product is a scalar and thus unaffected by Lorentz transformations. Setting $V^\mu = W^\mu$ and applying the inner product supplies the norm of the vector V. This number need not be positive-definite, thanks to the choice of metric in Minkowski spacetime:

$$\eta_{\mu\nu} V^\mu V^\nu = \begin{cases} < 0, & V^\mu \text{ is timelike} \\ = 0, & V^\mu \text{ is null} \\ > 0, & V^\mu \text{ is spacelike} \end{cases}$$

This in turn allows to define the notion of a length ds^2; setting $V^\mu = W^\mu = \mathrm{d}x^\mu$ yields

$$ds^2 = \eta_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu$$

Lastly, since the metric has non-zero determinant, it is invertible. The inverse metric $\eta^{\mu\nu}$ is a $\binom{2}{0}$ tensor; since it's an inverse, it satisfies the relation

$$\eta^{\alpha\beta} \eta_{\beta\sigma} = \delta^\alpha_\sigma$$

3.2 Metrics for general manifolds

Of course, the metric described above is a special case of the metric for flat spacetimes. In general, the metric

[3] I'm using the East Coast metric here for consistency w/ Carroll, despite how much it pains me

(1) supplies a notion of "past" and "future"

(2) allows the computation of path length and proper time

(3) determines the "shortest distance" between two points

(4) replaces the Newtonian gravitational field

(5) provides a notion of locally inertial frames and therefore a sense of "no rotation"

(6) determines causality, by defining a speed of light faster than which no signal can travel

(7) replaces the traditional Euclidean dot product of Newtonian mechanics

A few properties from before still hold. First, the metric $g_{\mu\nu}$ is still a (usually) non-degenerate symmetric $\binom{0}{2}$ tensor. The inverse metric $g^{\mu\nu}$ still satisfies

$$g^{\alpha\beta}g_{\beta\sigma} = \delta^{\alpha}_{\sigma}$$

The notion of a differential path length ds^2 remains the same

$$ds^2 = g_{\mu\nu}\mathrm{d}x^{\mu}\mathrm{d}x^{\nu}$$

In Minkowski spacetime, it is possible to choose a set of coordinates such that all the $g_{\mu\nu}$ are constant. In curved spacetime, it is not possible to find a set of coordinates such that $g_{\mu\nu}$ =constant everywhere; otherwise the space would be flat! Before proceeding, let's define the canonical form of the metric:

$$g_{\mu\nu} = \mathrm{diag}\,(-1, \ldots, -1, 1, \ldots, 1, 0, \ldots, 0)$$

In canonical form, the signature of the metric is plainly visible, in the number of $+1$ and -1. A Riemannian metric is one with all positive eigenvalues; a pseudo-Riemannian metric is one with exactly one negative eigenvalue.

3.3 Locally inertial coordinates

As stated before, it is not possible to choose a global coordinate system the metric such that all its values are constant everywhere; however, at a point p on the manifold, it is possible to choose coordinates (which we'll label with a hatted index, $x^{\hat{\mu}}$) such that the metric takes its canonical form *and* the first derivatives $\partial_{\hat{\sigma}}g_{\hat{\mu}\hat{\nu}} = 0$. In short,

$$g_{\hat{\mu}\hat{\nu}}(p) = \eta_{\hat{\mu}\hat{\nu}} \qquad\qquad \partial_{\hat{\sigma}}g_{\hat{\mu}\hat{\nu}}(p) = 0$$

The $x^{\hat{\mu}}$ are known as *locally ineritVal coordinates*; the associated basis vectors $\partial_{\hat{\mu}}$ are known as a *locally Lorentz frame.* This provides a rigorous definition to the notion of local patches of curved spacetime "looking flat": to first order, the metric is simply that of flat spacetime! It is worth stating that there is no difficulty in constructing a set of basis vectors so that $g_{\mu\nu}$ takes its canonical form; however, it is not possible to specify a set of coordinates x^{μ} so that this is true globally. Furthermore, it is not possible in general to choose coordinates such that the second derivative of the metric vanishes simultaneously with the first derivative and the metric assumes its canonical form; there simply aren't enough degrees of freedom to make that happen[4]

[4]see Carroll, pg. 74-75 for details

4 The Connection

4.1 General covariant derivatives

Suppose T is some arbitrary $\binom{2}{0}$ tensor, for example. In flat space the partial derivative of its components $\partial_\alpha T^{\mu\nu}$ would yield a $\binom{2}{1}$ tensor, whose components transform as follows

$$\begin{aligned}\partial_\alpha T^{\mu\nu} &= \frac{\partial x^{\alpha'}}{\partial x^\alpha}\frac{\partial}{\partial x^{\alpha'}}\left(\frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial x^\nu}{\partial x^{\nu'}}T^{\mu'\nu'}\right)\\ &= \frac{\partial x^{\alpha'}}{\partial x^\alpha}\frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial x^\nu}{\partial x^{\nu'}}\partial_{\alpha'}T^{\mu'\nu'} + \frac{\partial x^{\alpha'}}{\partial x^\alpha}\frac{\partial^2 x^\mu}{\partial x^{\alpha'}\partial x^{\mu'}}\frac{\partial x^\nu}{\partial x^{\nu'}}T^{\mu'\nu'} + \frac{\partial x^{\alpha'}}{\partial x^\alpha}\frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial^2 x^\nu}{\partial x^{\alpha'}\partial x^{\nu'}}T^{\mu'\nu'}\end{aligned}$$

In flat space, the second derivative factors would all vanish identically which gives the usual tensor transformation rule. But in curved space, that needn't be true at all. Thus it is necessary to generalize the partial derivative to what is known as the *covariant derivative*, ∇. Since ∇ is a derivative, it should be linear and obey the Leibniz rule

$$\nabla(T+S) = \nabla T + \nabla S \qquad\qquad \nabla(T\otimes S) = (\nabla T)\otimes S + T\otimes(\nabla S)$$

Furthermore, we will demand that the derivative commute with the Kronecker delta; this is equivalent to saying that the identity map is the same everywhere. Furthermore, the covariant derivative should reduce to the partial derivative on scalars. It can be shown that ∇ may be written as a partial derivative plus some correction that cancels the non-tensorial part of the partial derivative; this factor is known as the *connection*. Thus for each direction μ, the components of the covariant derivative ∇_μ will be given by ∂_μ plus a correction by n $n\times n$ matrices (indexed by μ, components given by ρ,σ), $(\Gamma^\mu)^\rho{}_\sigma$. In equation form this becomes

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\sigma}V^\sigma$$

So what are the components of $\Gamma^\nu_{\mu\sigma}$? We are demanding that $\nabla_\mu V^\nu$ be a tensor. In other words, it satisfies the relation

$$\nabla_\mu V^\nu = \frac{\partial x^{\mu'}}{\partial x^\mu}\frac{\partial x^\nu}{\partial x^{\nu'}}\nabla_{\mu'}V^{\nu'}$$

Basically, we're shunting the non-tensorial part of the partial derivative into $\Gamma^\nu_{\mu\sigma}$ so that the whole thing together transforms as a tensor. Note that this means that $\Gamma^\nu_{\mu\sigma}$ is **not** a tensor; therefore its indices cannot be raised and lowered meaningfully!

So what is the physical interpretation of the connection coefficients $\Gamma^\nu_{\mu\sigma}$? Clearly, the ν represents the component of the vector V^ν that is changing. μ corresponds to the partial derivative index ∂_μ, which indicates the direction in which the component is changing. Lastly, σ corresponds to the direction along which the vector is being moved. In summary, $\Gamma^\nu_{\mu\sigma}$ represents the change in the μ direction of the ν component of the vector V due to motion in the σ direction.

What is the action of ∇_μ on a dual vector? Very similarly to the case with a vector, the covariant derivative on a dual vector is given by

$$\nabla_\mu\omega_\nu = \partial_\mu\omega_\nu + \tilde{\Gamma}^\sigma_{\mu\nu}\omega_\sigma$$

where the $\sim$ over the Γ serves to distinguish the symbol from the corresponding vector symbol. What is the relation between $\Gamma^{\nu}_{\mu\sigma}$ and $\tilde{\Gamma}^{\sigma}_{\mu\nu}$? This can be found by considering the covariant derivative on a scalar formed by $V^{\sigma}\omega_{\sigma}$:

$$\begin{aligned}
\nabla_{\mu}\left(V^{\nu}\omega_{\nu}\right) &= (\nabla_{\mu}V^{\nu})\omega_{\nu} + V^{\nu}(\nabla_{\mu}\omega_{\nu}) \\
\partial_{\mu}(V^{\nu}\omega_{\nu}) &= \left(\partial_{\mu}V^{\nu} + \Gamma^{\nu}_{\mu\sigma}V^{\sigma}\right)\omega_{\nu} + V^{\nu}\left(\partial_{\mu}\omega_{\mu} + \tilde{\Gamma}^{\sigma}_{\mu\nu}\omega_{\sigma}\right) \\
(\partial_{\mu}V^{\nu})\omega_{\nu} + V^{\nu}(\partial_{\mu}\omega_{\nu}) &= (\partial_{\mu}V^{\nu})\omega_{\nu} + \Gamma^{\nu}_{\mu\sigma}V^{\sigma}\omega_{\nu} + V^{\nu}(\partial_{\mu}\omega_{\mu}) + V^{\nu}\tilde{\Gamma}^{\sigma}_{\mu\nu}\omega_{\sigma} \\
0 &= \Gamma^{\nu}_{\mu\sigma}V^{\sigma}\omega_{\nu} + \tilde{\Gamma}^{\nu}_{\mu\sigma}V^{\sigma}\omega_{\nu} \\
\Gamma^{\nu}_{\mu\sigma}V^{\sigma}\omega_{\nu} &= -\tilde{\Gamma}^{\nu}_{\mu\sigma}V^{\sigma}\omega_{\nu}
\end{aligned}$$

which gives

$$\Gamma^{\nu}_{\mu\sigma} = -\tilde{\Gamma}^{\nu}_{\mu\sigma}$$

Thus the covariant derivative for dual vectors is identical to that for vectors, except for the minus sign on the connection coefficients

$$\nabla_{\mu}\omega_{\nu} = \partial_{\mu}\omega_{\nu} - \Gamma^{\sigma}_{\mu\nu}\omega_{\sigma}$$

This generalizes for $\binom{m}{n}$ tensors very cleanly:

$$\begin{aligned}
\nabla_{\mu}T^{\alpha_1\ldots\alpha_m}{}_{\beta_1\ldots\beta_n} = {} & \partial_{\mu}T^{\alpha_1\ldots\alpha_m}{}_{\beta_1\ldots\beta_n} \\
& + \Gamma^{\alpha_1}_{\mu\nu}T^{\nu\alpha_2\ldots\alpha_m}{}_{\beta_1\ldots\beta_n} + \cdots + \Gamma^{\alpha_m}_{\mu\nu}T^{\alpha_1\ldots\nu\alpha_m}{}_{\beta_1\ldots\beta_n} \\
& - \Gamma^{\nu}_{\mu\beta_1}T^{\alpha_1\ldots\alpha_m}{}_{\nu\beta_2\ldots\beta_n} - \cdots - \Gamma^{\nu}_{\mu\beta_n}T^{\alpha_1\ldots\alpha_m}{}_{\beta_1\ldots\nu\beta_n}
\end{aligned}$$

Lastly, the difference between two covariant derivatives ∇_{μ} and $\hat{\nabla}_{\mu}$ generates a tensor $S^{\sigma}{}_{\mu\nu}$

$$\nabla_{\mu}V^{\sigma} - \hat{\nabla}_{\mu}V^{\sigma} = \partial_{\mu}V^{\sigma} + \Gamma^{\sigma}_{\mu\nu}V^{\nu} - \partial_{\mu}V^{\sigma} - \hat{\Gamma}^{\sigma}_{\mu\nu}V^{\nu} = \Gamma^{\sigma}_{\mu\nu}V^{\nu} - \hat{\Gamma}^{\sigma}_{\mu\nu}V^{\nu} = (\Gamma^{\sigma}_{\mu\nu} - \hat{\Gamma}^{\sigma}_{\mu\nu})V^{\nu} = S^{\sigma}{}_{\mu\nu}V^{\nu}$$

Since the term on the far left transforms as a tensor, then the factor on the far right must also transform as a tensor and since V^{ν} is arbitrary, $S^{\sigma}{}_{\mu\nu}$ must be a tensor itself. Thus a connection $\Gamma^{\sigma}_{\mu\nu}$ may be written in terms of some arbitrary connection $\hat{\Gamma}^{\sigma}_{\mu\nu}$ plus a tensorial correction. One such correction is simply given by setting $\hat{\Gamma}^{\sigma}_{\mu\nu} = \Gamma^{\sigma}_{\nu\mu}$; the corresponding correction tensor $S^{\sigma}{}_{\mu\nu} = T^{\sigma}{}_{\mu\nu}$ is known as the *torsion tensor*

$$T^{\sigma}{}_{\mu\nu} = \Gamma^{\sigma}{}_{\mu\nu} - \Gamma^{\sigma}{}_{\nu\mu} = 2\Gamma^{\sigma}_{[\mu\nu]}$$

The torsion effectively measures how much a vector 'twists' about itself when moved along another vector. For basic purposes, we won't use it much, but it may come in handy later. As a final note, the *contorsion tensor* is defined as the difference between a connection with torsion and one without.

4.2 The Levi-Civita connection

Thus far, the general conditions on the covariant derivative (and, by extension, the connection) have been

- $\nabla(T + S) = \nabla T + \nabla S$ linearity
- $\nabla(T \otimes S) = (\nabla T) \otimes S + T \otimes (\nabla S)$ satisfaction of the Leibniz rule
- $\nabla_\mu \delta^\alpha_\beta = 0$ constancy of the the identity
- $\nabla_\mu \phi = \partial_\mu \phi$ reduction to the partial derivative for scalars

For a (semi-)Riemannian manifold M, there are any number of connections that satisfy these conditions. There are two more conditions that when imposed simultaneously imply a unique connection with several nice properties. These are that the connection be

- $\Gamma^\sigma_{\mu\nu} = \Gamma^\sigma_{(\mu\nu)}$ torsion-free
- $\nabla_\mu g_{\alpha\beta} = 0$ metric compatible

A connection satisfying these properties is known as the *Levi-Civita* connection; the associated connection coefficients are known as the *Christoffel symbols*. The first of these conditions essentially allows us to create covariantly transforming objects out of objects containing partial derivatives (such as the Lie and exterior derivatives) by simply promoting the partials to covariant derivatives:

$$(\mathrm{d}\omega)_{\mu\nu} = 2\partial_{[\mu}\omega_{\nu]} = 2\nabla_{[\mu}\omega_{\nu]}$$

and

$$[X, Y]^\mu = X^\sigma \partial_\sigma Y^\mu - Y^\sigma \partial_\sigma X^\mu = X^\sigma \nabla_\sigma Y^\mu - Y^\sigma \nabla_\sigma X^\mu$$

The second of these conditions is weightier. It's the first explicit contact between an object with physical consequence (the metric) and geometry (the covariant derivative). In many ways, it represents a concretization of the Equivalence Principle. Metric compatibility means that index raising and lowering commutes with the derivative and it allows us to find explicit expressions for the Christoffel symbols. Consider the three covariant derivatives

$$\begin{aligned}
0 = \nabla_\rho g_{\mu\nu} &= \partial_\rho g_{\mu\nu} - \Gamma^\sigma_{\rho\mu} g_{\sigma\nu} - \Gamma^\sigma_{\rho\nu} g_{\mu\sigma} \\
0 = \nabla_\mu g_{\nu\rho} &= \partial_\mu g_{\nu\rho} - \Gamma^\sigma_{\mu\nu} g_{\sigma\rho} - \Gamma^\sigma_{\mu\rho} g_{\nu\sigma} \\
0 = \nabla_\nu g_{\rho\mu} &= \partial_\nu g_{\rho\mu} - \Gamma^\sigma_{\nu\rho} g_{\sigma\mu} - \Gamma^\sigma_{\nu\mu} g_{\rho\sigma}
\end{aligned}$$

Recall that $g_{\mu\nu} = g_{\nu\mu}$ and that we are demanding that $\Gamma^\sigma_{\mu\nu} = \Gamma^\sigma_{\nu\mu}$. Therefore the above may be written as

$$\begin{aligned}
0 = \nabla_\rho g_{\mu\nu} &= \partial_\rho g_{\mu\nu} - \Gamma^\sigma_{\mu\rho} g_{\nu\sigma} - \Gamma^\sigma_{\nu\rho} g_{\mu\sigma} \\
0 = \nabla_\mu g_{\nu\rho} &= \partial_\mu g_{\nu\rho} - \Gamma^\sigma_{\mu\nu} g_{\rho\sigma} - \Gamma^\sigma_{\mu\rho} g_{\nu\sigma} \\
0 = \nabla_\nu g_{\mu\rho} &= \partial_\nu g_{\rho\mu} - \Gamma^\sigma_{\nu\rho} g_{\mu\sigma} - \Gamma^\sigma_{\mu\nu} g_{\rho\sigma}
\end{aligned}$$

Subtracting the third equation from the first gives

$$0 = \partial_\rho g_{\mu\nu} - \Gamma^\sigma_{\mu\rho} g_{\nu\sigma} - \Gamma^\sigma_{\nu\rho} g_{\mu\sigma} - \partial_\nu g_{\rho\mu} + \Gamma^\sigma_{\nu\rho} g_{\mu\sigma} + \Gamma^\sigma_{\mu\nu} g_{\rho\sigma} = \partial_\rho g_{\mu\nu} - \partial_\nu g_{\mu\rho} - \Gamma^\sigma_{\mu\rho} g_{\nu\sigma} + \Gamma^\sigma_{\mu\nu} g_{\rho\sigma}$$

Adding the second equation to the above yields

$$\begin{aligned}
0 &= \partial_\rho g_{\mu\nu} - \partial_\nu g_{\mu\rho} - \Gamma^\sigma_{\mu\rho} g_{\nu\sigma} + \Gamma^\sigma_{\mu\nu} g_{\rho\sigma} + \partial_\mu g_{\nu\rho} - \Gamma^\sigma_{\mu\nu} g_{\rho\sigma} - \Gamma^\sigma_{\mu\rho} g_{\nu\sigma} \\
&= \partial_\rho g_{\mu\nu} - \partial_\nu g_{\mu\rho} + \partial_\mu g_{\nu\rho} - 2\Gamma^\sigma_{\mu\rho} g_{\nu\sigma} \\
2\Gamma^\sigma_{\mu\rho} g_{\nu\sigma} &= \partial_\rho g_{\mu\nu} - \partial_\nu g_{\mu\rho} + \partial_\mu g_{\nu\rho} \\
\Gamma^\sigma_{\mu\rho} g_{\nu\sigma} g^{\nu\tau} &= \frac{1}{2} g^{\nu\tau} \left(\partial_\rho g_{\mu\nu} - \partial_\nu g_{\mu\rho} + \partial_\mu g_{\nu\rho}\right) \\
\Gamma^\sigma_{\mu\rho} \delta^\tau_\sigma &= \frac{1}{2} g^{\nu\tau} \left(\partial_\rho g_{\mu\nu} - \partial_\nu g_{\mu\rho} + \partial_\mu g_{\nu\rho}\right) \\
\Gamma^\tau_{\mu\rho} &= \frac{1}{2} g^{\nu\tau} \left(\partial_\rho g_{\mu\nu} - \partial_\nu g_{\mu\rho} + \partial_\mu g_{\nu\rho}\right)
\end{aligned}$$

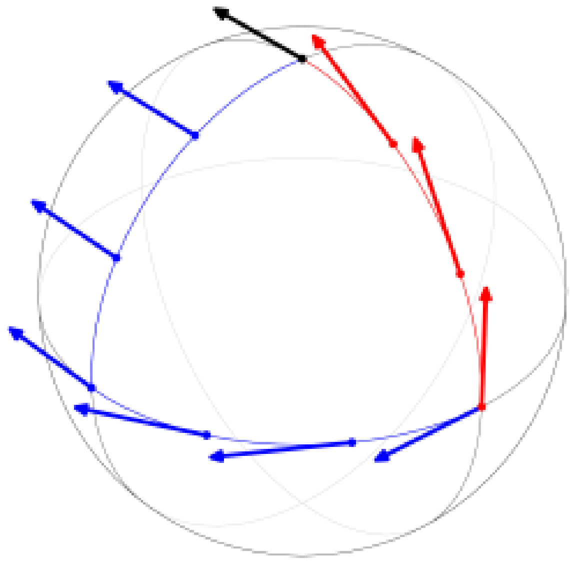
Cleaning up the indices, this yields an expression for the *Christoffel symbols of the second kind*

$$\Gamma^\sigma_{\mu\nu} = \frac{1}{2} g^{\sigma\rho} \left(\partial_\mu g_{\nu\rho} - \partial_\rho g_{\mu\nu} + \partial_\nu g_{\rho\mu}\right)$$

Notice that the Christoffel symbols are written in terms of derivatives of the metric. Recall that there exists a set of coordinates, the locally inertial coordinates, such that the metric and its first derivatives vanish at a point p. Thus it is possible to make the Christoffel symbols vanish at a point p. However, recall that although it is possible to make the metric and its first derivative vanish, it is not possible to make the second derivative vanish by choice of coordinates. Therefore, derivatives of the Christoffel symbols will not vanish, since they involve second derivatives of the metric. This will be important when defining the Riemann curvature tensor.

4.3 Parallel Transport

We're already familiar with the notion of parallel transport, even if we've never called it by that name. When adding/subtracting vectors using the parallelogram rule, we "move" one vector along another to put them "tail-to-tip," keeping the vector constant. This is known as parallel transport. Implicitly, whenever we do this in flat space, we've been using the Levi-Civita connection without realizing it. How does this process generalize to curved space and why does that matter? In flat space, we can parallel transport a vector along any path and it won't make a difference; in curved space, this is no longer true! To see this, consider the typical example involving vectors on a sphere, as shown the figure



Beginning with the black vector at the north pole, the vector is sent along two different paths (the red and blue vectors) to the same point on the equator; despite being kept "fixed," the vector has changed. Both paths represent parallel transport, which illustrates the fact that on an arbitrary curved space, *there is no canonical way to parallel transport a vector*. This is because parallel transport depends on the connection used. This in turn leads to the fact that there is no standard way to compare vectors that don't live in the same tangent space.

So what can we do? First, let's formalize the notion of parallel transport for a tensor on flat space. Recall that in flat space this means keeping the components of the tensor constant when moving it along a path. Let $x^\mu(\lambda)$ be the components of a vector field along which we're parallel transporting the tensor, where λ is a real parameter describing the vector field. Then the parallel transport condition is that the directional derivative along this vector field vanish:

$$\frac{d}{d\lambda}T^{\mu_1\cdots\mu_m}{}_{\nu_1\dots\nu_n} = \frac{dx^\mu}{d\lambda}\frac{\partial}{\partial x^\mu}T^{\mu_1\cdots\mu_m}{}_{\nu_1\dots\nu_n} = 0$$

Recall that to make an expression manifestly covariant, we replace the partial derivatives with covariant derivatives. This gives the *covariant directional derivative*:

$$\frac{d}{d\lambda} = \frac{dx^\mu}{d\lambda}\frac{\partial}{\partial x^\mu} \qquad \to \qquad \frac{D}{d\lambda} = \frac{dx^\mu}{d\lambda}\nabla_\mu$$

Notice that, as stated earlier, this derivative is not canonically defined independent of the connection used in the covariant derivative! The covariant directional derivative is a map defined along a path $x^\mu(\lambda)$ (and assuming a connection) from $\binom{m}{n}$ tensors to $\binom{m}{n}$ tensors. Then, parallel transport of the tensor T is the condition that the covariant directional derivative vanish:

$$\left(\frac{D}{d\lambda}T\right)^{\mu_1\cdots\mu_m}{}_{\nu_1\dots\nu_n} = \frac{dx^\mu}{d\lambda}\nabla_\mu T^{\mu_1\cdots\mu_m}{}_{\nu_1\dots\nu_n} = 0$$

Plugging in the form of the covariant derivative gives, for a vector, the following *equation* *for parallel transport*

$$0 = \left(\frac{D}{d\lambda}V\right)^{\mu} = \frac{dx^{\nu}}{d\lambda}\nabla_{\nu}V^{\mu} = \frac{dx^{\nu}}{d\lambda}\partial_{\nu}V^{\mu} + \Gamma^{\mu}_{\nu\sigma}\frac{dx^{\nu}}{d\lambda}V^{\sigma} = \frac{d}{d\lambda}V^{\mu} + \Gamma^{\mu}_{\nu\sigma}\frac{dx^{\nu}}{d\lambda}V^{\sigma}$$

This is a first-order differential equation. Given some initial value for the tensor at some point along the path, solutions to the differential equation are those that represent parallel transport.

Now let's specialize to the Levi-Civita connection. In this case, the covariant directional derivative vanishes, since the covariant derivative is metric-compatible:

$$\frac{D}{d\lambda}g_{\mu\nu} = \frac{dx^{\sigma}}{d\lambda}\nabla_{\sigma}g_{\mu\nu} = 0$$

This in turn implies that quantities built out the metric tensor (such as lengths or inner products) are preserved:

$$\frac{D}{d\lambda}(g_{\mu\nu}V^{\mu}W^{\nu}) = \frac{D}{d\lambda}(g_{\mu\nu})V^{\mu}W^{\nu} + g_{\mu\nu}\frac{D}{d\lambda}(V^{\mu})W^{\nu} + g_{\mu\nu}V^{\mu}\frac{D}{d\lambda}(W^{\nu}) = 0$$

4.4 Geodesics

A geodesic is a generalization of the notion of a straight line in Euclidean space. There are two possible definitions

> A geodesic is the path of shortest distance (or greatest proper time) between two points

or, when speaking of geodesics on a manifold equipped with a Levi-Civita connection

> A geodesic is the path that parallel transports its own tangent vector.

The second definition immediately yields an equation for a geodesic. A tangent vector to a path parameterized by λ is simply the directional derivative $\frac{dx^{\mu}}{d\lambda}$. Thus the parallel transport equation (and imposing the geodesic condition)

$$0 = \left(\frac{D}{d\lambda}\frac{dx^{\mu}}{d\lambda}\right) = \frac{dx^{\nu}}{d\lambda}\nabla_{\nu}\left(\frac{dx^{\mu}}{d\lambda}\right) = \frac{dx^{\nu}}{d\lambda}\partial_{\nu}\left(\frac{dx^{\mu}}{d\lambda}\right) + \Gamma^{\mu}_{\nu\sigma}\frac{dx^{\nu}}{d\lambda}\frac{dx^{\sigma}}{d\lambda} = \frac{d^2x^{\mu}}{d\lambda^2} + \Gamma^{\mu}_{\nu\sigma}\frac{dx^{\nu}}{d\lambda}\frac{dx^{\sigma}}{d\lambda}$$

In many ways, this is a generalization of the Newton force equation. A particle experiencing an external force would satisfy the above equation with the right-hand side equal to the sum of forces

$$\frac{1}{m}\sum F^{\mu} = \frac{d^2x^{\mu}}{d\lambda^2} + \Gamma^{\mu}_{\nu\sigma}\frac{dx^{\nu}}{d\lambda}\frac{dx^{\sigma}}{d\lambda}$$

A natural way to parameterize a vector is using the proper time τ; this may be generalized to what is known as an affine parameter so that $\lambda = a\tau + b$. Curves satisfying the geodesic equation are constrained in their parameterization to an affine parameter. But wait, what about null paths (aka those where the proper time vanishes). Is there no parameterization? There is, but as before it must be of the form $a\lambda + b$, along with an appropriate normalization condition (such as $p^{\mu} = \frac{dx^{\mu}}{d\lambda}$)

5 Curvature

5.1 Riemann Curvature Tensor

There are two (equivalent) approaches to the Riemann curvature tensor. The first is very physical and consists of describing the Riemann curvature tensor as a parameterization of the change δV^ρ in a vector V^σ when parallel transported along two vectors A^μ and B^ν

$$\delta V^\rho = R^\rho{}_{\sigma\mu\nu} V^\sigma A^\mu B^\nu$$

Note that in this view it is plainly obvious that $R^\rho{}_{\sigma\mu\nu}$ should be antisymmetric in its last two indices, since parallel transport along A then B is "backwards" relative to transport along B then A. While physically intuitive, this approach is computationally difficult. The second, more tractable approach, consists of computing the effect of a commutator of covariant derivatives on a vector V^ρ:

$$\begin{aligned}
[\nabla_\mu, \nabla_\nu] V^\rho &= \nabla_\mu \nabla_\nu V^\rho - \nabla_\nu \nabla_\mu V^\rho \\
&= \partial_\mu \left(\nabla_\nu V^\rho\right) + \Gamma^\rho_{\mu\sigma} \nabla_\nu V^\sigma - \Gamma^\lambda_{\mu\nu} \nabla_\lambda V^\rho - \partial_\nu \left(\nabla_\mu V^\rho\right) + \Gamma^\rho_{\nu\sigma} \nabla_\mu V^\sigma - \Gamma^\lambda_{\nu\mu} \nabla_\lambda V^\rho \\
&= \partial_\mu \left(\partial_\nu V^\rho + \Gamma^\rho_{\nu\sigma} V^\sigma\right) + \Gamma^\rho_{\mu\sigma} \left(\partial_\nu V^\sigma + \Gamma^\sigma_{\nu\lambda} V^\lambda\right) - \Gamma^\lambda_{\mu\nu} \left(\partial_\lambda V^\rho + \Gamma^\rho_{\lambda\sigma} V^\sigma\right) \\
&\qquad - \partial_\nu \left(\partial_\mu V^\rho + \Gamma^\rho_{\mu\sigma} V^\sigma\right) - \Gamma^\rho_{\nu\sigma} \left(\partial_\mu V^\sigma + \Gamma^\sigma_{\mu\lambda} V^\lambda\right) + \Gamma^\lambda_{\nu\mu} \left(\partial_\lambda V^\rho + \Gamma^\rho_{\lambda\sigma} V^\sigma\right) \\
\\
&= \partial_\mu \left(\partial_\nu V^\rho + \Gamma^\rho_{\nu\sigma} V^\sigma\right) + \Gamma^\rho_{\mu\sigma} \left(\partial_\nu V^\sigma + \Gamma^\sigma_{\nu\lambda} V^\lambda\right) \\
&\qquad - \partial_\nu \left(\partial_\mu V^\rho + \Gamma^\rho_{\mu\sigma} V^\sigma\right) - \Gamma^\rho_{\nu\sigma} \left(\partial_\mu V^\sigma + \Gamma^\sigma_{\mu\lambda} V^\lambda\right) - \left(\Gamma^\lambda_{\mu\nu} - \Gamma^\lambda_{\nu\mu}\right) \left(\partial_\lambda V^\rho + \Gamma^\rho_{\lambda\sigma} V^\sigma\right) \\
\\
&= (\partial_\mu \Gamma^\rho_{\nu\sigma}) V^\sigma + \Gamma^\rho_{\nu\sigma} \partial_\mu V^\sigma + \Gamma^\rho_{\mu\sigma} \left(\partial_\nu V^\sigma + \Gamma^\sigma_{\nu\lambda} V^\lambda\right) \\
&\qquad - (\partial_\nu \Gamma^\rho_{\mu\sigma}) V^\sigma - \Gamma^\rho_{\mu\sigma} \partial_\nu V^\sigma - \Gamma^\rho_{\nu\sigma} \left(\partial_\mu V^\sigma + \Gamma^\sigma_{\mu\lambda} V^\lambda\right) - 2\Gamma^\lambda_{[\mu\nu]} \left(\partial_\lambda V^\rho + \Gamma^\rho_{\lambda\sigma} V^\sigma\right) \\
\\
&= (\partial_\mu \Gamma^\rho_{\nu\sigma}) V^\sigma - (\partial_\nu \Gamma^\rho_{\mu\sigma}) V^\sigma + \Gamma^\rho_{\mu\sigma} \Gamma^\sigma_{\nu\lambda} V^\lambda - \Gamma^\rho_{\nu\sigma} \Gamma^\sigma_{\mu\lambda} V^\lambda - T^\lambda{}_{\mu\nu} \left(\partial_\lambda V^\rho + \Gamma^\rho_{\lambda\sigma} V^\sigma\right) \\
&= \left(\partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\nu \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda} \Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda} \Gamma^\lambda_{\mu\sigma}\right) V^\sigma - T^\lambda{}_{\mu\nu} \left(\partial_\lambda V^\rho + \Gamma^\rho_{\lambda\sigma} V^\sigma\right) \\
&= R^\rho{}_{\sigma\mu\nu} V^\sigma
\end{aligned}$$

where, in the last step, we've identified the *Riemann curvature tensor* as

$$R^\rho{}_{\sigma\mu\nu} = \partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\nu \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda} \Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda} \Gamma^\lambda_{\mu\sigma}$$

while neglecting the term involving the torsion tensor $T^\lambda{}_{\mu\nu}$ since we're assuming the connection is torsion-free. Notice that here the antisymmetry in the last two indices is manifest. Furthermore, this expression was constructed with no reference to the metric. Certain formulations of GR employ this fact to express the action with the connection as the degree of freedom rather than the connection. Let's express the tensor in coordinate-free language. The torsion tensor is a rank $\binom{1}{2}$ tensor and thus may be thought of as a map from two vector fields X and Y to a third vector field

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]$$

where the notation employed $\nabla_X = X^\mu \nabla_\mu$ means the covariant derivative in the direction of the vector field X. Similarly, the Riemann curvature tensor may be thought of as a map from three vector fields to a fourth vector field; connecting to the earlier viewpoint on the tensor, the first two vectors may be thought of as the vectors along which the third is parallel transported, resulting in the fourth vector field describing the change:

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

Flipping back to coordinates gives

$$R^\rho{}_{\sigma\mu\nu} X^\mu Y^\nu Z^\sigma = X^\mu \nabla_\mu \left(Y^\nu \nabla_\nu Z^\sigma\right) - Y^\nu \nabla_\nu \left(X^\mu \nabla_\mu Z^\sigma\right) - \left(X^\lambda \partial_\lambda Y^\eta - Y^\lambda \partial_\lambda X^\eta\right) \nabla_\eta Z^\sigma$$

If X^μ and Y^ν are coordinate vector fields (which satisfy $[X,Y]=0$, or that their Lie derivative vanishes), then we recover the expression above.

Now let's turn to the Riemann curvature tensor for the Levi-Civita connection. The Riemann curvature tensor is constructed out of derivatives of the Christoffel symbols, which themselves are constructed out of derivatives of the metric. Therefore, the Riemann tensor depends on second derivatives of the metric. However, it was discussed earlier that locally inertial coordinates at a point p could at most make the metric look take on its canonical form and make its first derivatives vanish. The second derivatives were not at all similarly constrained.

So, let's suppose that at a point p, not only were it possible to make the metric look canonical, but also to make its first *and* its second derivatives vanish. Thus the Christoffel symbols would vanish; but then so too would their derivatives and thus the Riemann tensor is $R^\rho{}_{\sigma\mu\nu} = 0$ at p. Since this is a tensor equation, however, it must be true everywhere. Thus if a coordinate system exists such that the Riemann tensor vanishes at that point, the entire space must be flat. This turns out to go the other way, namely that a vanishing tensor implies that a coordinate system exist such that the metric components are constant.[5]

5.2 Properties of the Riemann Curvature Tensor

At first blush, the Riemann tensor has n^4 degrees-of-freedom (n for each index). However, we already know that the last two are not independent, but related by antisymmetry, which reduces the degrees of freedom by $n^2 \to \frac{n(n-1)}{2}$ giving $\frac{n^3(n-1)}{2}$ possible degrees of freedom. We can hack these down further. It is possible to prove[6] that the Riemann tensor is antisymmetric in its first two indices as well

$$R^{\rho\sigma\mu\nu} = -R^{\sigma\rho\mu\nu}$$

[5] See Carroll, pg. 125

[6] See Carroll, pg. 126

This reduces the degrees of freedom to $\left(\frac{n(n-1)}{2}\right)^2$. It is also symmetric under exchange between its first and second pair of indices

$$R_{\mu\nu\rho\sigma} = R_{\rho\sigma\mu\nu}$$

Together, these constraints yield

$$\frac{1}{2}\left(\frac{n(n-1)}{2}\right)\left(\frac{n(n-1)}{2}+1\right) = \frac{1}{8}(n^4 - 2n^3 + 3n^2 - 2n)$$

degrees of freedom. The Riemann tensor, also satisfies the algebraic (first) Bianchi identity

$$R_{\rho\sigma\mu\nu} + R_{\rho\mu\nu\sigma} + R_{\rho\nu\sigma\mu} = R_{\rho[\sigma\mu\nu]} = 0$$

which implies that the totally antisymmetric part of the Riemann tensor vanishes

$$R_{[\rho\sigma\mu\nu]} = 0$$

This can be used to further cut the number of degrees of freedom down by subtracting off the antisymmetric parts that vanish, yielding:

$$\frac{1}{8}(n^4 - 2n^3 + 3n^2 - 2n) - \frac{1}{4!}n(n-1)(n-2)(n-3)(n-4) = \frac{1}{12}n^2(n^2-1)$$

total degrees of freedom. For $n = 4$, this gives 20 degrees of freedom. These have all been algebraic properties of the Riemann tensor. It also obeys differential constraints, linking its components at one point to those at other points. This is known as the (second) Bianchi identity, given by

$$\nabla_\lambda R_{\rho\sigma\mu\nu} + \nabla_\rho R_{\sigma\lambda\mu\nu} + \nabla_\sigma R_{\lambda\rho\mu\nu} = \nabla_{[\lambda} R_{\rho\sigma]\mu\nu} = 0$$

Recall that it is possible to decompose a rank-2 tensor into a symmetric and antisymmetric portion. The symmetric part can be further decomposed into a piece proportional to its trace and a traceless piece. Thus in total, a rank-2 tensor can be written as

$$A_{\mu\nu} = \frac{1}{n}\mathrm{Tr}[A]g_{\mu\nu} + \hat{A}_{\mu\nu} + A_{[\mu\nu]}$$

The equivalent decomposition for something like a rank $\binom{1}{3}$ tensor is of course not nearly as simple. However, such a decomposition may be made of the *Ricci tensor*

$$R_{\mu\nu} = R^\lambda{}_{\mu\lambda\nu}$$

The Ricci tensor is related to the average curvature in the neighborhood of a point. Somewhat like the Maxwell stress tensor which has EM field pressures on the diagonal and stresses on the off-diagonal terms, the Ricci tensor describes the evolution of geodesics near a point. Diagonal terms correspond to the divergence or convergence of field lines, while off-diagonal terms correspond to shearing of field lines. Notice that these are all volume-altering transformations. Diagonal terms lead to a 'squeezing' effect (thus completing the analogy to pressures) while off-diagonal terms lead to shearing effects (corresponding to stresses).

This, of course, is not the only contraction that could be made out of the Riemann tensor. However, for the Levi-Civita connection, all others vanish or are proportional to the Ricci tensor. The Ricci tensor is symmetric and its trace is known as the Ricci scalar $R = g^{\mu\nu}R_{\mu\nu}$. These may be used to subtract all traces from the Riemann tensor, leaving the *Weyl tensor*

$$C_{\rho\sigma\mu\nu} = R_{\rho\sigma\mu\nu} - \frac{2}{n-2}(g_{\rho[\mu}R_{\nu]\sigma} - g_{\sigma[\mu}R_{\nu]\rho}) + \frac{2}{(n-1)(n-2)}g_{\rho[\mu}g_{\nu]\sigma}R$$

The Weyl tensor is only defined in three or more spacetime dimensions, and vanishes identically for $n = 3$. It may be thought of as the traceless portion of the Riemann tensor and is invariant under conformal transformations, meaning that C computed for g and computed again for $\omega(x)g$ where ω is some scalar function of spacetime (a conformal transformation), the same answer is obtained. The Weyl tensor can be thought of the Riemann tensor with 'all the Ricci taken out', Since the Ricci tensor and scalar characterize the compression/shearing of volumes due to the presence of matter, the Weyl tensor corresponds to gravitational effects that preserve volumes, also known as tidal forces. If $T^{\mu\nu} = 0$, then $R^{\mu\nu} = 0 = R$ identically, yet $C_{\rho\sigma\mu\nu}$ does not; in that sense, the Weyl tensor is more fundamental than the Ricci tensor or scalar. The Weyl tensor in $d = 4$ contains 10 independent degrees of freedom; these are the 10 d.o.f. of the Riemann tensor not captured by the Ricci tensor (which itself captures 10 d.o.f.) In fact, the Weyl tensor may be used to describe spacetime dynamics in the absence of matter (such as during the propagation of gravitational waves). A vanishing Weyl tensor corresponds to a spacetime that is conformally flat, meaning one whose metric may be conformally transformed into that of a flat spacetime.

Now let's return to the Ricci tensor. Take the second Bianchi identity and trace over it twice, once to form the Ricci tensor and once to form the Ricci scalar:

$$\begin{aligned}
0 &= g^{\sigma\nu}g^{\lambda\mu}\left(\nabla_\lambda R_{\rho\sigma\mu\nu} + \nabla_\rho R_{\sigma\lambda\mu\nu} + \nabla_\sigma R_{\lambda\rho\mu\nu}\right) \\
&= \nabla_\lambda R_{\rho\sigma}{}^{\lambda\sigma} + \nabla_\rho R^{\nu\mu}{}_{\mu\nu} + \nabla_\sigma R_{\lambda\rho}{}^{\lambda\sigma} \\
&= \nabla^\lambda R^\sigma{}_{\lambda\sigma\rho} - \nabla_\rho R^{\mu\nu}{}_{\mu\nu} + \nabla^\sigma R^\lambda{}_{\sigma\lambda\rho} \\
&= \nabla^\lambda R_{\lambda\rho} - \nabla_\rho R^\nu{}_\nu + \nabla^\sigma R_{\sigma\rho} \\
&= 2\nabla^\lambda R_{\lambda\rho} - \nabla_\rho R \\
\nabla^\lambda R_{\lambda\rho} &= \frac{1}{2}\nabla_\rho R
\end{aligned}$$

This may be written slightly more cleanly

$$\begin{aligned}
\nabla^\mu R_{\mu\nu} &= \frac{1}{2}\nabla^\mu R g_{\mu\nu} \\
\nabla^\mu\left(R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}\right) &= 0 \\
\nabla^\mu G_{\mu\nu} &= 0
\end{aligned}$$

where, in the last step, the *Einstein tensor* was introduced. This version of the Bianchi identity implies that the Einstein tensor is traceless, a fact which will have physical implications down the road.

5.3 Symmetries and Killing Vectors

In order to expedite evaluating the Einstein field equations, as with all physical theories, we employ the use of symmetries. What exactly is a symmetry? We will focus on a particular class of symmetries, namely those symmetries of the geometry, encoded by the metric. Such symmetries are known as *isometries.* A manifold M has a symmetry if under some mapping $\phi : M \to M$, the metric remains the same. In coordinates this becomes

$$g'_{\alpha\beta}(x') = g_{\mu\nu}(x)\frac{\partial x^\mu}{\partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta} = g_{\alpha\beta}(x')$$

Notice the placement of the primes. Evaluating g' at x' should yield g if done under an isometry. In the same vein as a Lie algebra, consider two points related by an infinitesimal transformation

$$x^\mu \qquad \text{and} \qquad x'^\mu = x^\mu + \epsilon K^\mu$$

Inserting this into the coordinate expression for an isometry and keeping only terms linear in ϵ gives

$$\begin{aligned}
g_{\alpha\beta}(x') &= g_{\mu\nu}(x)\frac{\partial x^\mu}{\partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta} \\
g_{\alpha\beta}(x+\epsilon K) &= g_{\mu\nu}(x)\left(\delta^\mu_\alpha - \epsilon\partial_\alpha K^\mu\right)\left(\delta^\nu_\beta - \epsilon\partial_\beta K^\nu\right) \\
g_{\alpha\beta}(x) + \epsilon K^\mu\partial_\mu g_{\alpha\beta}(x) &= g_{\mu\nu}(x)\left(\delta^\mu_\alpha\delta^\nu_\beta - \epsilon\left(\delta^\mu_\alpha\partial_\beta K^\nu + \delta^\nu_\beta\partial_\alpha K^\mu\right) + O(\epsilon^2)\right) \\
g_{\alpha\beta}(x) + \epsilon K^\mu\partial_\mu g_{\alpha\beta}(x) &= g_{\alpha\beta}(x) - \epsilon\left(g_{\alpha\nu}(x)\partial_\beta K^\nu + g_{\mu\beta}(x)\partial_\alpha K^\mu\right)
\end{aligned}$$

Equating the ϵ terms yields

$$0 = K^\mu\partial_\mu g_{\alpha\beta} + g_{\alpha\nu}\partial_\beta K^\nu + g_{\mu\beta}\partial_\alpha K^\mu$$

This may be rewritten into a cleaner, manifestly covariant form. First note

$$\partial_\rho(g_{\mu\sigma}K^\mu) = \partial_\rho K_\sigma = K^\mu(\partial_\rho g_{\mu\sigma}) + g_{\mu\sigma}\partial_\rho K^\mu \qquad \Rightarrow \qquad g_{\mu\sigma}\partial_\rho K^\mu = \partial_\rho K_\sigma - K^\mu(\partial_\rho g_{\mu\sigma})$$

Substitute this into the above equation to give

$$\begin{aligned}
0 &= K^\mu\partial_\mu g_{\rho\sigma} + g_{\rho\nu}\partial_\sigma K^\nu + g_{\mu\sigma}\partial_\rho K^\mu \\
&= K^\mu\partial_\mu g_{\rho\sigma} + \partial_\sigma K_\rho - K^\mu(\partial_\sigma g_{\mu\rho}) + \partial_\rho K_\sigma - K^\mu(\partial_\rho g_{\mu\sigma}) \\
&= \partial_\sigma K_\rho + \partial_\rho K_\sigma - (-\partial_\mu g_{\rho\sigma} + \partial_\sigma g_{\mu\rho} + \partial_\rho g_{\mu\sigma})K^\mu \\
&= \partial_\sigma K_\rho + \partial_\rho K_\sigma - g^{\mu\nu}(-\partial_\mu g_{\rho\sigma} + \partial_\sigma g_{\mu\rho} + \partial_\rho g_{\mu\sigma})K_\nu \\
&= \partial_\sigma K_\rho + \partial_\rho K_\sigma - 2\Gamma^\nu_{\rho\sigma}K_\nu \\
&= \partial_\sigma K_\rho - \Gamma^\nu_{\sigma\rho}K_\nu + \partial_\rho K_\sigma - \Gamma^\nu_{\rho\sigma}K_\nu \\
&= \nabla_\sigma K_\rho + \nabla_\rho K_\sigma
\end{aligned}$$

Using the symmetrization notation this becomes

$$\nabla_{(\sigma}K_{\rho)} = 0$$

This is known as the *Killing equation*. Those fields K that solve this equation are known as *Killing vector fields*. This equation of course generalizes to Killing tensor fields

$$\nabla_{(\mu} K_{\nu_1 \dots \nu_l)} = 0$$

It is possible to show that the geometry is explicitly not changing when moving along a Killing field. The second derivative of a Killing field is given by

$$\nabla_\mu \nabla_\sigma K^\rho = R^\rho{}_{\sigma\mu\nu} K^\nu$$

Contracting this gives

$$\nabla_\rho \nabla_\mu K^\rho = R^\rho{}_{\mu\rho\nu} K^\nu = R_{\mu\nu} K^\nu$$

Together with the Bianchi identity, these give

$$K^\mu \nabla_\mu R = 0$$

which is a manifest demonstration that the geometry (as encoded in the curvature) is unchanging along Killing vector fields. Other symmetries may be encoded as well. Consider a geodesic $x^\mu(\tau)$, with corresponding velocity $v^\mu(\tau)$. Now form a scalar $\xi_\mu(x) v^\mu$; this quantity is conserved when parallel transported along a geodesic:

$$v^\mu D_\mu(\xi_\nu(x) v^\nu) = v^\mu(v^\nu D_\mu \xi_\nu + \xi_\nu D_\mu v^\nu) = v^\mu v^\nu D_\mu \xi_\nu = 0$$

where the last equality is due to the fact that $D_\mu \xi_\nu = -D_\nu \xi_\mu$. Thus isometries corresponding to the Killing vectors may be used to construct conserved quantities

Recall from section 1.6 that if a tensor T is invariant under some family of diffeomorphisms ϕ so that $\phi^* T = T$, then the Lie derivative vanishes, $\mathcal{L}_V T = 0$. This holds for isometries, in which case this becomes

$$\mathcal{L}_K g_{\mu\nu} = 0$$

which is the Killing equation in another form. This helps illuminate the fact that Killing fields are the infinitesimal generators of isometries. Moving the metric along a flow generated by a Killing vector will preserve the metric. Therefore, moving all points of an object the same distance along a Killing vector field will preserve distances on that object.

Lastly, there is the question of the number of solutions to the Killing equation. In general there are a solutions so that the vector fields are labeled $K^\mu_{(a)}$. What is a? Intuitively, in the neighborhood of a point, a maximally symmetric D-dimensional geometry will have $\frac{1}{2}D(D-1)$ rotations and D translations, for a total of $a = \frac{1}{2}D(D+1)$ possible Killing fields. Of course, not all spaces are maximally symmetric, but those that are have some nice properties.

5.4 Maximally symmetric spaces

5.5 Geodesic deviation

Geodesic deviation is nothing more than a fancy term for describing how badly the parallel postulate fails. Recall that in Euclidean space, parallel lines remain parallel to infinity, a postulate of the geometry. An equivalent way of describing that the lines remain parallel is to state that the separation between them does not change identically. In curved spacetime, there is no reason for this to be true, and in fact it isn't, in general. Let's quantify that. Consider a one-parameter family of geodesics $\gamma_s(t)$, with $s, t \in \mathbb{R}$. This means that t measures motion along the geodesic while s defines the particular geodesic being considered. This defines a $2D$ surface $x^\mu(s,t) \in M$. Two vector fields on this surface are

$$T^\mu = \frac{\partial x^\mu}{\partial t} \qquad \text{and} \qquad S^\mu = \frac{\partial x^\mu}{\partial s}$$

where T^μ is the vector field tangent to the geodesics and S^μ is the vector field perpendicular to the geodesic. S^μ can naturally be thought of as the separation between geodesics. Let's see how S^μ changes as we move along a given geodesic. This is of course given by parallel transporting S along T:

$$V^\mu = (\nabla_T S)^\mu = \nabla_\rho T^\rho S^\mu$$

which defines a vector V^μ, the "relative velocity of geodesics". It is possible to similarly define a "relative acceleration of geodesics" by parallel transporting V along T:

$$A^\mu = (\nabla_T V)^\mu = T^\rho \nabla_\rho V^\mu = T^\rho \nabla_\rho \left(T^\sigma \nabla_\sigma S^\mu\right)$$

Now let's examine the Lie derivative of S along T, specifying to the Levi-Civita connection (which is torsion-free):

$$\mathcal{L}_T S = [S, T] = S^\mu T^\nu [e_\mu, e_\nu] = 0$$

where in the last step, we've recognized the fact that local basis vector commute. This allows us to write

$$[S,T]^\mu = 0 = S^\lambda \nabla_\lambda T^\mu - T^\lambda \nabla_\lambda S^\mu \qquad \to \qquad S^\lambda \nabla_\lambda T^\mu = T^\lambda \nabla_\lambda S^\mu$$

Combining this expression with the relative acceleration expression gives

$$\begin{aligned}
A^\mu &= T^\rho \nabla_\rho \left(T^\sigma \nabla_\sigma S^\mu\right) \\
&= T^\rho \nabla_\rho \left(S^\sigma \nabla_\sigma T^\mu\right) \\
&= T^\rho (\nabla_\rho S^\sigma)(\nabla_\sigma T^\mu) + T^\rho S^\sigma \nabla_\rho \nabla_\sigma T^\mu \\
&= T^\rho (\nabla_\rho S^\sigma)(\nabla_\sigma T^\mu) + T^\rho S^\sigma \left(\nabla_\sigma \nabla_\rho T^\mu + R^\mu{}_{\nu\rho\sigma} T^\nu\right) \\
&= T^\rho (\nabla_\rho S^\sigma)(\nabla_\sigma T^\mu) + T^\rho S^\sigma \nabla_\sigma \nabla_\rho T^\mu + R^\mu{}_{\nu\rho\sigma} T^\nu T^\rho S^\sigma \\
&= T^\rho (\nabla_\rho S^\sigma)(\nabla_\sigma T^\mu) + S^\sigma \nabla_\sigma (T^\rho \nabla_\rho T^\mu) - S^\sigma (\nabla_\sigma T^\rho)(\nabla_\rho T^\mu) + R^\mu{}_{\nu\rho\sigma} T^\nu T^\rho S^\sigma \\
&= T^\rho (\nabla_\rho S^\sigma)(\nabla_\sigma T^\mu) - T^\sigma (\nabla_\sigma S^\rho)(\nabla_\rho T^\mu) + R^\mu{}_{\nu\rho\sigma} T^\nu T^\rho S^\sigma \\
&= R^\mu{}_{\nu\rho\sigma} T^\nu T^\rho S^\sigma
\end{aligned}$$

This

$$A^\mu = \frac{D^2}{dt^2} S^\mu = R^\mu{}_{\nu\rho\sigma} T^\nu T^\rho S^\sigma$$

is the *geodesic deviation equation*. We will use it to describe how collections of particles behave in the presence of gravitational waves. Of course, things can be more complicated when describing a multidimensional set of geodesics (so there is more than one parameter s that describes the families of geodesics). These are known as geodesic congruences.

6 Tetrad formalism*

6.1 Noncoordinate bases

Up until now, we've been simply using derivatives and gradients of the coordinate functions x^μ as the bases to expand vectors and forms, respectively, Of course, this is not the only choice that may be made, and indeed in some cases, may actually complicate things. We may instead use non-coordinate (or non-holonomic) bases; at each point p in the manifold it is possible to choice orthonormal coordinates $\hat{e}_{(a)}$ that satisfy

$$g(\hat{e}_{(a)}, \hat{e}_{(b)}) = \eta_{ab}$$

where the (a) indices indicate that the basis vectors are unrelated to the coordinate functions. These may also be known as tetrads, vielbein or vierbein. All vectors may be decomposed in terms of any basis we choose, be that a coordinate or non-coordinate basis. This includes coordinate basis vectors, which may be decomposed in terms of non-coordinate basis vectors as

$$\hat{e}_{(\mu)} = e_\mu{}^a \hat{e}_{(a)}$$

where the $e_\mu{}^a$ are $n \times n$ matrices. The converse in terms of inverse matrices is

$$\hat{e}_{(a)} = e^\mu{}_a \hat{e}_{(\mu)}$$

Notice that the spacetime index always goes first, followed by the 'internal' non-coordinate indices. Since $e_\mu{}^a$ and $e^\mu{}_a$ are inverses, they satisfy

$$e^\mu{}_a e_\nu{}^a = \delta^\mu_\nu \qquad \text{and} \qquad e_\mu{}^a e^\mu{}_b = \delta^a_b$$

Lastly, the metric may be expanded in terms of these matrices

$$\begin{aligned}
g(\hat{e}_{(a)}, \hat{e}_{(b)}) &= \eta_{ab} \\
g(e^\mu{}_a \hat{e}_{(\mu)}, e^\nu{}_a \hat{e}_{(\nu)}) &= \eta_{ab} \\
e^\mu{}_a e^\nu{}_a g(\hat{e}_{(\mu)}, \hat{e}_{(\nu)}) &= \eta_{ab} \\
e^\mu{}_a e^\nu{}_a g_{\mu\nu} &= \eta_{ab} \\
g_{\mu\nu} &= e_\mu{}^a e^b_\nu \eta_{ab}
\end{aligned}$$

A similar non-coordinate basis of one-forms $\hat{\theta}^{(a)}$ may be chosen. Expanding the coordinate basis one-forms in terms of these non-coordinate one-forms yields

$$\hat{\theta}^{(\mu)} = e^\mu{}_a \hat{\theta}^{(a)} \qquad \text{and} \qquad \hat{\theta}^{(a)} = e_\mu{}^a \hat{\theta}^{(\mu)}$$

These may be chosen such that the usual orthogonal inner product is satisfied

$$\hat{\theta}^{(a)}(\hat{e}_{(b)}) = \delta^a_b$$

As stated (and restated) any vector or form may be decomposed in terms of the noncoordinate basis, which may be switched into the coordinate basis by means of the matrices $e^\mu{}_a$ (which, for brevity, I'll start referring to as the tetrad components)

$$V = V^\mu \hat{e}_\mu = V^\mu e_\mu{}^a \hat{e}_a = V^a \hat{e}_a \qquad \Rightarrow \qquad V^a = V^\mu e_\mu{}^a$$

Thus the tetrad components allow us to switch back and forth from Latin and Greek indices. The manifold metric $g_{\mu\nu}$ and the coordinate metric η_{ab} may be used to raise and lower indices, as usual

$$e_\mu{}^a = g_{\mu\nu}\eta^{ab}e^\nu{}_b$$

Now, it's worth noting that although $\{\mu\}$ and $\{a\}$ live in different spaces, the $e_\mu{}^a$ may be thought of as a $\binom{1}{1}$ tensor:

$$e = e_\mu{}^a \mathrm{d}x^\mu \otimes \hat{e}_a$$

In particular, it's nothing more than the identity operator, since acting with the tetrad components on vector components returns those same vector components, but in a different basis. Since it's a tensor, it must have a well-defined transformation property. We already know how the $\{\mu\}$ indices transform under general coordinate transformations (GCTs), how do the $\{a\}$ indices transform? Recall that these coordinates are flat; their metric is simply the Minkowski metric. Furthermore, they may be transformed separately from the coordinates. What is the most general transformation that may be executed upon flat coordinates? Simply the Lorentz transformation:

$$\hat{e}_{a'} = \Lambda^a{}_{a'}(x)\hat{e}_a$$

These are what are known as local Lorentz transformations (LLTs); note that the matrices representing the Lorentz transformation are spacetime dependent. Furthermore, these must leave the metric invariant (as a normal Lorentz transformation would)

$$\Lambda^a{}_{a'}\Lambda^b{}_{b'}\eta_{ab} = \eta_{a'b'}$$

Thus we know how each type of index transforms and therefore a mixed tensor like $e^\mu{}_a$ transforms as

$$e^{\mu'}{}_{a'} = \Lambda^a{}_{a'}\frac{\partial x^{\mu'}}{\partial x^\mu}e^\mu{}_a$$

6.2 Spin connections

Recall that the covariant derivative was defined to have the proper tensor transformation property; the partial derivative alone did not transform properly, so a 'corrective factor', the connection coefficients $\Gamma^\rho_{\mu\nu}$ was added to absorb the non-tensorial part. A similar exercise may be undertaken when computing the covariant derivative of tensor components decomposed in a non-coordinate basis. Consider for example a $\binom{1}{1}$ tensor X:

$$\nabla_\mu X^a{}_b = \partial X^a{}_b + \omega_\mu{}^a{}_c X^c b - \omega_\mu{}^c{}_b X^a{}_c$$

Here the $\omega_\mu{}^c{}_b$ are components of the *spin connection.* Note that spin connection factors are added/subtracted in the usual sense (plus for vector-like indices, minus for form-like indices). What is the relation between the Γ and ω? Let's expand the covariant derivative in each basis to find out

$$
\begin{aligned}
\nabla X = \nabla_\mu X^\nu \partial_\nu \otimes \mathrm{d}x^\mu &= \left(\partial_\mu X^\nu + \Gamma^\nu_{\mu\rho} X^\rho\right) \mathrm{d}x^\mu \otimes \partial_\nu \\
\nabla X = \nabla_\mu X^a \mathrm{d}x^\mu \otimes \hat{e}_a &= \left(\partial_\mu X^a + \omega_\mu{}^a{}_b X^b\right) \mathrm{d}x^\mu \otimes \hat{e}_a \\
&= \left(\partial_\mu \left(e_\nu{}^a X^\nu\right) + \omega_\mu{}^a{}_b \left(e_\rho{}^b X^\rho\right)\right) \mathrm{d}x^\mu \otimes \left(e^\sigma{}_a \partial_\sigma\right) \\
&= e^\sigma{}_a \left(X^\nu \left(\partial_\mu e_\nu{}^a\right) + e_\nu{}^a \partial_\mu X^\nu + \omega_\mu{}^a{}_b \left(e_\rho{}^b X^\rho\right)\right) \mathrm{d}x^\mu \otimes \partial_\sigma \\
&= e^\sigma{}_a e_\nu{}^a \left(X^\rho e^\nu{}_a \left(\partial_\mu e_\rho{}^a\right) + \partial_\mu X^\nu + \omega_\mu{}^a{}_b e^\nu{}_a e_\rho{}^b X^\rho\right) \mathrm{d}x^\mu \otimes \partial_\sigma \\
&= \delta^\sigma_\nu \left(\partial_\mu X^\nu + \left(e^\nu{}_a \left(\partial_\mu e_\rho{}^a\right) + \omega_\mu{}^a{}_b e^\nu{}_a e_\rho{}^b\right) X^\rho\right) \mathrm{d}x^\mu \otimes \partial_\sigma \\
&= \left(\partial_\mu X^\nu + \left(e^\nu{}_a \left(\partial_\mu e_\rho{}^a\right) + \omega_\mu{}^a{}_b e^\nu{}_a e_\rho{}^b\right) X^\rho\right) \mathrm{d}x^\mu \otimes \partial_\mu
\end{aligned}
$$

Comparing the two expressions, we see that

$$
\Gamma^\nu_{\mu\rho} = e^\nu{}_a \left(\partial_\mu e_\rho{}^a\right) + \omega_\mu{}^a{}_b e^\nu{}_a e_\rho{}^b
$$

This expression has an important conclusion. Compute the covariant derivative of the tetrad components

$$
\begin{aligned}
\nabla_\mu e_\nu{}^a &= \partial_\mu e_\nu{}^a - \Gamma^\rho_{\mu\nu} e_\rho{}^a + \omega_\mu{}^a{}_b e_\nu{}^b \\
&= \partial_\mu e_\nu{}^a - \left(e^\rho{}_a \left(\partial_\mu e_\nu{}^a\right) + \omega_\mu{}^a{}_b e^\rho{}_a e_\nu{}^b\right) e_\rho{}^a + \omega_\mu{}^a{}_b e_\nu{}^b \\
&= \partial_\mu e_\nu{}^a - \left(\partial_\mu e_\nu{}^a\right) - \omega_\mu{}^a{}_b e_\nu{}^b + \omega_\mu{}^a{}_b e_\nu{}^b \\
&= 0
\end{aligned}
$$

Thus we see that the covariant derivative of the tetrad components vanish identically with this definition of the spin connection. A couple of points: (1) since the tetrad components may be viewed as nothing more than the identity map, it makes sense that its covariant derivative should vanish identically, (2) this identity (sometimes known as the tetrad postulate) is satisfied without reference to a particular connection (i.e., no assumptions made about torsion, etc.) and is thus very general. Lastly, this expression may be inverted to give the spin connection in terms of the connection coefficients:

$$
\begin{aligned}
\Gamma^\nu_{\mu\rho} &= e^\nu{}_a \left(\partial_\mu e_\rho{}^a\right) + \omega_\mu{}^a{}_b e^\nu{}_a e_\rho{}^b \\
e_\nu{}^a \Gamma^\nu_{\mu\rho} e^\rho{}_b &= \left(e_\nu{}^a e^\nu{}_a\right) \left(\partial_\mu e_\rho{}^a\right) e^\rho{}_b + \omega_\mu{}^a{}_b \left(e_\nu{}^a e^\nu{}_a\right) \left(e_\rho{}^b e^\rho{}_b\right) \\
e_\nu{}^a \Gamma^\nu_{\mu\rho} e^\rho{}_b &= \left(\partial_\mu e_\rho{}^a\right) e^\rho{}_b + \omega_\mu{}^a{}_b \\
\Rightarrow \omega_\mu{}^a{}_b &= e_\nu{}^a \Gamma^\nu_{\mu\rho} e^\rho{}_b - e^\rho{}_b \left(\partial_\mu e_\rho{}^a\right)
\end{aligned}
$$

Now, recall that the connection coefficients did not transform tensorially; the same is true of the spin connection components. The covariant derivative is already invariant under GCTs (by construction); it should also be invariant under LLTs:

$$
\nabla_\mu (X_{a'}) = \nabla_\mu (\Lambda^a{}_{a'}(x) X_a) = \Lambda^a{}_{a'}(x) \nabla_\mu X_a + X^a \nabla_\mu \Lambda^a{}_{a'}(x)
$$

which is only satisfied when $\nabla_\mu \Lambda^a{}_{a'}(x) = 0$. This imposes the following restriction on the spin connection:

$$\begin{aligned}
\nabla_\mu \Lambda^a{}_b(x) = 0 &= \partial_\mu \Lambda^a{}_b(x) + \omega_\mu{}^a{}_c \Lambda^c{}_b - \omega_\mu{}^c{}_b \Lambda^a{}_c \\
-\omega_\mu{}^a{}_c \Lambda^c{}_b \Lambda^b{}_d &= \left(\partial_\mu \Lambda^a{}_b\right) \Lambda^b{}_d - \omega_\mu{}^c{}_b \Lambda^a{}_c \Lambda^b{}_d \\
\omega_\mu{}^a{}_c \delta^c_d &= -\left(\partial_\mu \Lambda^a{}_b\right) \Lambda^b{}_d + \omega_\mu{}^c{}_b \Lambda^a{}_c \Lambda^b{}_d \\
\omega_\mu{}^a{}_d &= \omega_\mu{}^c{}_b \Lambda^a{}_c \Lambda^b{}_d - \left(\partial_\mu \Lambda^a{}_b\right) \Lambda^b{}_d
\end{aligned}$$

Thus the presence of the second factor clearly indicates that the spin connection transforms non-tensorially.

Let's now specify the spin connection in the case of the Levi-Civita connection. First, since the derivative (both the partial and covariant varieties) of the flat metric η_{ab} vanishes, this enforces a symmetry condition on the spin connection components:

$$\begin{aligned}
\nabla_\mu \eta_{ab} = 0 &= \partial_\mu \eta_{ab} - \omega_{\mu a}{}^c \eta_{cb} - \omega_{\mu b}{}^c \eta_{ac} \\
0 &= -\omega_{\mu a}{}^c \eta_{cb} - \omega_{\mu b}{}^c \eta_{ac} \\
\omega_{\mu ab} &= -\omega_{\mu ba}
\end{aligned}$$

Thus we see that the spin connection coefficients are antisymmetric in the non-coordinate indices. It is possible to derive explicit expressions for the spin connection coefficients. Starting with metric compatibility, contract the covariant derivative with the metric, expanding using the tetrad components and spin connection:

$$\begin{aligned}
0 &= \nabla^\alpha g_{\alpha\beta} \\
0 &= g^{\rho\alpha} \nabla_\rho g_{\alpha\beta} \\
&= g^{\rho\alpha} \left[\partial_\rho g_{\alpha\beta} - \Gamma^\sigma_{\rho\alpha} g_{\sigma\beta} - \Gamma^\sigma_{\rho\beta} g_{\sigma\alpha}\right] \\
&= \\
&= e^\alpha{}_b e^\beta{}_c \left(\partial_\alpha e_{\beta a} - \partial_\beta e_{\alpha a} + \omega_{\alpha a}{}^d e_{\beta d} - \omega_{\beta a}{}^d e_{\alpha d}\right)
\end{aligned}$$

6.3 Cartan geometry

Let's now shift our view on these oddly mixed tensors, such as the tetrad components $e_\mu{}^a$. As described in the previous section, the tetrad components can be viewed as a $\binom{1}{1}$ tensor; our shift consists in thinking of the spacetime indices as designating the tensor as a 1-form that takes vector values in the 'internal', non-coordinate basis. As such, these generalized tensors $X_{\mu_1 \ldots \mu_n}{}^{a_1 \ldots a_p}{}_{b_1 \ldots b_q}$ may be thought of as $\binom{p}{q}$-valued n-forms. As a few examples, $e_\mu{}^a$ are vector-valued 1-forms, and $A_{\mu\nu}{}^a{}_b$ are $\binom{1}{1}$ tensor-valued 2-form. Thus even simple 1-forms may be thought of as scalar-valued forms.

This shift of perspective to viewing the generalized tensor as tensor bundle-valued forms leads to the question as to whether a generalized exterior derivative may be taken. The answer is yes, but of course with certain modifications to ensure the appropriate transformation rule. A first, naive attempt at computing the exterior derivative might be

$$(DX)_{\mu\nu}{}^a = \partial_\mu X_\nu{}^a - \partial_\nu X_\mu{}^a$$

But does this transform properly under both GCTs and LLTs? Under GCTs the answer is yes, for the same reason that the ordinary exterior derivative does (namely the antisymmetrization of the spacetime indices). Under LLTs, the answer is not quite; the exterior derivative should transform as

$$(DX)_{\mu\nu}{}^{a'} = \Lambda^{a'}{}_a (DX)_{\mu\nu}{}^a$$

but that clearly doesn't happen

$$\begin{aligned}
(DX)_{\mu\nu}{}^{a'} &= \partial_\mu X_\nu{}^{a'} - \partial_\nu X_\mu{}^{a'} \\
&= \partial_\mu \left(\Lambda^{a'}{}_a X_\nu{}^a\right) - \partial_\nu \left(\Lambda^{a'}{}_a X_\mu{}^a\right) \\
&= \left(\partial_\mu \Lambda^{a'}{}_a\right) X_\nu{}^a + \Lambda^{a'}{}_a \left(\partial_\mu X_\nu{}^a\right) - \left(\partial_\nu \Lambda^{a'}{}_a\right) X_\mu{}^a - \Lambda^{a'}{}_a \left(\partial_\nu X_\mu{}^a\right) \\
&= \Lambda^{a'}{}_a \left(\partial_\mu X_\nu{}^a - \partial_\nu X_\mu{}^a\right) + \left(\partial_\mu \Lambda^{a'}{}_a\right) X_\nu{}^a - \left(\partial_\nu \Lambda^{a'}{}_a\right) X_\mu{}^a \\
&\neq \Lambda^{a'}{}_a \left(\partial_\mu X_\nu{}^a - \partial_\nu X_\mu{}^a\right)
\end{aligned}$$

Let's instead try the same expression using the covariant derivative:

$$\begin{aligned}
(DX)_{\mu\nu}{}^a &= \nabla_\mu X_\nu{}^a - \nabla_\nu X_\mu{}^a \\
&= \partial_\mu X_\nu{}^a + \omega_\mu{}^a{}_b X_\nu{}^b - \partial_\nu X_\mu{}^a - \omega_\nu{}^a{}_b X_\mu{}^b \\
&= \partial_\mu X_\nu{}^a - \partial_\nu X_\mu{}^a + \omega_\mu{}^a{}_b X_\nu{}^b - \omega_\nu{}^a{}_b X_\mu{}^b \\
&= (\mathrm{d}X)_{\mu\nu}{}^a + (\omega \wedge X)_{\mu\nu}{}^a
\end{aligned}$$

In coordinate-free notation this becomes

$$DX = \mathrm{d}X + \omega \wedge X$$

It can be shown (after some lengthy algebra) that this exhibits the appropriate transformation rule.

Now that we have a covariant way to take derivatives of these mixed tensors, let's recast some of the geometric structures we built using coordinate bases, such the torsion and Riemann tensors. Let's start with something basic; take the exterior covariant derivative of the tetrad components:

$$\begin{aligned}
(De)_{\mu\nu}{}^a &= \partial_\mu e_\nu{}^a - \partial_\nu e_\mu{}^a + \omega_\mu{}^a{}_b e_\nu{}^b - \omega_\nu{}^a{}_b e_\mu{}^b \\
&= e_\rho{}^a(\underbrace{e^\rho{}_a \partial_\mu e_\nu{}^a + \omega_\mu{}^a{}_b e^\rho{}_a e_\nu{}^b}_{\Gamma^\rho_{\mu\nu}} - \underbrace{(e^\rho{}_a \partial_\nu e_\mu{}^a + \omega_\nu{}^a{}_b e^\rho{}_a e_\mu{}^b)}_{\Gamma^\rho_{\nu\mu}}) \\
&= e_\rho{}^a T_{\mu\nu}{}^\rho
\end{aligned}$$

Thus we've identified the torsion tensor in terms of noncoordinate basis as simply

$$T^a = \mathrm{d}e^a + \omega^a{}_b \wedge e^b$$

where the spacetime indices have been suppressed (by convention). The other object of which we can take the exterior covariant derivative simply is the spin connection itself:

$$(D\omega_b)_{\mu\nu}{}^a = \partial_\mu \omega_\nu{}^a{}_b - \partial_\nu \omega_\mu{}^a{}_b + \omega_\mu{}^a{}_c \omega_\nu{}^c{}_b - \omega_\nu{}^a{}_b \omega_\mu{}^c{}_b = R^a{}_{b\mu\nu}$$

where we've identified the Riemann tensor through a bit of malice of forethought (the algebra to prove this is rather nasty). We have thus arrived at what are known as the *Cartan structure equations*, in summary

$$\begin{aligned} T^a &= \mathrm{d}e^a + \omega^a{}_b \wedge e^b \\ R^a{}_b &= \mathrm{d}\omega^a{}_b + \omega^a{}_c \wedge \omega^c{}_b \end{aligned}$$

Lastly, these expressions lend themselves to expressing identities that previously were esoteric, in particular the Bianchi identities. The exterior covariant derivative of the torsion yields

$$(DT)_{\rho\mu\nu}{}^a = (\mathrm{d}^2 e)_{\rho\mu\nu}{}^a + \omega^a{}_c \mathrm{d}e^c$$

Part II
Gravitation

7 The Einstein Field Equations

7.1 Derivation

There are three main ingredients that go into constructing the Einstein Field Equations.

1. We seek to generalize the Newtonian Gauss law, $\nabla^2\Phi = -4\pi G_N \rho$; whatever our solution, it should reduce to this law in flat (or nearly-flat) spacetime

2. Since the metric may be brought into normal form and have vanishing first derivatives at a point p by a judicious choice of coordinates (and this is always true), the EOMs for the metric tensor must at least involve second derivatives of the metric.

3. Energy-momentum conservation should be encoded. In flat spacetime, this has the form $\partial_\mu T^{\mu\nu} = 0$; this may be "promoted" to curved spacetime by upgrading the partial derivative to a covariant derivative, $\nabla_\mu T^{\mu\nu} = 0$

Since $\Phi \sim g$ and $\rho \sim T$ in a general sense, (1) seems to indicate that the EOMs should have the schematic form

$$Lg = \kappa T$$

where L is some differential operator acting on the metric. Furthermore, (2) suggests that $L \sim \nabla^2$. Therefore the LHS of the above equation should be built out of tensors that contain second derivatives of the metric. The Riemann tensor is one such tensor. However, it has the wrong number of indices; the Ricci tensor, on the other hand does not. Therefore, a good first guess is

$$R_{\mu\nu} = \kappa T_{\mu\nu}$$

However, this guess is incompatible with (3); taking the divergence of both sides gives

$$\nabla^\mu R_{\mu\nu} = \kappa \nabla^\mu T_{\mu\nu} = 0$$

This in general is not true; what is true however is

$$\nabla^\mu R_{\mu\nu} = \frac{1}{2}\nabla_\nu R$$

as shown in section 5.2. This fact was employed to construct the Einstein tensor, which by definition satisfies the requisite divergence condition. Therefore the (correct) guess for the EFEs is

$$G_{\mu\nu} + f_{\mu\nu} = \kappa T_{\mu\nu}$$

where $f_{\mu\nu}$ is any divergenceless tensor. Typically, this is taken to be 0 (in the simplest case) or $f_{\mu\nu} = \Lambda g_{\mu\nu}$. We'll usually toss this term, but specific note will be made when it is employed. Thus the EFEs are

$$R_{\mu\nu} - \frac{1}{2}R g_{\mu\nu} = \kappa T_{\mu\nu}$$

with κ a normalization to be derived in a proceeding section. This equation may be rearranged to give $R_{\mu\nu}$ as a function of $T_{\mu\nu}$ (which is ultimately what we want, since we're interested in equations for the components of g). Taking the trace of the above equation yields

$$\begin{aligned} g^{\mu\nu}R_{\mu\nu} - \frac{1}{2}Rg^{\mu\nu}g_{\mu\nu} &= \kappa g^{\mu\nu}T_{\mu\nu} \\ R - \frac{1}{2}R(4) &= \kappa T \\ R &= -\kappa T \end{aligned}$$

Inserting this into the EFEs gives

$$\begin{aligned} R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} &= \kappa T_{\mu\nu} \\ R_{\mu\nu} + \frac{\kappa}{2}Tg_{\mu\nu} &= \kappa T_{\mu\nu} \\ R_{\mu\nu} &= \kappa\left(T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu}\right) \end{aligned}$$

In a vacuum, where $T_{\mu\nu} = T = 0$, this reduces to

$$R_{\mu\nu} = 0$$

Physically, this corresponds to geometries where the geodesics neither compress nor contract, nor shear due to the curvature of spacetime.

7.2 Recovering Newtonian gravity

Let's now find the normalization constant by completing the above program and ensuring that the EFEs satisfy (1). In order to do so, we may expand the metric perturbatively, meaning the dynamics occur on a Minkowski background

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

Recall from Section 4.4 that free-falling particles move along geodesics

$$0 = \frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\sigma}\frac{dx^\nu}{d\tau}\frac{dx^\sigma}{d\tau}$$

where, in this case, the affine parameter λ was taken to be the proper time itself. In the Newtonian limit, particles are treated as moving slowly so that

$$\frac{dx^i}{d\tau} \ll \frac{dt}{d\tau}$$

This simplifies the geodesic equation quite a bit

$$0 = \frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{00}\left(\frac{dt}{d\tau}\right)^2$$

Now let's assume that the test particle is small compared to the mass of the object generating the gravitational field. This corresponds to the static solution and therefore the metric is essentially independent of time; thus $\partial_0 g_{\mu\nu} = 0$ which gives

$$\Gamma^{\mu}_{00} = \frac{1}{2} g^{\mu\nu} \left(-\partial_\nu g_{00} + \partial_0 g_{\nu 0} + \partial_0 g_{0\nu} \right) = -\frac{1}{2} \partial^\mu h_{00}$$

Therefore the geodesic equation becomes

$$\begin{aligned} 0 &= \frac{d^2 x^\mu}{d\tau^2} + \Gamma^{\mu}_{00} \left(\frac{dt}{d\tau} \right)^2 \\ &= \frac{d^2 x^\mu}{d\tau^2} - \frac{1}{2} \partial^\mu h_{00} \left(\frac{dt}{d\tau} \right)^2 \\ &= \frac{d^2 x^0}{d\tau^2} - \frac{1}{2} \partial^0 h_{00} \left(\frac{dt}{d\tau} \right)^2 + \frac{d^2 x^i}{d\tau^2} - \frac{1}{2} \partial^i h_{00} \left(\frac{dt}{d\tau} \right)^2 \end{aligned}$$

This implies two separate equations

$$0 = \frac{d^2 x^0}{d\tau^2} - \frac{1}{2} \partial^0 h_{00} \left(\frac{dt}{d\tau} \right)^2 \qquad\qquad 0 = \frac{d^2 x^i}{d\tau^2} - \frac{1}{2} \partial^i h_{00} \left(\frac{dt}{d\tau} \right)^2$$

Of these, focus on the second; multiplying throughout by $\left(\frac{d\tau}{dt}\right)^2$ gives

$$\frac{d^2 x^i}{dt^2} = \frac{1}{2} \partial^i h_{00}$$

This looks an awful lot like Newton's second law; if we identify the RHS with $-\partial^i \Phi$ this implies that $h_{00} = -2\Phi$.

Now let's assume that the energy-momentum in the near vicinity of some object is due solely to the mass of said object, so that $T_{00} = \rho$ and $T_{0i} = T_{0j} = T_{ij} = 0$. Thus the trace of T is simply $T = -\rho$ and the EFEs become

$$R_{00} = \kappa \left(\rho - \frac{1}{2} (-\rho) g_{00} \right) = \frac{1}{2} \kappa \rho$$

Furthermore, since this is a static system, there should be no time dependence; therefore anything involving time-derivatives should vanish identically. Therefore, the R_{00} component is given in terms of the Riemann tensor by

$$\begin{aligned} R_{00} = R^{\rho}{}_{0\rho 0} &= R^{0}{}_{000} + R^{i}{}_{0i0} \\ &= \partial_i \Gamma^{i}_{00} - \partial_0 \Gamma^{i}_{i0} + \Gamma^{i}_{i\lambda} \Gamma^{\lambda}_{00} - \Gamma^{i}_{0\lambda} \Gamma^{\lambda}_{0i} \\ &= \partial_i \Gamma^{i}_{00} - \partial_t \Gamma^{i}_{j0} + O(h^2) \\ &= \partial_i \frac{1}{2} g^{i\lambda} \left(-\partial_\lambda g_{00} + \partial_0 g_{\lambda 0} + \partial_0 g_{0\lambda} \right) \\ &= \frac{1}{2} \partial_i \left(-\partial^i h_{00} \right) \\ &= -\frac{1}{2} \nabla^2 h_{00} \end{aligned}$$

Inserting this expression into the EFEs and making the identification $h_{00} = -2\Phi$ as derived before gives

$$\begin{aligned}
R_{00} &= \frac{1}{2}\kappa\rho \\
-\frac{1}{2}\nabla^2 h_{00} &= \frac{1}{2}\kappa\rho \\
2\nabla^2\Phi &= -\kappa\rho \\
\nabla^2\Phi &= -\frac{1}{2}\kappa\rho
\end{aligned}$$

This yields the desired Poisson equation when $\kappa = 8\pi G_N$. Thus the EFEs (restoring the requisite factors of c) are

$$\boxed{R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G_N}{c^4}T_{\mu\nu}}$$

It is worth commenting on the nature of the EFEs. These give 10 equations which seemingly determine the 10 d.o.f. of the metric tensor. This however, is not quite right. Remember that the Einstein tensor must also be covariantly divergenceless

$$\nabla_\mu G^{\mu\nu} = 0$$

which yields four constraints. Thus the EFEs do not correspond to 10 independent equations of motion, but rather only $10 - 4 = 6$. This leaves four d.o.f. undetermined; these of course are the d.o.f. that may be freely chosen by coordinate transformations.

This situation is analogous to that of Maxwell electrodynamics. Superficially, the inhomogenenous Maxwell equations

$$j^\alpha = \partial_\beta F^{\beta\alpha} = \partial_\beta\left(\partial^\beta A^\alpha - \partial^\alpha A^\beta\right) = -\Box A^\alpha + \partial^\alpha\partial_\beta A^\beta$$

correspond to four equations of motion, seemingly for the four d.o.f. of the four-vector potential A^μ. However, the four-vector potential also satisfies the Bianchi identity, of which there is one, meaning $4 - 1 = 3$ independent EOM. This free d.o.f. corresponds to the freedom of gauge choice.

8 Lagrangian Formulation

8.1 Einstein-Hilbert action

The derivation of the EFEs in the previous section was somewhat heuristic. Although faithful to the origins of GR, there exist more systematic methods of deriving the EOMs for the metric. These employ the use of the action principle, as in classical field theory. Recall that in classical field theory, the EOMs for the field Φ are given by the Euler-Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial \Phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi)} \right) = 0$$

where $\mathcal{L}$ is a Lagrangian density given by the action

$$S = \int \mathcal{L} d^d x$$

In plain classical field theory, the above action integral is sufficient since the integration measure is a scalar, and thus manifestly invariant. In curved spacetime, it is not, since it is technically a tensor density and thus must be multiplied by a factor $\sqrt{|g|}$ to be made manifestly invariant (see Section 2.3). Therefore, the action integral becomes

$$S = \int \hat{\mathcal{L}} \sqrt{-|g|} d^d x$$

Note the ˆ on the Lagrangian density; this is to emphasize the fact that $\mathcal{L}$ is a *density* whose product with $d^d x$ is properly a tensor. Since $\sqrt{|g|} d^d x$ is a scalar, then $\hat{\mathcal{L}}$ is properly a scalar as well. Similarly, the EL equations undergo a slight modification in curved spacetime

$$\frac{\partial \hat{L}}{\partial \Phi} - \nabla_\mu \left(\frac{\partial \hat{L}}{\partial(\nabla_\mu \Phi)} \right) = 0$$

Now let's turn to GR. The dynamical field is now the metric $g_{\mu\nu}$. What scalars may be built out of the metric? Recall that it is always possible at a point p to choose coordinates such that the metric assumes its canonical form and its first derivatives vanish. Thus any scalar we employ should not depend on the metric or its first derivatives directly.

We must therefore deal with second derivatives of the metric, such as those contained in the Riemann tensor. However, as discussed previously, in the Levi-Civita connection, the only unique scalar built out of the Riemann tensor is the Ricci scalar R (all other contractions are proportional to R). Thus the simplest (and it turns out, correct) guess for the Lagrangian $\hat{\mathcal{L}}$ is simply

$$S_{\text{EH}} = \int \sqrt{-g} R d^d x$$

This is known as the Einstein-Hilbert action. To find the EOMs, we may vary this action and employ the action principle:

$$\frac{\delta}{\delta g_{\rho\tau}} S_H = 0 = \frac{\delta}{\delta g_{\rho\tau}} \int \left(\sqrt{-g} R \right) d^d x$$

This would be a mess to evaluate by brute force so let's be clever. First, let's note that R is composed of exactly two derivatives of the metric, whether they be two first derivatives multiplied together or one second derivative. Variations with respect to $\delta g_{\rho\tau}$ produce terms of the form $\sqrt{-g}H^{\rho\tau}$; thus $H^{\rho\tau}$ must be symmetric under $\rho \leftrightarrow \tau$ interchange. Thus we can write in full generality

$$\frac{\delta}{\delta g_{\rho\tau}}\delta\left(\sqrt{-g}R\right)d^dx = \sqrt{-g}H^{\rho\tau} = \frac{\delta}{\delta g_{\rho\tau}}\int \mathcal{H}^{\gamma\alpha\beta\sigma\mu\nu}(g_{\mu\nu})\partial_\gamma g_{\alpha\beta}\partial_\sigma g_{\mu\nu}d^dx + \frac{\delta}{\delta g_{\rho\tau}}\int \mathcal{G}^{\alpha\beta\mu\nu}(g_{\mu\nu})\partial_\alpha\partial_\beta g_{\mu\nu}d^dx$$

where $\mathcal{H}$ and $\mathcal{G}$ are functions of $g_{\mu\nu}$. Now we can integrate the second term by parts. This gives a total derivative (which vanishes on the boundary) minus a term which has the partial derivative moved onto the $\mathcal{G}$ function; but this simply creates terms that are quadratic in first derivatives of $g_{\mu\nu}$, which may be absorbed into the $\mathcal{H}$ term with no loss of generality. Now we can perform the variation:

$$\begin{aligned}
\int \frac{\delta}{\delta g_{\rho\tau}}\left(\sqrt{-g}R\right)d^dx &= \frac{\delta}{\delta g_{\rho\tau}}\int \mathcal{H}^{\gamma\alpha\beta\sigma\mu\nu}(g_{\epsilon\zeta})\partial_\gamma g_{\alpha\beta}\partial_\sigma g_{\mu\nu}d^dx \\
&= \int \frac{\delta \mathcal{H}^{\gamma\alpha\beta\sigma\mu\nu}(g_{\epsilon\zeta})}{\delta g_{\rho\tau}}\partial_\gamma g_{\alpha\beta}\partial_\sigma g_{\mu\nu}d^dx + 2\int \mathcal{H}^{\gamma\alpha\beta\sigma\mu\nu}(g_{\epsilon\zeta})\partial_\gamma g_{\alpha\beta}\frac{\delta}{\delta g_{\rho\tau}}(\partial_\sigma g_{\mu\nu})d^dx \\
&= \int \frac{\delta \mathcal{H}^{\gamma\alpha\beta\sigma\mu\nu}(g_{\epsilon\zeta})}{\delta g_{\rho\tau}}\partial_\gamma g_{\alpha\beta}\partial_\sigma g_{\mu\nu}d^dx - 2\int \partial_\sigma\left(\mathcal{H}^{\gamma\alpha\beta\sigma\mu\nu}(g_{\epsilon\zeta})\partial_\gamma g_{\alpha\beta}\right)\frac{\delta g_{\mu\nu}}{\delta g_{\rho\tau}}d^dx \\
&= \int \frac{\partial \mathcal{H}^{\gamma\alpha\beta\sigma\mu\nu}}{\partial g_{\epsilon\zeta}}\frac{\delta g_{\epsilon\zeta}}{\delta g_{\rho\tau}}\partial_\gamma g_{\alpha\beta}\partial_\sigma g_{\mu\nu}d^dx - 2\int \partial_\sigma\left(\mathcal{H}^{\gamma\alpha\beta\sigma\mu\nu}(g_{\epsilon\zeta})\partial_\gamma g_{\alpha\beta}\right)\frac{\delta g_{\mu\nu}}{\delta g_{\rho\tau}}d^dx \\
&= \frac{\partial \mathcal{H}^{\gamma\alpha\beta\sigma\mu\nu}}{\partial g_{\rho\tau}}\partial_\gamma g_{\alpha\beta}\partial_\sigma g_{\mu\nu} - 2\partial_\sigma\left(\mathcal{H}^{\gamma\alpha\beta\sigma\rho\tau}(g_{\epsilon\zeta})\partial_\gamma g_{\alpha\beta}\right) \\
&= \frac{\partial \mathcal{H}^{\gamma\alpha\beta\sigma\mu\nu}}{\partial g_{\rho\tau}}\partial_\gamma g_{\alpha\beta}\partial_\sigma g_{\mu\nu} - 2\frac{\partial \mathcal{H}^{\gamma\alpha\beta\sigma\rho\tau}}{\partial g_{\epsilon\zeta}}\partial_\sigma g_{\epsilon\zeta}\partial_\gamma g_{\alpha\beta} - 2\mathcal{H}^{\gamma\alpha\beta\sigma\rho\tau}(g_{\epsilon\zeta})\partial_\sigma\partial_\gamma g_{\alpha\beta}
\end{aligned}$$

Thus we see that $H^{\rho\tau}$ depends only on g (through its dependence on $\mathcal{H}$), on two first derivatives of g (and no second derivatives) or one second derivative of g (and no first derivatives). Furthermore, it has exactly two derivatives (i.e., no plain g terms). This greatly constraints $H^{\rho\tau}$. We can constrain it further by choosing a particular Lorentz frame where normal coordinates are employed. Therefore the first derivatives of the metric vanish. Thus $H^{\mu\nu}$ can only really depend on the metric or its second derivatives:

$$H^{\mu\nu} = c_1 R^{\mu\nu} + c_2 R g^{\mu\nu}$$

Now let the metric perturbation be a gauge transformation $\delta g_{\mu\nu} = -2\nabla_{(\mu}\xi_{\nu)}$ (This is derived in the following section). Then we get

$$\begin{aligned}
\delta S_H = 0 &= \delta \int \left(\sqrt{-g}R\right) d^d x \\
&= \int \sqrt{-g} H^{\mu\nu} \nabla g_{\mu\nu} d^d x \\
&= -2 \int \sqrt{-g} H^{\mu\nu} \nabla_{(\mu} \xi_{\nu)} d^d x \\
&= - \int \sqrt{-g} H^{\mu\nu} \nabla_\mu \xi_\nu d^d x \\
&= - \int \sqrt{-g} \nabla_\mu \left(H^{\mu\nu} \xi_\nu\right) d^d x + \int \sqrt{-g} (\nabla_\mu H^{\mu\nu}) \xi_\nu d^d x \\
&= \int \sqrt{-g} (\nabla_\mu H^{\mu\nu}) \xi_\nu d^d x
\end{aligned}$$

This vanishes in general only if $\nabla_\mu H^{\mu\nu} = 0$. Employing our general form for $H^{\mu\nu}$ and applying this condition, we get

$$\begin{aligned}
0 &= \nabla_\mu H^{\mu\nu} \\
&= c_1 \nabla_\mu R^{\mu\nu} + c_2 \nabla_\mu (g^{\mu\nu} R) \\
c_1 \nabla_\mu R^{\mu\nu} &= c_2 g^{\mu\nu} \nabla_\mu R
\end{aligned}$$

Recall that the Bianchi identity may be contracted to yield the condition $\nabla_\mu R^{\mu\nu} = \frac{1}{2} g^{\mu\nu} \nabla_\mu R$. Thus to satisfy the Bianchi identity, we see that $c_1 = 1$ and $c_2 = -\frac{1}{2}$. But then $H^{\mu\nu} = G^{\mu\nu}$, the Einstein tensor! Thus we recover the vacuum EFEs, $G^{\mu\nu} = 0$. What about the EFEs in non-vacuum? Well, let's consider a slightly expanded action

$$S = \frac{1}{16\pi G_N} S_{\text{EH}} + S_{\text{matter}}$$

The matter action is of the form

$$S_{\text{matter}} = \int \Lambda \sqrt{-g} d^4 x$$

where Λ is any scalar constructed out of matter fields. For example, the corresponding term for Maxwell electrodynamics would be

$$\Lambda = -\frac{1}{4} g_{\mu\alpha} g_{\nu\beta} F^{\mu\nu} F^{\alpha\beta}$$

How is this related to the energy-momentum tensor? And how will this recover the inhomogenous EFEs?

8.2 Gauge invariance

In Maxwell electrodynamics and other Yang-Mills theories, the equations of motion do not correspond one-to-one to the degrees of freedom of the theory. This redundancy manifests itself as invariance under redefinitions corresponding to gauge transformations. By the Noether theorem, symmetries of the action correspond to conserved currents; invariance under gauge transformations leads to charge conservation.

A similar situation exists in general relativity. Physics should be invariant under choice of coordinates; this corresponds to diffeomorphism invariance. As described in Section 2.2, diffeomorphisms amount to coordinate transformations. Consider an infinitesimal coordinate transformation $x'^{\mu} = x^{\mu} + \xi^{\mu}$. Variations of the Einstein-Hilbert Lagrangian with respect to the metric (including such coordinate transformations) yield the vacuum EFEs as described in the previous section; satisfying the EFEs leaves the EH action invariant under variations of the metric. The matter Lagrangian should similarly be unaffected; what sort of consequence will this have? In other words, what are the consequences of setting

$$\delta S_{\text{matter}} = 0$$

Using the general form described in the previous section, variations of the matter Lagrangian with respect to the metric take the form

$$\delta S_{\text{matter}} = \int \left(\frac{\partial(\Lambda\sqrt{-g})}{\partial g_{\mu\nu}} + \partial_\sigma \left(\frac{\partial(\Lambda\sqrt{-g})}{\partial(\partial_\sigma g_{\mu\nu})} \right) \right) \delta g_{\mu\nu} d^4x = \frac{1}{2} \int T^{\mu\nu} \delta g_{\mu\nu} d^4x$$

Let's find out how the metric tensor changes under infinitesimal coordinate transformations

$$\begin{aligned}
g'_{\alpha\beta}(x'^{\sigma}) &= g_{\mu\nu}(x^{\sigma}) \frac{\partial x^{\mu}}{\partial x'^{\alpha}} \frac{\partial x^{\nu}}{\partial x'^{\beta}} \\
&= g_{\mu\nu}(x^{\sigma}) \left(\delta^{\mu}_{\alpha} - \partial_\alpha \xi^{\mu}\right) \left(\delta^{\nu}_{\beta} - \partial_\beta \xi^{\nu}\right) \\
&\approx g_{\mu\nu}(x^{\sigma}) \left(\delta^{\mu}_{\alpha}\delta^{\nu}_{\beta} - \delta^{\mu}_{\alpha}\partial_\beta \xi^{\nu} - \delta^{\nu}_{\beta}\partial_\alpha \xi^{\mu} + O(\xi^2)\right) \\
&= g_{\mu\nu} - g_{\alpha\nu}\partial_\beta \xi^{\nu} - g_{\mu\beta}\partial_\alpha \xi^{\mu}
\end{aligned}$$

Since we would like to compare the components of the metric in the same coordinate system, expand the LHS to give

$$g'_{\alpha\beta}(x'^{\sigma}) = g'_{\alpha\beta}(x^{\sigma} + \xi^{\sigma}) \approx g'_{\alpha\beta}(x^{\sigma}) + \xi^{\sigma} \frac{\partial g'_{\alpha\beta}}{\partial x^{\sigma}} \approx g'_{\alpha\beta}(x^{\sigma}) + \xi^{\sigma}\partial_\sigma g_{\alpha\beta}$$

where in the last step, we recognize the fact that the second term is already first-order in ξ so that $g'^{\alpha\beta} \approx g^{\alpha\beta}$ (since the next term is itself first-order in ξ and thus would be second-order overall). Thus we're left with

$$\begin{aligned}
g'_{\alpha\beta}(x'^{\sigma}) &= g_{\mu\nu}(x^{\sigma}) \frac{\partial x^{\mu}}{\partial x'^{\alpha}} \frac{\partial x^{\nu}}{\partial x'^{\beta}} \\
g'_{\alpha\beta}(x^{\sigma}) + \xi^{\sigma}\partial_\sigma g_{\alpha\beta} &\approx g_{\mu\nu} - g_{\alpha\nu}\partial_\beta \xi^{\nu} - g_{\mu\beta}\partial_\alpha \xi^{\mu} \\
g'_{\alpha\beta}(x^{\sigma}) &= g_{\alpha\beta}(x^{\sigma}) - \xi^{\sigma}\partial_\sigma g_{\alpha\beta} - g_{\alpha\nu}\partial_\beta \xi^{\nu} - g_{\mu\beta}\partial_\alpha \xi^{\mu}
\end{aligned}$$

Therefore, under a coordinate transformation, the effect on the metric $\delta g^{\alpha\beta}$ is

$$\delta g_{\alpha\beta} = -\xi^{\sigma}\partial_\sigma g_{\alpha\beta} - g_{\alpha\mu}\partial_\beta \xi^{\mu} - g_{\mu\beta}\partial_\alpha \xi^{\mu}$$

This can be rewritten into a nicer form after some algebra, similar to that of Section 5.3. Consider first the derivative

$$\partial_\alpha (g_{\mu\beta}\xi^\mu) = \partial_\alpha \xi_\beta = (\partial_\alpha g_{\mu\beta})\xi^\mu + g_{\mu\beta}\partial_\alpha \xi^\mu \qquad \Rightarrow \qquad g_{\mu\beta}\partial_\alpha \xi^\mu = \partial_\alpha \xi_\beta - (\partial_\alpha g_{\mu\beta})\xi^\mu$$

which gives

$$\begin{aligned}
\delta g_{\alpha\beta} &= -\xi^\sigma \partial_\sigma g_{\alpha\beta} - g_{\alpha\mu}\partial_\beta \xi^\mu - g_{\mu\beta}\partial_\alpha \xi^\mu \\
&= -\xi^\mu \partial_\mu g_{\alpha\beta} - \partial_\alpha \xi_\beta + (\partial_\alpha g_{\mu\beta})\xi^\mu - \partial_\beta \xi_\alpha + (\partial_\beta g_{\mu\alpha})\xi^\mu \\
&= -\partial_\alpha \xi_\beta - \partial_\beta \xi_\alpha + \xi^\mu \left(\partial_\mu g_{\alpha\beta} + \partial_\alpha g_{\mu\beta} + \partial_\beta g_{\alpha\mu}\right) \\
&= -\partial_\alpha \xi_\beta - \partial_\beta \xi_\alpha + \xi_\sigma g^{\sigma\mu} \left(\partial_\mu g_{\alpha\beta} + \partial_\alpha g_{\mu\beta} + \partial_\beta g_{\alpha\mu}\right) \\
&= -\partial_\alpha \xi_\beta - \partial_\beta \xi_\alpha + 2\Gamma^\sigma_{\alpha\beta}\xi_\sigma \\
&= -\nabla_\alpha \xi_\beta - \nabla_\beta \xi_\alpha \\
&= -2\nabla_{(\alpha}\xi_{\beta)}
\end{aligned}$$

Thus the transformation of the metric under coordinate transformations is simply

$$g'_{\mu\nu} = g_{\mu\nu} - 2\nabla_{(\alpha}\xi_{\beta)}$$

We can use this to compute the variations w.r.t. the metric:

$$\delta S_{\text{matter}} = \int \left(\frac{\partial(\Lambda\sqrt{-g})}{\partial g_{\mu\nu}} \delta g_{\mu\nu} + \frac{\partial(\Lambda\sqrt{-g})}{\partial(\partial_\sigma g_{\mu\nu})} \delta(\partial_\sigma g_{\mu\nu}) \right) d^4x = \int \left(\frac{\partial(\Lambda\sqrt{-g})}{\partial g_{\mu\nu}} - \partial_\sigma \left(\frac{\partial(\Lambda\sqrt{-g})}{\partial(\partial_\sigma g_{\mu\nu})} \right) \right) \delta g_{\mu\nu} d^4x$$

where the second term was integrated by parts (assuming the variation of the metric vanishes at infinity). Now define $T^{\mu\nu}$ as proportional to the term in parenthesis

$$\frac{1}{2}\sqrt{-g}T^{\mu\nu} = \frac{\partial(\Lambda\sqrt{-g})}{\partial g_{\mu\nu}} - \partial_\sigma \left(\frac{\partial(\Lambda\sqrt{-g})}{\partial(\partial_\sigma g_{\mu\nu})} \right)$$

where the factors of $\frac{1}{2}$ and $\sqrt{-g}$ are by convention and to make the integral invariant, respectively. Thus variations in the matter Lagrangian become

$$\delta S_{\text{matter}} = \int \left(\frac{\partial(\Lambda\sqrt{-g})}{\partial g_{\mu\nu}} - \partial_\sigma \left(\frac{\partial(\Lambda\sqrt{-g})}{\partial(\partial_\sigma g_{\mu\nu})} \right) \right) \delta g_{\mu\nu} d^4x = \frac{1}{2}\int \sqrt{-g}T^{\mu\nu}\delta g_{\mu\nu} d^4x$$

Now, generally, at this point we'd say we want variations to vanish since the physics should be independent of choice of coordinates and thus the coefficients of the variations should vanish, since the variations are arbitrary. But in this case, that's not true! This only works if the variations are truly arbitrary; $\delta g_{\mu\nu}$ has ten components, but not all of them are independent due to the nature of coordinate transformations, of which there are only four. Thus we can't just say $T^{\mu\nu} = 0$ (and if we could, that would be a bad conclusion). Now substitute $\delta g_{\mu\nu}$

into the variations to give

$$\begin{aligned}
\delta S_{\text{matter}} &= \frac{1}{2}\int \sqrt{-g}T^{\mu\nu}\delta g_{\mu\nu}d^4x \\
&= \frac{1}{2}\int \sqrt{-g}T^{\mu\nu}\left(-2\nabla_{(\mu}\xi_{\nu)}\right)d^4x \\
&= -\int T^{\mu\nu}\nabla_\mu\xi_\nu\sqrt{-g}d^4x \\
&= -\int \left(\nabla_\mu(T^{\mu\nu}\xi_\nu) - (\nabla_\mu T^{\mu\nu})\xi_\nu\right)\sqrt{-g}d^4x \\
&= -\int \nabla_\mu(T^{\mu\nu}\xi_\nu)\sqrt{-g}d^4x + \int(\nabla_\mu T^{\mu\nu})\xi_\nu\sqrt{-g}d^4x
\end{aligned}$$

The vector identity

$$\nabla_\mu A^\mu = \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}A^\mu)$$

allows us to rewrite the first integral as

$$\int \nabla_\mu(T^{\mu\nu}\xi_\nu)\sqrt{-g}d^4x = \int \partial_\mu(\sqrt{-g}T^{\mu\nu}\xi_\nu)d^4x$$

The integral above is a total divergence and may be performed over the three hypersurfaces (the boundary of 4-space) where the variations ξ_ν vanish; thus the integral vanishes identically. Therefore the matter Lagrangian variation can be written as

$$\delta S_{\text{matter}} = \int(\nabla_\mu T^{\mu\nu})\xi_\nu\sqrt{-g}d^4x = 0$$

Since the variations are truly arbitrary, then the above integral vanishes only when the term in parenthesis vanishes, or when

$$\nabla_\mu T^{\mu\nu} = 0$$

which is energy-momentum conservation, as desired. Thus diffeomorphism invariance, like gauge invariance, implies a conservation law, namely energy-momentum conservation. In summary, the full action for general relativity is

$$S = \frac{1}{16\pi G_N}\int R\sqrt{-g}d^4x + \int \Lambda\sqrt{-g}d^4x$$

with equations of motion

$$R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R = 8\pi G_N T^{\mu\nu}$$

where

$$T^{\mu\nu} = \frac{2}{\sqrt{-g}}\left[\frac{\partial(\Lambda\sqrt{-g})}{\partial g_{\mu\nu}} - \partial_\sigma\left(\frac{\partial(\Lambda\sqrt{-g})}{\partial(\partial_\sigma g_{\mu\nu})}\right)\right]$$

and energy-momentum conservation $\nabla_\mu T^{\mu\nu} = 0$ is encoded in the EOMs. Note that, unlike electrodynamics, $T^{\mu\nu}$ shows explicit dependence on $g^{\mu\nu}$. This manifests the fact that the gravitational field, unlike the photon field, is gravitationally "charged" and thus interacting, leading to the non-linearity of the theory.

8.3 Palatini formalism*

8.4 Tetradic palatini action*

Part III
Applications

9 Linearized gravity

9.1 Linearized EFEs

The EFEs are of course quite difficult to solve in full generality. We will explore some exact solutions in the proceeding sections. However, before we do that, there is some insight to be gained from simplifying the theory perturbatively. Employing the Newtonian limit from section 7.2, where the metric tensor may be expanded as a perturbation atop a Minkowskian metric, the Christoffel symbols become

$$\begin{aligned}
\Gamma^{\rho}_{\mu\nu} &= \frac{1}{2} g^{\rho\sigma}\left(-\partial_\sigma g_{\mu\nu} + \partial_\nu g_{\sigma\mu} + \partial_\mu g_{\nu\sigma}\right) \\
&= \frac{1}{2}\left(\eta^{\rho\sigma} - h^{\rho\sigma}\right)\left(-\partial_\sigma\left(\eta_{\mu\nu} + h_{\mu\nu}\right) + \partial_\nu\left(\eta_{\sigma\mu} + h_{\sigma\mu}\right) + \partial_\mu\left(\eta_{\nu\sigma} + h_{\nu\sigma}\right)\right) \\
&= \frac{1}{2}\left(\eta^{\rho\sigma} - h^{\rho\sigma}\right)\left(-\partial_\sigma h_{\mu\nu} + \partial_\nu h_{\sigma\mu} + \partial_\mu h_{\nu\sigma}\right) \\
&= \frac{1}{2}\eta^{\rho\sigma}\left(-\partial_\sigma h_{\mu\nu} + \partial_\nu h_{\sigma\mu} + \partial_\mu h_{\nu\sigma}\right) + O(h^2)
\end{aligned}$$

The Riemann tensor may also be written down. However, the expression for the Riemann tensor contains products of Christoffel symbols; since these are linear in h, the products, which are quadratic in h, vanish at linear order in h. Thus the Riemann tensor becomes

$$\begin{aligned}
R^{\alpha}{}_{\beta\mu\nu} &= \partial_\mu \Gamma^{\alpha}_{\beta\nu} - \partial_\nu \Gamma^{\alpha}_{\mu\beta} + \Gamma^{\alpha}_{\mu\lambda}\Gamma^{\lambda}_{\beta\nu} - \Gamma^{\alpha}_{\nu\lambda}\Gamma^{\lambda}_{\beta\mu} \\
&= \frac{1}{2}\partial_\mu\left(\eta^{\alpha\sigma}\left(-\partial_\sigma h_{\beta\nu} + \partial_\nu h_{\sigma\beta} + \partial_\beta h_{\nu\sigma}\right)\right) - \frac{1}{2}\partial_\nu\left(\eta^{\alpha\sigma}\left(-\partial_\sigma h_{\mu\beta} + \partial_\mu h_{\sigma\beta} + \partial_\beta h_{\mu\sigma}\right)\right) + O(h^2) \\
&= \frac{1}{2}\eta^{\alpha\sigma}\left(-\partial_\mu\partial_\sigma h_{\beta\nu} + \partial_\mu\partial_\nu h_{\sigma\beta} + \partial_\mu\partial_\beta h_{\nu\sigma} + \partial_\nu\partial_\sigma h_{\mu\beta} - \partial_\nu\partial_\mu h_{\sigma\beta} - \partial_\nu\partial_\beta h_{\mu\sigma}\right) \\
&= \frac{1}{2}\eta^{\alpha\sigma}\left(\partial_\mu\partial_\beta h_{\nu\sigma} - \partial_\nu\partial_\beta h_{\mu\sigma} - \partial_\mu\partial_\sigma h_{\beta\nu} + \partial_\nu\partial_\sigma h_{\mu\beta}\right)
\end{aligned}$$

Of course, the Ricci tensor can be easily constructed by taking the appropriate trace:

$$\begin{aligned}
R_{\beta\nu} = R^{\alpha}{}_{\beta\alpha\nu} &= \frac{1}{2}\eta^{\alpha\sigma}\left(\partial_\alpha\partial_\beta h_{\nu\sigma} - \partial_\nu\partial_\beta h_{\alpha\sigma} - \partial_\alpha\partial_\sigma h_{\beta\nu} + \partial_\nu\partial_\sigma h_{\alpha\beta}\right) \\
&= \frac{1}{2}\left(\partial_\beta\partial^\sigma h_{\nu\sigma} - \partial_\nu\partial_\beta h - \Box h_{\beta\nu} + \partial_\nu\partial^\alpha h_{\alpha\beta}\right)
\end{aligned}$$

and lastly the Ricci scalar:

$$\begin{aligned}
R = g^{\beta\nu}R_{\beta\nu} &= \eta^{\beta\nu}R_{\beta\nu} + O(h^2) \\
&= \frac{1}{2}\eta^{\beta\nu}\left(\partial_\beta\partial^\sigma h_{\nu\sigma} - \partial_\nu\partial_\beta h - \Box h_{\beta\nu} + \partial_\nu\partial^\alpha h_{\alpha\beta}\right) \\
&= \frac{1}{2}\left(\partial^\nu\partial^\sigma h_{\nu\sigma} - \Box h - \Box h + \partial^\beta\partial^\alpha h_{\alpha\beta}\right) \\
&= \partial^\alpha\partial^\beta h_{\alpha\beta} - \Box h
\end{aligned}$$

Therefore the Einstein tensor is

$$\begin{aligned}
G_{\mu\nu} &= R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \\
&= R_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} R + O(h^2) \\
&= \frac{1}{2} \left(\partial_\mu \partial^\sigma h_{\nu\sigma} - \partial_\mu \partial_\nu h - \Box h_{\mu\nu} + \partial_\nu \partial^\sigma h_{\sigma\mu} \right) - \frac{1}{2} \eta_{\mu\nu} \left(\partial^\alpha \partial^\beta h_{\alpha\beta} - \Box h \right)
\end{aligned}$$

Thus the EFEs in linearized gravity are

$$\partial_\mu \partial^\sigma h_{\nu\sigma} + \partial_\nu \partial^\sigma h_{\sigma\mu} - \Box h_{\mu\nu} - \eta_{\mu\nu} \partial^\alpha \partial^\beta h_{\alpha\beta} - \partial_\mu \partial_\nu h + \eta_{\mu\nu} \Box h = 16\pi G_N T_{\mu\nu}$$

This looks like a mess! Even something as simple as solving for the metric in the vacuum looks intractable. Fortunately, there are a myriad simplifications we can make. It turns out that our perturbative decomposition is not unique. Suppose we switch from Cartesian to polar coordinates. Now the Minkowski metric $\eta_{\mu\nu}$ is no longer in its canonical form and thus the perturbation matrix must change accordingly; but the dynamics won't change at all (nor should they!). The dynamics are invariant under a set of reparameterizations, and the actual degrees of freedom are dramatically restricted by this redundancy. This, of course, is a manifestation of gauge invariance.

9.2 Gauge transformations

Suppose we treat the actual physical spacetime as a manifold M_p. A metric $g_{\alpha\beta}$ obeying the EFEs lives on this manifold. Now suppose further that there's a spacetime where the metric is Minkowski, $\eta_{\mu\nu}$; let's call this the background manifold, M_b. Lastly, suppose there is a diffeomorphism ϕ from the background to the physical manifold, $\phi : M_b \to M_p$. It is possible to pull the physical metric back to M_b. Thus $\eta_{\mu\nu}$ and $(\phi^* g)_{\mu\nu}$ both live in the same manifold and may be compared directly

$$h_{\mu\nu} = (\phi^* g)_{\mu\nu} - \eta_{\mu\nu}$$

For some ϕ, the above equation will yield sufficiently small components of $h_{\mu\nu}$, $|h_{\mu\nu}| \ll 1$. Now consider a vector field $\xi^\mu(x)$ living on M_b; this vector field generates a one-parameter family of diffeomorphisms, $\psi_\epsilon : M_b \to M_b$ (see section 1.6). Think, for example, of a simple coordinate redefinition given by

$$x'^\mu = x^\mu + \epsilon \xi^\mu$$

Diffeomorphisms may of course be concatenated and thus $(\phi \circ \psi_\epsilon)^*$ is also a diffeomorphism, in this case one from M_p to M_b. This allows us to define a family of $h_{\mu\nu}$, parameterized by ϵ:

$$\begin{aligned}
h^{(\epsilon)}_{\mu\nu} &= [(\phi \circ \psi_\epsilon)^* g]_{\mu\nu} - \eta_{\mu\nu} \\
&= [\psi^*_\epsilon (\phi^* g)]_{\mu\nu} - \eta_{\mu\nu} \\
&= [\psi^*_\epsilon (h + \eta)]_{\mu\nu} - \eta_{\mu\nu} \\
&= \psi^*_\epsilon (h_{\mu\nu}) + \psi^*_\epsilon (\eta_{\mu\nu}) - \eta_{\mu\nu}
\end{aligned}$$

Now let's assume that ϵ is small, such that $\psi_\epsilon^*(h_{\mu\nu}) \approx h_{\mu\nu} + O(h^2)$. Thus $h_{\mu\nu}^{(\epsilon)}$ may be written as

$$\begin{aligned} h_{\mu\nu}^{(\epsilon)} &= \psi_\epsilon^*(h_{\mu\nu}) + \psi_\epsilon^*(\eta_{\mu\nu}) - \eta_{\mu\nu} \\ &= h_{\mu\nu} + \epsilon \left(\frac{\psi_\epsilon^*(\eta_{\mu\nu}) - \eta_{\mu\nu}}{\epsilon} \right) \\ &\to h_{\mu\nu} + \epsilon \mathcal{L}_\xi \eta_{\mu\nu} \end{aligned}$$

As given in Section 1.6, the Lie derivative of a metric tensor may be written as

$$\mathcal{L}_\xi \eta_{\mu\nu} = 2\nabla_{(\mu}\xi_{\nu)} \to 2\partial_{(\mu}\xi_{\nu)}$$

Note that in the last step above, the covariant derivative was replaced with partial derivative since the metric is flat. Thus the perturbation matrix as a function of ϵ is

$$h_{\mu\nu}^{(\epsilon)} = h_{\mu\nu} + 2\epsilon\partial_{(\mu}\xi_{\nu)}$$

These are the gauge transformations of the perturbation matrix. They are analogous to the gauge transformations of classical electrodynamics where $A_\mu \to A_\mu + \partial_\mu\alpha$, which leaves the Maxwell equations unchanged; certain judicious choices of α make solving the Maxwell equations significantly simpler. This concept can be lifted to linearized gravity, in that certain gauge choices will make solving the linearized EFEs much simpler.

9.3 Degrees of freedom

Recall that the gauge transformations may be used to isolate the propagating degrees of freedom by choosing a gauge such that the redundancies vanish. A propagating degree of freedom is one with non-zero time-derivative. To that end, we should explore the $h_{\mu\nu}$ tensor to see what its actual physical degrees of freedom are. This follows in the same vein as identifying the entries of the EM field tensor as $F^{0i} = E^i$ and $B^i = \epsilon^{ijk}F_{jk}$. Of course, this strict identification is one Lorentz transformation away from being incorrect and the same will hold for our analysis of $h_{\mu\nu}$, but that should not stop us from proceeding with the knowledge that the elements of the tensor may be mixed up (or brought into the appropriate form) by some arbitrary transformation.

First, $h_{\mu\nu}$ is a symmetric tensor. This means that although it has 16 entries (in $(3+1)D$), only 10 of these are truly independent. The tensor may be decomposed into h_{00} (1 degree), $h_{0i} = h_{i0}$ (3 degrees) and a symmetric spatial subtensor h_{ij} (6 degrees):

$$h_{\mu\nu} = h_{00} + h_{0i} + h_{i0} + h_{ij}$$

The h_{00} entry is a singlet under rotations and thus may be identified with a scalar potential $h_{00} = -2\Phi$. The h_{0i} are a vector under rotations and thus may be identified with a vector potential $h_{0i} = w_i$. The spatial subtensor may be decomposed into a trace Ψ times an identity matrix and a trace-free matrix s_{ij} (known as the strain)

$$h_{ij} = 2s_{ij} - 2\Psi\delta_{ij}$$

Occasionally, this decomposition isn't quite so useful, so it may be dropped. In summary

$$h_{\mu\nu} = \left(\begin{array}{c|c} -2\Phi & w_i \\ \hline w_i & h_{ij} \end{array}\right)$$

Let's use this decomposition to analyze the EFEs. To do so, let's construct the Einstein tensor; this requires the Ricci tensor, and therefore the Riemann tensor. The entries of the Riemann tensor are

$$\begin{aligned}
R_{0\mu0\nu} &= \frac{1}{2}\left(\partial_0\partial_\mu h_{\nu0} - \partial_\nu\partial_\mu h_{00} - \partial_0\partial_0 h_{\mu\nu} + \partial_\nu\partial_0 h_{0\mu}\right) \\
R_{0j0l} &= \frac{1}{2}\left(\partial_0\partial_j h_{l0} - \partial_l\partial_j h_{00} - \partial_0\partial_0 h_{jl} + \partial_l\partial_0 h_{0j}\right) \\
&= \frac{1}{2}\left(\partial_0\partial_j w_l - \partial_l\partial_j(-2\Phi) - \partial_0\partial_0 h_{jl} + \partial_l\partial_0 w_j\right) \\
&= \partial_j\partial_l\Phi + \partial_0\partial_{(j}w_{l)} - \frac{1}{2}\partial_0\partial_0 h_{jl}
\end{aligned}$$

$$\begin{aligned}
R_{0jkl} &= \frac{1}{2}\left(\partial_k\partial_j h_{l0} - \partial_l\partial_j h_{k0} - \partial_k\partial_0 h_{jl} + \partial_l\partial_0 h_{kj}\right) \\
&= \frac{1}{2}\left(\partial_j\left(\partial_k w_l - \partial_l w_k\right) - \partial_0\left(\partial_k h_{jl} - \partial_l h_{jk}\right)\right) \\
&= \partial_j\partial_{[k}w_{l]} - \partial_0\partial_{[k}h_{l]j}
\end{aligned}$$

$$\begin{aligned}
R_{ijkl} &= \frac{1}{2}\left(\partial_k\partial_j h_{li} - \partial_l\partial_j h_{ki} - \partial_k\partial_i h_{jl} + \partial_l\partial_i h_{kj}\right) \\
&= \frac{1}{2}\left(\partial_j\left(\partial_k h_{li} - \partial_l h_{ki}\right) - \partial_i\left(\partial_k h_{jl} - \partial_l h_{kj}\right)\right) \\
&= \partial_j\partial_{[k}h_{l]i} - \partial_i\partial_{[k}h_{l]j}
\end{aligned}$$

Now let's construct the Ricci tensor

$$R_{\mu\nu} = R^\rho{}_{\mu\rho\nu} = R^0{}_{\mu0\nu} + R^i{}_{\mu i\nu}$$

$$\begin{aligned}
R_{00} &= R^0{}_{000} + R^i{}_{0i0} \\
&= \eta^{ij}R_{j0i0} \\
&= \eta^{ij}\left(\partial_j\partial_i\Phi + \partial_0\partial_{(j}w_{i)} - \frac{1}{2}\partial_0\partial_0 h_{ji}\right) \\
&= \nabla^2\Phi + \partial_0\partial_i w^i - \frac{1}{2}\partial_0\partial_0(-6\Psi) \\
&= \nabla^2\Phi + \partial_0\partial_i w^i + 3\partial_0\partial_0\Psi
\end{aligned}$$

$$
\begin{aligned}
R_{0j} &= R^0{}_{00j} + R^i{}_{0ij} \\
&= \eta^{0i} R_{i00j} + \eta^{ki} R_{k0ij} \\
&= -\eta^{ki} R_{0kij} \\
&= -\eta^{ki} \left(\partial_k \partial_{[i} w_{j]} - \partial_0 \partial_{[i} h_{j]k}\right) \\
&= -\frac{1}{2}\eta^{ki} \left(\partial_k \left(\partial_i w_j - \partial_j w_i\right) - \partial_0 \left(\partial_i h_{jk} - \partial_j h_{ik}\right)\right) \\
&= -\frac{1}{2} \left(\nabla^2 w_j - \partial_j \partial_i w_i - \partial_0 \left(\partial_i h_{ij} - \partial_j h\right)\right) \\
&= -\frac{1}{2} \left(\nabla^2 w_j - \partial_j \partial_i w_i - \partial_0 \left(\partial_i \left(2 s_{ij} - 2\Psi \delta_{ij}\right) - \partial_j (-6\Psi)\right)\right) \\
&= -\frac{1}{2} \left(\nabla^2 w_j - \partial_j \partial_i w_i - \partial_0 \left(2\partial_i s_{ij} - 2\partial_j \Psi + 6\partial_j \Psi\right)\right) \\
&= -\frac{1}{2}\nabla^2 w_j + \frac{1}{2}\partial_j \partial_i w_i + \partial_0 \partial_i s_{ij} + 2\partial_0 \partial_j \Psi
\end{aligned}
$$

$$
\begin{aligned}
R_{jk} &= R^0{}_{j0k} + R^i{}_{jik} \\
&= -\eta^{0i} R_{0jik} - \eta^{li} R_{ljik} \\
&= -\eta^{0i} R_{0jik} - \eta^{li} R_{ikjl} \\
&= -\eta^{0i} \left(\partial_j \partial_{[i} w_{k]} - \partial_0 \partial_{[i} h_{k]j}\right) - \eta^{li} \left(\partial_k \partial_{[j} h_{l]i} - \partial_i \partial_{[j} h_{l]k}\right) \\
&= -\frac{1}{2}\eta^{0i} \left(\partial_j \left(\partial_i w_k - \partial_k w_i\right) - \partial_0 \left(\partial_i h_{kj} - \partial_k h_{ij}\right)\right) - \frac{1}{2}\eta^{li} \left(\partial_k \left(\partial_j h_{li} - \partial_l h_{ji}\right) - \partial_i \left(\partial_j h_{kl} - \partial_l h_{jk}\right)\right) \\
&= -\frac{1}{2} \left(\partial_j \left(\partial_0 w_k - \partial_k w_0\right) - \partial_0 \left(\partial_0 h_{kj} - \partial_k h_{0j}\right)\right) - \frac{1}{2} \left(\partial_k \left(\partial_j h - \partial_i h_{ij}\right) - \partial_i \left(\partial_j h_{ki} - \partial_i h_{jk}\right)\right) \\
&= -\frac{1}{2} \left(\partial_j \partial_0 w_k - \partial_j \partial_k (-2\Phi) - \partial_0 \partial_0 h_{kj} + \partial_k \partial_0 w_j + \partial_k \partial_j h - \partial_i \partial_k h_{ij} - \partial_i \partial_j h_{ki} + \nabla^2 h_{jk}\right) \\
&= -\frac{1}{2} \left(\partial_0 \left(\partial_j w_k + \partial_k w_j\right) + 2\partial_j \partial_k \Phi + \partial_k \partial_j (-6\Psi) - \partial_i \partial_k h_{ij} - \partial_i \partial_j h_{ki} + \left(-\partial_0^2 + \nabla^2\right) h_{jk}\right) \\
&= -\frac{1}{2} \left(2\partial_0 \partial_{(j} w_{k)} + 2\partial_j \partial_k \Phi - 6\partial_k \partial_j \Psi - \partial_i \partial_k (2 s_{ij} - 2\Psi \delta_{ij}) - \partial_i \partial_j (2 s_{ik} - 2\Psi \delta_{ik}) + \Box h_{jk}\right) \\
&= -\frac{1}{2} \left(2\partial_0 \partial_{(j} w_{k)} + 2\partial_j \partial_k \Phi - 6\partial_k \partial_j \Psi - 2\partial_i \partial_k s_{ij} + 2\partial_j \partial_k \Psi - 2\partial_i \partial_j s_{ik} + 2\partial_k \partial_j \Psi + \Box h_{jk}\right) \\
&= -\frac{1}{2} \left(2\partial_0 \partial_{(j} w_{k)} + 2\partial_j \partial_k \Phi - 2\partial_k \partial_j \Psi - 2\partial_i \left(\partial_k s_{ij} + \partial_i \partial_j s_{ik}\right) + \Box h_{jk}\right) \\
&= -\frac{1}{2} \left(2\partial_0 \partial_{(j} w_{k)} + 2\partial_j \partial_k (\Phi - \Psi) - 4\partial_i \partial_{(j} s_{k)i} + \Box (2 s_{jk} - 2\Psi \delta_{jk})\right) \\
&= -\partial_j \partial_k (\Phi - \Psi) - \partial_0 \partial_{(j} w_{k)} - \Box s_{jk} + \Box \Psi \delta_{jk} + 2\partial_i \partial_{(j} s_{k)i}
\end{aligned}
$$

Whew! That was quite an exercise. But we're not done yet; we still need the Ricci scalar:

$$
\begin{aligned}
R = R^{\mu}{}_{\nu} &= \eta^{\mu\nu} R_{\mu\nu} \\
&= \eta^{00} R_{00} + \eta^{0j} R_{0j} + \eta^{i0} R_{i0} + \eta^{ij} R_{ij} \\
&= \eta^{00} R_{00} + 2\eta^{0j} R_{0j} + \eta^{ij} R_{ij} \\
&= \left(\nabla^2 \Phi + \partial_0 \partial_i w^i + 3\partial_0^2 \Psi\right) + 2\eta^{0j} \left(-\frac{1}{2}\nabla^2 w_j + \frac{1}{2}\partial_j \partial_i w_i + \partial_0 \partial_i s_{ij} + 2\partial_0 \partial_j \Psi\right) \\
&\qquad - \eta^{ij} \left(-\partial_i \partial_j (\Phi - \Psi) - \partial_0 \partial_{(i} w_{j)} - \Box s_{ij} + \Box \Psi \delta_{ij} + 2\partial_k \partial_{(i} s_{j)k}\right)
\end{aligned}
$$

Now we can construct the Einstein tensor

$$
\begin{aligned}
G_{00} &= 2\nabla^2 \Psi + \partial_i \partial_j s^{ij} \\
G_{0j} &= -\frac{1}{2}\nabla^2 w_j + \frac{1}{2}\partial_j \partial_k w^k + 2\partial_0 \partial_j \Psi + \partial_0 \partial_k s_j{}^k \\
G_{ij} &= (\delta_{ij}\nabla^2 - \partial_i \partial_j)(\Phi - \Psi) + \delta_{ij}\partial_0 \partial_k w^k - \partial_0 \partial_{(i} w_{j)} + 2\delta_{ij}\partial_0^2 \Psi - \Box s_{ij} + 2\partial_k \partial_{(i} s_{j)}{}^k - \delta_{ij}\partial_k \partial_l s^{kl}
\end{aligned}
$$

Now let's see what the EFEs give us. The 00 equation gives

$$
\begin{aligned}
G_{00} &= 8\pi G_N T_{00} \\
2\nabla^2 \Psi + \partial_i \partial_j s^{ij} &= 8\pi G_N T_{00} \\
\nabla^2 \Psi &= 4\pi G_N T_{00} - \frac{1}{2}\partial_i \partial_j s^{ij}
\end{aligned}
$$

which is a constraint on Ψ in terms of T_{00} and s_{ij}. The $\{0j\}$ equations give

$$
\begin{aligned}
G_{0j} &= 8\pi G_N T_{0j} \\
-\frac{1}{2}\nabla^2 w_j + \frac{1}{2}\partial_j \partial_k w^k + 2\partial_0 \partial_j \Psi + \partial_0 \partial_k s_j{}^k &= 8\pi G_N T_{0j} \\
(\delta_{jk}\nabla^2 - \partial_j \partial_k) w^k &= -16\pi G_N T_{0j} + 4\partial_0 \partial_j \Psi + 2\partial_0 \partial_k s_j{}^k
\end{aligned}
$$

which is a constraint on the w_i in terms of T_{0j}, Ψ and $s_j{}^k$. Lastly, the $\{ij\}$ equations give

$$
\begin{aligned}
8\pi G_N T_{ij} &= G_{ij} \\
8\pi G_N T_{ij} &= (\delta_{ij}\nabla^2 - \partial_i \partial_j)(\Phi - \Psi) + \delta_{ij}\partial_0 \partial_k w^k - \partial_0 \partial_{(i} w_{j)} + 2\delta_{ij}\partial_0^2 \Psi \\
&\qquad - \Box s_{ij} + 2\partial_k \partial_{(i} s_{j)}{}^k - \delta_{ij}\partial_k \partial_l s^{kl}
\end{aligned}
$$

$$
\begin{aligned}
(\delta_{ij}\nabla^2 - \partial_i \partial_j)\Phi &= -8\pi G_N T_{ij} - (\delta_{ij}\nabla^2 - \partial_i \partial_j)\Psi + \delta_{ij}\partial_0 \partial_k w^k - \partial_0 \partial_{(i} w_{j)} + 2\delta_{ij}\partial_0^2 \Psi \\
&\qquad - \Box s_{ij} + 2\partial_k \partial_{(i} s_{j)}{}^k - \delta_{ij}\partial_k \partial_l s^{kl}
\end{aligned}
$$

which is a constraint on Φ in terms of T_{ij}, Ψ, w_k and s_{ij}. Note that in the above equations, Ψ, Φ and w_i do not have time-derivatives acting on them. This means that their values are fixed by boundary conditions and the s_{ij}; they do not represent propagating degrees of freedom. Thus, the strain, s_{ij} represents the only propagating degrees of freedom in the linearized EFEs.

There is one other way the decomposition of the perturbation matrix may be viewed. Recall the fundamental theorem of vector calculus that a vector V^i (which represents three degrees of freedom) may be decomposed into a divergenceless (transverse) vector plus a curl-free (longitudinal) vector

$$V^i = D^i + C^i$$

where D^i and C^i satisfy the relations

$$\partial_i D^i = 0 \qquad \text{and} \qquad \epsilon_{ijk}\partial^j C^k = 0$$

These expressions imply the following

$$D^i = \epsilon^{ijk}\partial_j \xi_k \qquad \text{and} \qquad C^i = \partial^i \lambda$$

Thus the vector V^i may actually be specified by some scalar λ and a vector ξ^i. But wait, doesn't that represent *four* degrees of freedom? No, because the vector ξ^i is non-unique and may be redefined by $\xi_i \to \xi_i + \partial_i \omega$. We may proceed similarly for a symmetric, traceless 3×3 matrix s_{ij}, which represents 5 degrees of freedom (6 normally, but the fixed trace eliminates one). Such a tensor may be decomposed into transverse, solenoidal and longitudinal tensors:

$$s^{ij} = s^{ij}_T + s^{ij}_S + s^{ij}_L$$

The divergence of s_T vanishes (hence it's transverse) while the divergence of s_S yields a vector that is itself divergenceless and similarly, the divergence of s_L yields a vector that is curl-free:

$$\partial_j\left(\partial_i s^{ij}_S\right) = \partial_j D^j = 0 \qquad \text{and} \qquad \epsilon_{klj}\partial^l\left(\partial_i s^{ij}_L\right) = \epsilon_{klj}\partial^l C^j = 0$$

As before, this means that s_S may be derived from a (non-unique) transverse vector ζ^i and that s_L may derived from some scalar θ:

$$s^{ij}_S = \partial^{(i}\zeta^{j)} \qquad \text{and} \qquad s^{ij}_L = \left(\partial^i\partial^j - \frac{1}{3}\delta^{ij}\nabla^2\right)\theta$$

Thus the longitudinal part corresponds to one degree of freedom, θ; the solenoidal part to two, ζ^i and the transverse part to two as well, for a total of five.

We can put these results together to fully characterize $h_{\mu\nu}$, which is a symmetric rank-2 tensor and thus has 10 d.o.f. The scalar Φ represents one degree. The vector $h_{i0} = w_i$ may be decomposed as described above and characterized by λ (one degree) and ξ^i (two degrees). The spatial submatrix h_{ij} may be decomposed into its trace Ψ (one degree) and the strain s_{ij}. The strain is a symmetric, traceless 3×3 matrix. It thus may be further decomposed as described above and characterized by θ (one degree), ζ^i (two degrees) and s_T (two degrees), giving the total of ten degrees of freedom, as expected. In summary, $h_{\mu\nu}$ is completely characterized by the scalar degrees of freedom $(\Phi, \Psi, \lambda, \theta)$, the vector degrees of freedom (ξ^i, ζ^i) and the symmetric, traceless, transverse matrix s^{ij}_T. These may be referred to as the scalar, vector and tensor modes, respectively.

9.4 Linearized gravity gauges

We can combine the results of the two previous sections to greatly simplify evaluating the EFEs. Let's recall the definition of gauge transformations from section 9.2

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

and let's see the effects of these gauge transformations on the various components of the decomposition from section 9.3:

$$\begin{aligned} h_{00} &\to h_{00} + \partial_0 \xi_0 + \partial_0 \xi_0 \\ -2\Phi &\to -2\Phi + 2\partial_0 \xi_0 \\ \Phi &\to \Phi + \partial_0 \xi^0 \end{aligned}$$

$$\begin{aligned} h_{0j} &\to h_{0j} + \partial_0 \xi_j + \partial_j \xi_0 \\ w_j &\to w_j + \partial_0 \xi^j - \partial_j \xi^0 \end{aligned}$$

$$\begin{aligned} h_{ij} &\to h_{ij} + \partial_i \xi_j + \partial_j \xi_i \\ \delta^{ij}(2s_{ij} - 2\Psi\delta_{ij}) &\to \delta^{ij}(2s_{ij} - 2\Psi\delta_{ij}) + \delta^{ij}\partial_i \xi_j + \delta^{ij}\partial_j \xi_i \\ -6\Psi &\to -6\Psi + 2\partial_i \xi^i \\ \Psi &\to \Psi - \frac{1}{3}\partial_i \xi^i \end{aligned}$$

$$\begin{aligned} s_{ij} &= \frac{1}{2}h_{ij} + \Psi\delta_{ij} \\ &\to \frac{1}{2}\left(h_{ij} + \partial_i \xi_j + \partial_j \xi_i\right) + \Psi - \frac{1}{3}\partial_k \xi^k \delta_{ij} \\ &\to \frac{1}{2}h_{ij} + \Psi + \frac{1}{2}\left(\partial_i \xi_j + \partial_j \xi_i\right) - \frac{1}{3}\partial_k \xi^k \delta_{ij} \\ &\to s_{ij} + \partial_{(i}\xi_{j)} - \frac{1}{3}\partial_k \xi^k \delta_{ij} \end{aligned}$$

In summary, the components behave as

$$\begin{aligned} \Phi &\to \Phi + \partial_0 \xi^0 \\ \Psi &\to \Psi - \frac{1}{3}\partial_i \xi^i \\ w_i &\to w_i + \partial_0 \xi^i - \partial_i \xi^0 \\ s_{ij} &\to s_{ij} + \partial_{(i}\xi_{j)} - \frac{1}{3}\partial_k \xi^k \delta_{ij} \end{aligned}$$

9.4.1 Transverse gauge

Now, on to the gauge choices. The first possible choice is known as the *transverse gauge*. This gauge is a generalization of the Coulomb gauge from EM, $\partial_i A^i = 0$. In this gauge, the ξ_i are chosen such that the strain is transverse

$$\partial_i s^{ij} = 0$$

is satisfied. Plugging in the gauge transformations gives

$$\begin{aligned}
0 &= \partial_i s^{ij} \\
&\to \partial_i \left(s^{ij} + \frac{1}{2}(\partial^i \xi^j + \partial^j \xi^i) - \frac{1}{3}\partial_k \xi^k \delta^{ij} \right) \\
&= \partial_i s^{ij} + \frac{1}{2}\nabla^2 \xi^j + \frac{1}{2}\partial_i \partial^j \xi^i - \frac{1}{3}\partial^j \partial^k \xi_k \\
&= \partial_i s^{ij} + \frac{1}{2}\nabla^2 \xi^j + \frac{1}{2}\partial^i \partial^j \xi_i - \frac{1}{3}\partial^j \partial^k \xi_k \\
&= \partial_i s^{ij} + \frac{1}{2}\nabla^2 \xi^j + \frac{1}{6}\partial^j \partial^k \xi_k \\
-2\partial_i s^{ij} &= \nabla^2 \xi^j + \frac{1}{3}\partial^j \partial^k \xi_k
\end{aligned}$$

which is a constraint the ξ^i must satisfy. However, this leaves the ξ^0 component unconstrained. This is chosen so that the w_i are transverse as well

$$\partial_i w^i = 0$$

Together, these constitute the transverse gauge condition. Using these, the linearized EFEs simplify greatly. The $\{00\}$ equation becomes

$$\nabla^2 \Psi = 4\pi G_N T_{00} - \frac{1}{2}\partial_i \partial_j s^{ij} = 4\pi G_N T_{00}$$

The $\{0j\}$ equations become

$$\begin{aligned}
(\delta_{jk}\nabla^2 - \partial_j \partial_k) w^k &= -16\pi G_N T_{0j} + 4\partial_0 \partial_j \Psi + 2\partial_0 \partial_k {s_j}^k \\
\nabla^2 w_j &= -16\pi G_N T_{0j} + 4\partial_0 \partial_j \Psi
\end{aligned}$$

Lastly, the $\{ij\}$ equations become

$$\begin{aligned}
(\delta_{ij}\nabla^2 - \partial_i \partial_j)\Phi = -8\pi G_N T_{ij} - (\delta_{ij}\nabla^2 - \partial_i \partial_j)\Psi + \delta_{ij}\partial_0 \partial_k w^k - \partial_0 \partial_{(i} w_{j)} + 2\delta_{ij}\partial_0^2 \Psi \\
- \Box s_{ij} + 2\partial_k \partial_{(i} {s_{j)}}^k - \delta_{ij}\partial_k \partial_l s^{kl}
\end{aligned}$$

$$= -8\pi G_N T_{ij} - (\delta_{ij}\nabla^2 - \partial_i \partial_j)\Psi - \partial_0 \partial_{(i} w_{j)} + 2\delta_{ij}\partial_0^2 \Psi - \Box s_{ij}$$

Remember that in electrodynamics, certain gauge choices render some of the Maxwell equations almost trivial to solve while leaving the other horrendously intractable. As we can see, the same result occurs here with two of the equations becoming relatively simple while the third is horrible.

9.4.2 Synchronous gauge

Another choice is the *synchronous gauge*, which is a generalization of the Weyl (or temporal gauge) from electrodynamics, $\phi = 0$. In this gauge, the ξ^i are chosen so that $\Phi = 0$. Plugging in the gauge transformations gives

$$0 \to \Phi + \partial_0 \xi^0 \qquad \Rightarrow \qquad \partial_0 \xi^0 = -\Phi$$

Thus the ξ^0 is constrained. This of course leaves the ξ^i free to choose as we please. We'll choose it so that $w_i = 0$, which gives the gauge transformations

$$0 = w_i + \partial_0 \xi_i - \partial_i \xi_0 \qquad \Rightarrow \qquad \partial_0 \xi_i = -w_i + \partial_i \xi_0$$

This means the perturbation matrix takes the simple form $h_{\mu\nu} = h_{ij}$ and thus the metric is simply

$$ds^2 = -\mathrm{d}^2 t + (\delta_{ij} + h_{ij}) \mathrm{d}x^i \mathrm{d}x^j$$

The EFEs similarly take on a simplified character. The $\{00\}$ equation does not simplify much, but the $0j$ equations do

$$\begin{aligned} (\delta_{jk} \nabla^2 - \partial_j \partial_k) w^k &= -16\pi G_N T_{0j} + 4\partial_0 \partial_j \Psi + 2\partial_0 \partial_k {s_j}^k \\ 0 &= -16\pi G_N T_{0j} + 4\partial_0 \partial_j \Psi + 2\partial_0 \partial_k {s_j}^k \end{aligned}$$

as do the $\{ij\}$ equations

$$\begin{aligned} (\delta_{ij} \nabla^2 - \partial_i \partial_j) \Phi = -8\pi G_N T_{ij} - (\delta_{ij} \nabla^2 - \partial_i \partial_j) \Psi + \delta_{ij} \partial_0 \partial_k w^k - \partial_0 \partial_{(i} w_{j)} + 2\delta_{ij} \partial_0^2 \Psi \\ - \Box s_{ij} + 2\partial_k \partial_{(i} {s_{j)}}^k - \delta_{ij} \partial_k \partial_l s^{kl} \end{aligned}$$

$$0 = -8\pi G_N T_{ij} - (\delta_{ij} \nabla^2 - \partial_i \partial_j) \Psi + 2\delta_{ij} \partial_0^2 \Psi - \Box s_{ij} + 2\partial_k \partial_{(i} {s_{j)}}^k - \delta_{ij} \partial_k \partial_l s^{kl}$$

The last two gauges we'll consider are the *transverse-traceless (TT) gauge* and the *Lorenz/harmonic gauge*. These gauges are particularly useful for studying gravitational radiation, so their discussion will be postponed until the next chapter.

9.5 Photon deflection

We now have the tools to consider a simple application of linearized gravity, namely the angle by which a photon emitted by a far-away source is deflected by a quiescent (time-independent) gravitating body on its way to a far-away observer. Let $x^\mu(\lambda)$ be the actual path, a null geodesic, taken by the photon from the source to the observer. Since we're working in perturbation theory, it is appropriate to consider the actual geodesic as a background geodesic $x_0^\mu(\lambda)$ and a perturbed geodesic $x_1^\mu(\lambda)$. The background geodesic solves the geodesic equation in the background (meaning $g_{\mu\nu} = \eta_{\mu\nu}$ and thus $x_0^\mu(\lambda)$ is just a straight line). Thus we need to find an equation governing $x_1^\mu(\lambda)$. We suspect this will be derived from the geodesic equation. Let's start there and plug in our perturbative expansions, keeping things to first order in the perturbations:

$$\begin{aligned}
0 &= \frac{d^2x^\mu}{d\lambda^2}(x_0^\mu + x_1^\mu) + \Gamma^\mu_{\nu\rho}\frac{dx^\nu}{d\lambda}\frac{dx^\rho}{d\lambda} \\
&= \frac{d^2}{d\lambda^2}(x_0^\mu + x_1^\mu) + \Gamma^\mu_{\nu\rho}\frac{d}{d\lambda}(x_0^\nu + x_1^\nu)\frac{d}{d\lambda}(x_0^\rho + x_1^\rho) \\
&= \frac{d^2x_0^\mu}{d\lambda^2} + \frac{d^2x_1^\mu}{d\lambda^2} + \Gamma^\mu_{\nu\rho}\frac{dx_0^\nu}{d\lambda}\frac{dx_0^\rho}{d\lambda} + \text{H.O.T.} \\
\Rightarrow \frac{d^2x_1^\mu}{d\lambda^2} &= -\Gamma^\mu_{\nu\rho}\frac{dx_0^\nu}{d\lambda}\frac{dx_0^\rho}{d\lambda}
\end{aligned}$$

where we've recognized that $\frac{d^2x_0^\mu}{d\lambda^2} = 0$ since the background geodesic is simply a straight line. For clarity, let

$$k^\mu = \frac{dx_0^\mu}{d\lambda} \qquad \text{and} \qquad \ell^\mu = \frac{dx_1^\mu}{d\lambda}$$

Thus the geodesic equation becomes

$$\frac{d\ell^\mu}{d\lambda} = -\Gamma^\mu_{\nu\rho}k^\nu k^\rho$$

Of course, this is not the only condition that $x_1^\mu(\lambda)$ must satisfy; it must also "conspire" with $x_0^\mu(\lambda)$ so that the entire path $x^\mu(\lambda)$ is a null geodesic

$$g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = 0$$

As with the geodesic equation, this may be expanded perturbatively:

$$\begin{aligned}
0 &= g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} \\
&= (\eta_{\mu\nu} + h_{\mu\nu})(k^\mu + \ell^\mu)(k^\nu + \ell^\nu) \\
&= k\cdot k + 2k\cdot\ell + \ell\cdot\ell + h_{\mu\nu}k^\mu k^\nu + \text{H.O.T.} \\
&\approx 2k\cdot\ell + h_{\mu\nu}k^\mu k^\nu
\end{aligned}$$

Thus the two equations governing the perturbation geodesic are

$$\frac{d\ell^\mu}{d\lambda} = -\Gamma^\mu_{\nu\rho}k^\nu k^\rho \qquad \text{and} \qquad 2k\cdot\ell = -h_{\mu\nu}k^\mu k^\nu$$

To proceed any further we explicitly need the metric. Recall the EFEs in the transverse gauge ($\partial_i s^{ij} = 0$, $\partial_i w^i = 0$):

$$\begin{aligned}
\nabla^2\Psi &= 4\pi G_N T_{00} \\
\nabla^2 w_j &= -16\pi G_N T_{0j} + 4\partial_0\partial_j\Psi \\
(\delta_{ij}\nabla^2 - \partial_i\partial_j)\Phi &= -8\pi G_N T_{ij} - (\delta_{ij}\nabla^2 - \partial_i\partial_j)\Psi - \partial_0\partial_{(i}w_{j)} + 2\delta_{ij}\partial_0^2\Psi - \Box s_{ij}
\end{aligned}$$

Now as stated in the construction of the problem, the gravitating source is time-independent and stationary (if it's not, we can always Lorentz-transform to a frame where it is); thus the energy-momentum tensor is simply

$$T_{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

This together with the fact that the source is time-independent (so that time-derivatives vanish) allows us to simplify the linearized EFEs

$$\begin{aligned} \nabla^2\Psi &= 4\pi G_N \rho \\ \nabla^2 w_j &= 0 \\ (\delta_{ij}\nabla^2 - \partial_i\partial_j)\Phi &= -(\delta_{ij}\nabla^2 - \partial_i\partial_j)\Psi - \partial_0\partial_{(i}w_{j)} - \Box s_{ij} \end{aligned}$$

The second equation above implies that $w_j = 0$ since the solution must be well-behaved at infinity; taking the trace of the third equation (remembering that the strain is traceless) yields

$$\begin{aligned} \delta^{ij}(\delta_{ij}\nabla^2 - \partial_i\partial_j)\Phi &= -\delta^{ij}(\delta_{ij}\nabla^2 - \partial_i\partial_j)\Psi - \delta^{ij}\Box s_{ij} \\ (3\nabla^2 - \nabla^2)\Phi &= (3\nabla^2 - \nabla^2)\Psi \\ 2\nabla^2(\Phi - \Psi) &= 0 \end{aligned}$$

This is satisfied for $\Psi = \Phi$. Returning this to the third equation gives

$$\begin{aligned} (\delta_{ij}\nabla^2 - \partial_i\partial_j)\Phi &= -(\delta_{ij}\nabla^2 - \partial_i\partial_j)\Psi - \Box s_{ij} \\ (\delta_{ij}\nabla^2 - \partial_i\partial_j)\Phi &= -(\delta_{ij}\nabla^2 - \partial_i\partial_j)\Phi + \nabla^2 s_{ij} \\ 0 &= \nabla^2 s_{ij} \end{aligned}$$

which again implies that $s_{ij} = 0$. Thus the metric perturbation simplifies greatly to

$$h_{\mu\nu} = \begin{pmatrix} -2\Phi & 0 & 0 & 0 \\ 0 & -2\Phi & 0 & 0 \\ 0 & 0 & -2\Phi & 0 \\ 0 & 0 & 0 & -2\Phi \end{pmatrix}$$

This is the metric we can use in the above equations. This requires the Christoffel symbols, which in linearized gravity are simply

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2}\eta^{\rho\sigma}\left(-\partial_\sigma h_{\mu\nu} + \partial_\nu h_{\sigma\mu} + \partial_\mu h_{\nu\sigma}\right)$$

Since the perturbation metric is diagonal, computing the coefficients of the connection is relatively simple. Let's do so term-by-term, recalling that time-derivatives vanish:

$$\Gamma^0_{00} = \frac{1}{2}\eta^{0\sigma}\left(-\partial_\sigma h_0 + \partial_0 h_{\sigma 0} + \partial_0 h_{0\sigma}\right) = \frac{1}{2}\eta^{00}\left(-\partial_0 h_0\right) = 0$$

$$\Gamma^0_{0\nu} = \frac{1}{2}\eta^{0\sigma}\left(-\partial_\sigma h_{0\nu} + \partial_\nu h_{\sigma 0} + \partial_0 h_{\nu\sigma}\right) = -\frac{1}{2}(-\partial_0 h_{0\nu} + \partial_\nu h_{00}) = -\frac{1}{2}\partial_\nu(-2\Phi) = \partial_\nu\Phi$$

Recalling that time-derivatives vanish, this becomes

$$\Gamma^0_{0i} = \Gamma^0_{i0} = \partial_i\Phi$$

$$\Gamma^0_{ij} = \frac{1}{2}\eta^{0\sigma}\left(-\partial_\sigma h_{ij} + \partial_j h_{\sigma i} + \partial_i h_{j\sigma}\right) = \frac{1}{2}(-1)\left(-\partial_0 h_{ij} + \partial_j h_{0i} + \partial_i h_{j0}\right) = 0$$

Remembering that $h_{\mu\nu}$ is diagonal, we get

$$\Gamma^i_{00} = \frac{1}{2}\eta^{i\sigma}\left(-\partial_\sigma h_{00} + \partial_0 h_{\sigma 0} + \partial_0 h_{0\sigma}\right) = -\frac{1}{2}\partial^i(-2\Phi) = \partial^i\Phi$$

$$\Gamma^i_{0j} = \frac{1}{2}\eta^{i\sigma}\left(-\partial_\sigma h_{0j} + \partial_0 h_{\sigma 0} + \partial_j h_{0\sigma}\right) = \frac{1}{2}\left(-\partial^i h_{0j} + \partial_j h_0{}^i\right) = 0$$

$$\begin{aligned}
\Gamma^i_{jk} &= \frac{1}{2}\eta^{i\sigma}\left(-\partial_\sigma h_{jk} + \partial_k h_{\sigma j} + \partial_j h_{k\sigma}\right) \\
&= \frac{1}{2}(1)\left(-\partial_i h_{jk} + \partial_k h_{ij} + \partial_j h_{ki}\right) \\
&= \frac{1}{2}\left(-\partial_i(-2\Phi\delta_{jk}) + \partial_k(-2\Phi\delta_{ij}) + \partial_j(-2\Phi\delta_{ki})\right) \\
&= \delta_{jk}\partial^i\Phi - \delta^i_k\partial_j\Phi - \delta^i_j\partial_k\Phi
\end{aligned}$$

Now we can use these expressions in the geodesic equation. First let's determine how time is distorted along the perturbed geodesic:

$$\frac{d\ell^0}{d\lambda} = -\Gamma^0_{\nu\rho}k^\nu k^\rho = -\Gamma^0_{00}k^0k^0 - \Gamma^0_{0\rho}k^0k^\rho - \Gamma^0_{\nu 0}k^\nu k^0 - \Gamma^0_{ij}k^ik^j = -2k_0\partial_i\Phi k^i = -2k(\vec{k}\cdot\nabla)\Phi$$

Integrating both sides gives

$$\begin{aligned}
\frac{d\ell^0}{d\lambda} &= -2k(\vec{k}\cdot\nabla)\Phi \\
\ell^0 &= -2k\int \vec{k}\cdot\nabla\Phi d\lambda \\
&= -2k\int \frac{d\vec{x}}{d\lambda}\cdot\nabla\Phi d\lambda \\
&= -2k\int \nabla\Phi\cdot d\vec{x} \\
&= -2k\Phi
\end{aligned}$$

Plugging this into the null geodesic condition gives

$$\begin{aligned}
0 &= 2k\cdot\ell + h_{\mu\nu}k^\mu k^\nu \\
&= -2k^0\ell^0 - 2\vec{k}\cdot\vec{\ell} + (-2\Phi)\delta^{\mu\nu}k^\mu k^\nu \\
&= 2k^2\Phi - 2\vec{k}\cdot\vec{\ell} - 2k^2\Phi \\
&= 2\vec{k}\cdot\vec{\ell}
\end{aligned}$$

Thus we see that $\vec{k} \perp \vec{\ell}$; thus the perturbation lives in a plane perpendicular to the background geodesic at every point along the straight line. This allows us to define a triangle with a given point as one vertex, the vector $\vec{k}$ as one side, the vector $\vec{\ell}$ as the opposite side and the hypotenuse forming an angle α with the adjacent side. Of course, this is only true differentially; to find the total deflection, we must integrate over all such triangles, which means integrating $\frac{d\vec{\ell}}{d\lambda}$ over all λ. Thus we require an expression for $\frac{d\vec{\ell}}{d\lambda}$ which fortunately is the other part of the geodesic equation:

$$\begin{aligned}
\frac{d\ell^i}{d\lambda} &= -\Gamma^i_{\nu\rho} k^\nu k^\rho \\
&= -\Gamma^i_{00} k^0 k^0 - \Gamma^i_{0k} k^0 k^k - \Gamma^i_{j0} k^j k^0 - \Gamma^i_{jk} k^j k^k \\
&= -\partial^i \Phi k^0 k^0 - \left(\delta_{jk}\partial^i\Phi - \delta^{ik}\partial_j\Phi - \delta^i_j \partial_k \Phi\right) k^j k^k \\
&= -\partial^i \Phi k^2 - k^2 \partial^i \Phi + k^i k^j \partial_j \Phi + k^i k^k \partial_k \Phi \\
&= -2k^2 \partial^i \Phi + k^i (k^j \partial_j \Phi) \\
\frac{d\vec{\ell}}{d\lambda} &= -2k^2 (\nabla\Phi - \frac{1}{k^2}\vec{k}(\vec{k}\cdot\nabla\Phi)) \\
&= -2k^2 \nabla_\perp \Phi
\end{aligned}$$

where, in the lasts step, we've defined the $\vec{k}$-transverse gradient as

$$\nabla_\perp \Phi = \nabla\Phi - \frac{1}{k^2}\vec{k}(\vec{k}\cdot\nabla\Phi)$$

We now have an expression to compute the deflection angle α:

$$\alpha = -\frac{\Delta\vec{\ell}}{k} = -\frac{1}{k}\int \frac{d\vec{\ell}}{d\lambda} d\lambda = -\frac{1}{k}(-2k^2)\int \nabla_\perp \Phi d\lambda = 2\int \nabla_\perp \Phi ds$$

Lastly, we need an expression for the scalar potential. Suppose that the gravitating source of mass M may be treated in Newtonian mechanics

$$\Phi = -\frac{G_N M}{r} = -\frac{G_N M}{\sqrt{b^2 + x^2}}$$

where b is the distance of the background geodesic from the source and x is the position along this geodesic. Thus we must integrate over all x to find the deflection:

$$\begin{aligned}
\alpha &= 2\int_{-\infty}^{\infty} \nabla_\perp \Phi dx \\
&= 2\int_{-\infty}^{\infty} \nabla_\perp \left(-\frac{G_N M}{\sqrt{b^2+x^2}}\right) dx \\
&= 2G_N M b \int_{-\infty}^{\infty} \frac{dx}{(b^2+x^2)^{\frac{3}{2}}} = \frac{4G_N M}{b}
\end{aligned}$$

10 Gravitational radiation

10.1 Linearized plane wave solutions

Another simple application of linearized gravity is in the study of gravitational waves and their production. Let's first focus on gravitational waves propagating in free space far from any sources. Recall the EFEs in the transverse gauge ($\partial_i s^{ij} = 0$, $\partial_i w^i = 0$):

$$\begin{aligned}
\nabla^2 \Psi &= 4\pi G_N T_{00} \\
\nabla^2 w_j &= -16\pi G_N T_{0j} + 4\partial_0 \partial_j \Psi \\
(\delta_{ij}\nabla^2 - \partial_i \partial_j)\Phi &= -8\pi G_N T_{ij} - (\delta_{ij}\nabla^2 - \partial_i \partial_j)\Psi - \partial_0 \partial_{(i} w_{j)} + 2\delta_{ij}\partial_0^2 \Psi - \Box s_{ij}
\end{aligned}$$

Now we seek to find how gravitational waves propagate in free space, so we can set the energy-momentum tensor to 0. Thus the EFEs simplify even further:

$$\begin{aligned}
\nabla^2 \Psi &= 0 \\
\nabla^2 w_j &= 4\partial_0 \partial_j \Psi \\
(\delta_{ij}\nabla^2 - \partial_i \partial_j)\Phi &= -(\delta_{ij}\nabla^2 - \partial_i \partial_j)\Psi - \partial_0 \partial_{(i} w_{j)} + 2\delta_{ij}\partial_0^2 \Psi - \Box s_{ij}
\end{aligned}$$

The first equation, together with the fact that oscillations in the gravitational field should vanish at infinity, implies that $\Psi = 0$. Plugging this into the second equation similarly gives $w_j = 0$. Altogether, these conditions simplify the last equation drastically:

$$\begin{aligned}
(\delta_{ij}\nabla^2 - \partial_i \partial_j)\Phi &= -(\delta_{ij}\nabla^2 - \partial_i \partial_j)\Psi - \partial_0 \partial_{(i} w_{j)} + 2\delta_{ij}\partial_0^2 \Psi - \Box s_{ij} \\
(\delta_{ij}\nabla^2 - \partial_i \partial_j)\Phi &= -\Box s_{ij}
\end{aligned}$$

Taking the trace of both sides (recall that s_{ij} is traceless) yields

$$\begin{aligned}
\delta^{ij}(\delta_{ij}\nabla^2 - \partial_i \partial_j)\Phi &= -\delta^{ij}\Box s_{ij} \\
\left(3\nabla^2 - \nabla^2\right)\Phi &= 0 \\
-2\nabla^2 \Phi &= 0
\end{aligned}$$

which, as with the other equations, implies that $\Phi = 0$. This leaves the following equation

$$\Box s_{ij} = 0$$

This is a dramatic simplification! (And all with a judicious choice of gauge). Taking $h_{\mu\nu}$ together with $\Psi = \Phi = w_j = 0$ defines another gauge, the *transverse-traceless gauge*, or in matrix language

$$h_{\mu\nu}^{TT} = \left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & & & \\ 0 & & 2s_{ij} & \\ 0 & & & \end{array}\right)$$

The perturbation matrix $h_{\mu\nu}^{TT}$ satisfies these conditions

$$h_{0\nu}^{TT} = 0 \qquad \eta^{\mu\nu} h_{\mu\nu}^{TT} = 0 \qquad \partial^\mu h_{\mu\nu}^{TT} = 0$$

The EOM for the perturbation matrix in transverse-traceless gauge is

$$\Box h_{\mu\nu}^{TT} = 0$$

As from electrodynamics, this has solutions of the form

$$h_{\mu\nu}^{TT} = C_{\mu\nu} \exp[ik_\sigma x^\sigma]$$

where $C_{\mu\nu}$ must be symmetric, traceless and purely spatial. The EOM imposes the following condition:

$$\begin{aligned} 0 &= \Box h_{\mu\nu}^{TT} \\ &= \Box \left(C_{\mu\nu} \exp[ik_\sigma x^\sigma] \right) \\ &= -k_\sigma k^\sigma C_{\mu\nu} \end{aligned}$$

This implies either $C_{\mu\nu} = 0$ (which is a trivial solution) or $k_\sigma k^\sigma = 0$. Thus we're left with the fact that the wavevector for gravitational waves are null, meaning they must travel at c. Noting that $k^\sigma = (\omega, k^i)$ this translates to

$$0 = k_\sigma k^\sigma = \omega^2 - \delta_{ij} k^i k^j \qquad \Rightarrow \qquad \omega^2 = \delta_{ij} k^i k^j = k^2$$

The transversality equation gives

$$\begin{aligned} 0 &= \partial^\mu h_{\mu\nu}^{TT} \\ &= \partial^\mu \left(C_{\mu\nu} \exp[ik_\sigma x^\sigma] \right) \\ &= ik^\mu C_{\mu\nu} \exp[ik_\sigma x^\sigma] \end{aligned}$$

which yields $k^\mu C_{\mu\nu} = 0$. Thus $C_{\mu\nu}$ must be orthogonal to k_μ, in addition to being symmetric, traceless, and purely spatial. Let's choose a particular direction for the wavevector to propagate $k^\mu = (\omega, 0, 0, \omega)$. Using this definition, we can nail down $C_{\mu\nu}$:

$$0 = k^\mu C_{\mu\nu} \quad \rightarrow \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & C_{11} & C_{12} & C_{13} \\ 0 & C_{12} & C_{22} & C_{23} \\ 0 & C_{13} & C_{23} & C_{33} \end{pmatrix} \begin{pmatrix} \omega \\ 0 \\ 0 \\ \omega \end{pmatrix} \quad \rightarrow \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \omega \begin{pmatrix} 0 \\ C_{13} \\ C_{23} \\ C_{33} \end{pmatrix}$$

Thus we see that $C_{13} = C_{23} = C_{33} = 0$. In order to satisfy the traceless condition, $C_{11} = -C_{22}$. By convention, $C_{11} = h_+$ and $C_{12} = h_\times$. This gives

$$C_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Thus a gravitational wave propagating in the z-direction is completely characterized by three numbers, the frequency ω, h_+ and $h_\times$.

So, how does a passing gravitational wave affect a collection of particles? The motion of a collection of particles at rest is characterized by $U^\rho = (1,0,0,0)$ (they still move through time). The relations between these particles may be described using geodesic deviation:

$$\frac{D^2}{d\tau^2}S^\mu = R^\mu{}_{\nu\rho\sigma}U^\nu U^\rho S^\sigma = R^\mu{}_{00\sigma}S^\sigma$$

We'll need $R^\mu{}_{00\sigma} = \eta^{\mu\nu}R_{\nu 00\sigma} = -\eta^{\mu\nu}R_{0\nu 0\sigma}$. Fortunately, we already calculated this in Ch. 9

$$\begin{aligned}
R^\alpha{}_{\beta\mu\nu} &= \frac{1}{2}\eta^{\alpha\sigma}\left(\partial_\mu\partial_\beta h^{TT}_{\nu\sigma} - \partial_\nu\partial_\beta h^{TT}_{\mu\sigma} - \partial_\mu\partial_\sigma h^{TT}_{\beta\nu} + \partial_\nu\partial_\sigma h^{TT}_{\mu\beta}\right)\\
\eta_{\alpha\gamma}R^\alpha{}_{\beta\mu\nu} &= \frac{1}{2}\eta_{\alpha\gamma}\eta^{\alpha\sigma}\left(\partial_\mu\partial_\beta h^{TT}_{\nu\sigma} - \partial_\nu\partial_\beta h^{TT}_{\mu\sigma} - \partial_\mu\partial_\sigma h^{TT}_{\beta\nu} + \partial_\nu\partial_\sigma h^{TT}_{\mu\beta}\right)\\
R_{\gamma\beta\mu\nu} &= \frac{1}{2}\delta^\sigma_\gamma\left(\partial_\mu\partial_\beta h^{TT}_{\nu\sigma} - \partial_\nu\partial_\beta h^{TT}_{\mu\sigma} - \partial_\mu\partial_\sigma h^{TT}_{\beta\nu} + \partial_\nu\partial_\sigma h^{TT}_{\mu\beta}\right)\\
\rightarrow R_{0\beta 0\nu} &= \frac{1}{2}\delta^\sigma_0\left(\partial_0\partial_\beta h^{TT}_{\nu\sigma} - \partial_\nu\partial_\beta h^{TT}_{0\sigma} - \partial_0\partial_\sigma h^{TT}_{\beta\nu} + \partial_\nu\partial_\sigma h^{TT}_{0\beta}\right)\\
&= \frac{1}{2}\left(\partial_0\partial_\beta h^{TT}_{\nu 0} - \partial_\nu\partial_\beta h^{TT}_{00} - \partial_0\partial_0 h^{TT}_{\beta\nu} + \partial_\nu\partial_0 h^{TT}_{0\beta}\right)\\
&= -\frac{1}{2}\partial_0^2 h^{TT}_{\beta\nu}\\
R_{\beta 00\nu} &= \frac{1}{2}\partial_0^2 h^{TT}_{\beta\nu}\\
R^\mu{}_{00\nu} &= \frac{1}{2}\partial_0^2 h^{TT\mu}{}_\nu
\end{aligned}$$

Thus the geodesic deviation is simply

$$\frac{D^2}{d\tau^2}S^\mu = R^\mu{}_{00\nu}S^\nu \qquad \rightarrow \qquad \frac{D^2}{d\tau^2}S^\mu = \frac{1}{2}\partial_0^2 h^{TT\mu}{}_\nu S^\nu$$

To first order, $\tau \approx x^0 = t$ and thus we get

$$\frac{\partial^2}{\partial t^2}S^\mu = \frac{1}{2}\frac{\partial^2}{\partial t^2}h^{TT\mu}{}_\nu S^\nu = \frac{1}{2}\frac{\partial^2}{\partial t^2}C^\mu{}_\nu \exp[ik_\sigma x^\sigma]S^\nu$$

Since most of $h^{TT}_{\mu\nu} = 0$ we get only a couple of equations

$$\begin{aligned}
\frac{\partial^2}{\partial t^2}S^1 &= \frac{1}{2}\frac{\partial^2}{\partial t^2}\left(C^1{}_1\exp[ik_\sigma x^\sigma]S^1 + C^1{}_2\exp[ik_\sigma x^\sigma]S^2\right) = \frac{1}{2}\frac{\partial^2}{\partial t^2}\left(h_+\exp[ik_\sigma x^\sigma]S^1 + h_\times\exp[ik_\sigma x^\sigma]S^2\right)\\
\frac{\partial^2}{\partial t^2}S^2 &= \frac{1}{2}\frac{\partial^2}{\partial t^2}\left(C^2{}_1\exp[ik_\sigma x^\sigma]S^1 + C^2{}_2\exp[ik_\sigma x^\sigma]S^2\right) = \frac{1}{2}\frac{\partial^2}{\partial t^2}\left(h_\times\exp[ik_\sigma x^\sigma]S^1 - h_+\exp[ik_\sigma x^\sigma]S^2\right)
\end{aligned}$$

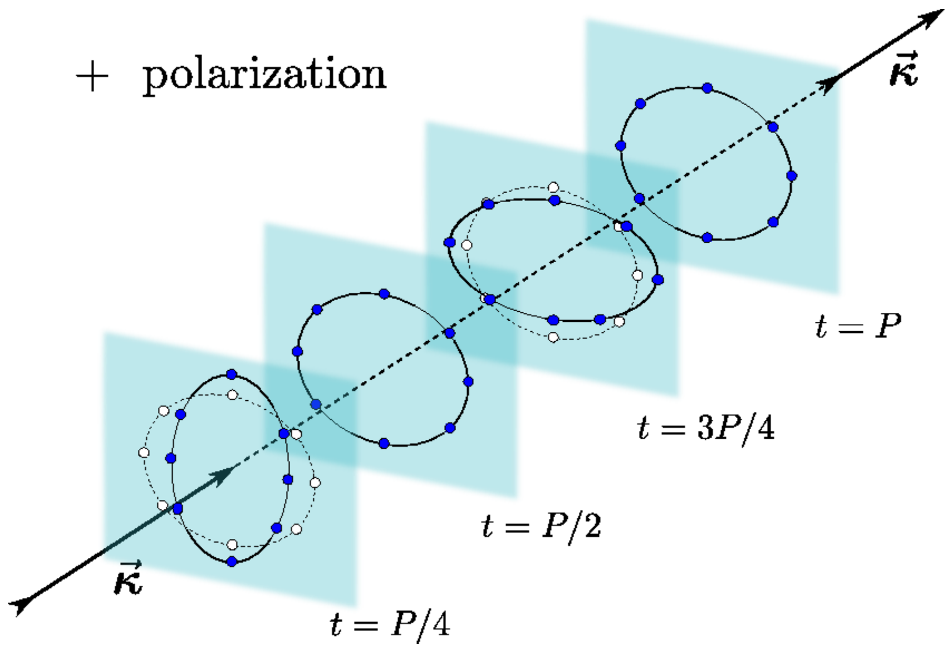
There are two separate cases to consider. In the case $h_+ \neq 0$, $h_\times = 0$, the two equations separate

$$\frac{\partial^2}{\partial t^2}S^1 = \frac{1}{2}\frac{\partial^2}{\partial t^2}\left(h_+\exp[ik_\sigma x^\sigma]S^1\right) \qquad\qquad \frac{\partial^2}{\partial t^2}S^2 = -\frac{1}{2}\frac{\partial^2}{\partial t^2}\left(h_+\exp[ik_\sigma x^\sigma]S^2\right)$$

To first order (meaning the derivatives on the RHS may be neglected), this yields (taking the real portion of the solution)

$$S^1 = \left(1 + \frac{1}{2}\Re[h_+ \exp(k_\sigma x^\sigma)]\right) S^1(0) \qquad\qquad S^2 = \left(1 - \frac{1}{2}\Re[h_+ \exp(k_\sigma x^\sigma)]\right) S^2(0)$$

This is what is known as a *plus-polarized gravitational wave*; its effects are illustrated in the figure below



The effect of these waves is to make the particles oscillate in an up-down/side-side way and the relative phase between the particles is determined their initial positions.

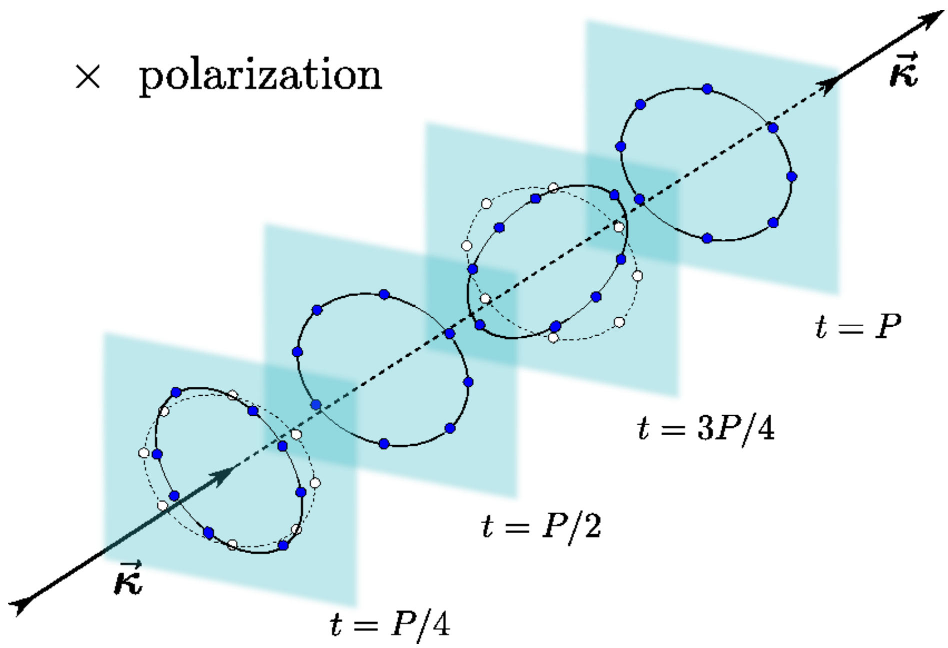
In the opposite case $h_+ = 0$, $h_\times \neq 0$, we get coupled differential equations

$$\frac{\partial^2}{\partial t^2} S^1 = \frac{1}{2}\frac{\partial^2}{\partial t^2}\left(h_\times \exp[ik_\sigma x^\sigma] S^2\right) \qquad\qquad \frac{\partial^2}{\partial t^2} S^2 = \frac{1}{2}\frac{\partial^2}{\partial t^2}\left(h_\times \exp[ik_\sigma x^\sigma] S^1\right)$$

Again solving at first-order, we get

$$S^1 = S^1(0) + \frac{1}{2}\Re[h_\times \exp(k_\sigma x^\sigma)] S^2(0) \qquad\qquad S^2 = S^2(0) + \frac{1}{2}\Re\exp[h_\times \exp(k_\sigma x^\sigma)] S^1(0)$$

which is a *cross-polarized gravitational wave*.



As with plane EM waves, where x and y polarizations may be combined into left- and right-circular polarizations, the plus- and cross-polarizations may be combined to yield left- and right-circularly polarized waves

$$h_R = \frac{1}{\sqrt{2}}(h_+ + ih_\times) \qquad\qquad h_L = \frac{1}{\sqrt{2}}(h_+ - ih_\times)$$

10.2 Wave production/Multipole expansion

Gravitational waves of course must be generated somewhere. Some energy-momentum density described by $T^{\mu\nu}$ will create disturbances in the metric which in turn will propagate through s_{ij} as shown in the previous section. However, this also means that Φ, Ψ and w_i are not necessarily 0. As such, let's work with the full $h_{\mu\nu}$. Once more, recall the linearized EFEs:

$$\partial_\mu \partial^\sigma h_{\nu\sigma} + \partial_\nu \partial^\sigma h_{\sigma\mu} - \Box h_{\mu\nu} - \eta_{\mu\nu}\partial^\alpha \partial^\beta h_{\alpha\beta} - \partial_\mu \partial_\nu h + \eta_{\mu\nu}\Box h = 16\pi G_N T_{\mu\nu}$$

This can be simplified by introducing the *trace-reversed perturbation*

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}h\eta_{\mu\nu}$$

The name refers to the fact that

$$\bar{h} = \eta^{\mu\nu}\bar{h}_{\mu\nu} = \eta^{\mu\nu}\left(h_{\mu\nu} - \frac{1}{2}h\eta_{\mu\nu}\right) = h - \frac{1}{2}h(4) = h - 2h = -h$$

In the limit that $h \to 0$ (meaning that $\Phi = \Psi = 0$) as in the traceless-transverse gauge (which is acceptable in vacuum far away from the source), then $h^{TT}_{\mu\nu} = \bar{h}^{TT}_{\mu\nu}$. Inverting this expression to give $h_{\mu\nu}$ in terms of $\bar{h}_{\mu\nu}$ and plugging this into the EFEs gives:

$$\begin{aligned}
16\pi G_N T_{\mu\nu} &= \partial_\mu \partial^\sigma h_{\nu\sigma} + \partial_\nu \partial^\sigma h_{\sigma\mu} - \Box h_{\mu\nu} - \eta_{\mu\nu}\partial^\alpha \partial^\beta h_{\alpha\beta} - \partial_\mu \partial_\nu h + \eta_{\mu\nu}\Box h \\
&= \partial_\mu \partial^\sigma \left(\bar{h}_{\nu\sigma} + \frac{1}{2}\eta_{\nu\sigma} h\right) + \partial_\nu \partial^\sigma \left(\bar{h}_{\sigma\mu} + \frac{1}{2}\eta_{\sigma\mu} h\right) - \Box \left(\bar{h}_{\mu\nu} + \frac{1}{2}\eta_{\mu\nu} h\right) \\
&\qquad - \eta_{\mu\nu}\partial^\alpha \partial^\beta \left(\bar{h}_{\alpha\beta} + \frac{1}{2}\eta_{\alpha\beta} h\right) - \partial_\mu \partial_\nu h + \eta_{\mu\nu}\Box h \\
&= \partial_\mu \partial^\sigma \bar{h}_{\nu\sigma} + \frac{1}{2}\partial_\mu \partial_\nu h + \partial_\nu \partial^\sigma \bar{h}_{\sigma\mu} + \frac{1}{2}\partial_\mu \partial_\nu h - \Box \bar{h}_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\Box h \\
&\qquad - \eta_{\mu\nu}\partial^\alpha \partial^\beta \bar{h}_{\alpha\beta} - \frac{1}{2}\eta_{\mu\nu}\Box h - \partial_\mu \partial_\nu h + \eta_{\mu\nu}\Box h \\
&= \partial_\mu \partial^\sigma \bar{h}_{\nu\sigma} + \partial_\nu \partial^\sigma \bar{h}_{\sigma\mu} - \eta_{\mu\nu}\partial^\alpha \partial^\beta \bar{h}_{\alpha\beta} - \Box \bar{h}_{\mu\nu}
\end{aligned}$$

This is much nicer! It would be nice if we could condense this equation further; notice that all three terms without a D'Alembertian contain a divergence $\partial_\mu \bar{h}^{\mu\nu}$. We have not yet chosen a gauge; does one exist such that these divergences vanish? First, let's see how the trace-reversed metric perturbation behaves under gauge transformations. We know how $h_{\mu\nu}$ transforms, so let's take the trace and see how it transforms

$$\begin{aligned}
h &= \eta^{\mu\nu} h_{\mu\nu} \\
&\to \eta^{\mu\nu}(h_{\mu\nu} + 2\partial_{(\mu}\xi_{\nu)}) \\
&= h + 2\partial_\mu \xi^\mu
\end{aligned}$$

Thus $\bar{h}_{\mu\nu}$ transforms as

$$\begin{aligned}
\bar{h}_{\mu\nu} &= h_{\mu\nu} - \frac{1}{2} h \eta_{\mu\nu} \\
&\to h_{\mu\nu} + 2\partial_{(\mu}\xi_{\nu)} - \frac{1}{2}\eta_{\mu\nu}\left(h + 2\partial_\sigma \xi^\sigma\right) \\
&= h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} h + 2\partial_{(\mu}\xi_{\nu)} - \eta_{\mu\nu}\partial_\sigma \xi^\sigma \\
&= \bar{h}_{\mu\nu} + 2\partial_{(\mu}\xi_{\nu)} - \eta_{\mu\nu}\partial_\sigma \xi^\sigma
\end{aligned}$$

Now take the divergence of both sides and impose the condition that $\partial_\mu \bar{h}^{\mu\nu} = 0$:

$$\begin{aligned}
\bar{h}_{\mu\nu} &= \bar{h}_{\mu\nu} + 2\partial_{(\mu}\xi_{\nu)} - \eta_{\mu\nu}\partial_\sigma \xi^\sigma \\
\partial^\mu \bar{h}_{\mu\nu} &= \partial^\mu \bar{h}_{\mu\nu} + \partial^\mu \left(\partial_\mu \xi_\nu + \partial_\nu \xi_\mu\right) - \partial_\nu \partial_\sigma \xi^\sigma \\
0 &= \partial^\mu \bar{h}_{\mu\nu} + \Box \xi_\nu + \partial_\nu \partial_\mu \xi^\mu - \partial_\nu \partial_\mu \xi^\mu \\
\Box \xi_\nu &= -\partial^\mu \bar{h}_{\mu\nu}
\end{aligned}$$

Thus we must choose ξ_μ that satisfy the above equation. This is known as the *harmonic* or *Lorenz gauge*. With this choice, the linearized EFEs become

$$\Box \bar{h}_{\mu\nu} = -16\pi G_N T_{\mu\nu}$$

This is nothing more than 10 inhomogenous wave equations (one for each component of $\bar{h}_{\mu\nu}$). We can of course solve this using a Green function

$$G_r(x^\mu, y^\mu) = \frac{1}{4\pi\|\vec{x}-\vec{y}\|}\delta\left(\|\vec{x}-\vec{y}\| - (x^0 - y^0)\right)\Theta(x^0 - y^0)$$

where we have chosen the retarded solution so that the field at x^μ responds to variations made in the past lightcone. Thus given $T_{\mu\nu}$ the metric (using the δ-function to perform the y_0 integral) is

$$\begin{aligned}
\bar{h}_{\mu\nu}(x_0, \vec{x}) &= 4G_N \int \frac{1}{\|\vec{x}-\vec{y}\|}\delta\left(\|\vec{x}-\vec{y}\| - (x^0 - y^0)\right)\Theta(x^0 - y^0)T_{\mu\nu}(y_0, \vec{y})dy_0 d^3y \\
&= 4G_N \int \frac{1}{\|\vec{x}-\vec{y}\|} T_{\mu\nu}(x_0 - \|\vec{x}-\vec{y}\|, \vec{y})d^3y
\end{aligned}$$

It is possible to identify the retarded time $t_r = x_0 - \|\vec{x}-\vec{y}\|$. As in radiation theory in electrodynamics, it is possible to consider different regions of space based on the relative length scales of observation $\vec{x} = r$ compared to the source length scale $\vec{y} = d$ and wavelength λ:

- $d \ll r \ll \lambda$ near zone
- $d \ll r \sim \lambda$ intermediate zone
- $d \ll \lambda \ll r$ far/radiation zone

We'll consider solutions in the radiation zone, emitted by a small, localized, non-relativistic source. In order to employ the above perturbation, it is necessary to use the Fourier transform

$$\begin{aligned}
\tilde{\bar{h}}_{\mu\nu}(\omega, \vec{x}) &= \int \bar{h}_{\mu\nu}(t, \vec{x})e^{i\omega t}\frac{dt}{\sqrt{2\pi}} \\
&= \frac{4G_N}{\sqrt{2\pi}} \int\int \frac{e^{i\omega t}}{\|\vec{x}-\vec{y}\|} T_{\mu\nu}(t - \|\vec{x}-\vec{y}\|, \vec{y})dt d^3y \\
&= \frac{4G_N}{\sqrt{2\pi}} \int\int \frac{e^{i\omega(t_r + \|\vec{x}-\vec{y}\|)}}{\|\vec{x}-\vec{y}\|} T_{\mu\nu}(t_r, \vec{y})dt_r d^3y \\
&= \frac{4G_N}{\sqrt{2\pi}} \int \frac{e^{i\omega\|\vec{x}-\vec{y}\|}}{\|\vec{x}-\vec{y}\|} \int e^{i\omega t_r} T_{\mu\nu}(t_r, \vec{y})dt_r d^3y \\
&= 4G_N \int \frac{e^{i\omega\|\vec{x}-\vec{y}\|}}{\|\vec{x}-\vec{y}\|} \tilde{T}_{\mu\nu}(\omega, \vec{y})d^3y
\end{aligned}$$

Now employ the far field approximation

$$\|\vec{x}-\vec{y}\| \approx r - \hat{n}\cdot\vec{y} \qquad \Rightarrow \qquad \frac{e^{i\omega\|\vec{x}-\vec{y}\|}}{\|\vec{x}-\vec{y}\|} \approx \frac{e^{i\omega r}}{r}$$

Thus the Fourier transform becomes

$$\tilde{\bar{h}}_{\mu\nu}(\omega, \vec{x}) = 4G_N \frac{e^{i\omega r}}{r} \int \tilde{T}_{\mu\nu}(\omega, \vec{y}) d^3y$$

We can use the gauge condition to reduce our work further:

$$\partial^\mu \bar{h}_{\mu\nu} = 0 = \partial_0 \bar{h}^{0\nu} + \partial_i \bar{h}^{i\nu} \qquad \rightarrow \qquad \partial_0 \bar{h}^{0\nu} = -\partial_i \bar{h}^{i\nu}$$

Using the above approximated form of the Fourier transform and the gauge condition yields

$$\begin{aligned}
\partial_0 \bar{h}^{0\nu} &= -\partial_i \bar{h}^{i\nu} \\
\partial_0 \left(4G_N \frac{e^{i\omega r}}{r} \int \tilde{T}^{0\nu}(\omega, \vec{y}) d^3y \right) &= -\partial_i \bar{h}^{i\nu} \\
i\omega \left(4G_N \frac{e^{i\omega r}}{r} \int \tilde{T}^{0\nu}(\omega, \vec{y}) d^3y \right) &= -\partial_i \bar{h}^{i\nu} \\
i\omega h^{0\nu} &= -\partial_i \bar{h}^{i\nu} \\
h^{0\nu} &= \frac{i}{\omega} \partial_i \bar{h}^{i\nu}
\end{aligned}$$

Breaking this up into temporal and spatial components gives

$$\bar{h}^{00} = \frac{i}{\omega} \partial_i \bar{h}^{i0} \qquad \bar{h}^{0j} = \frac{i}{\omega} \partial_i \bar{h}^{ij}$$

Since we know that $\bar{h}^{i0} = \bar{h}^{0i}$ this means that it is possible to recover the temporal part of $\bar{h}^{\mu\nu}$ entirely by determining $\bar{h}^{ij}$. Thus we can restrict ourselves to the spatial part of the energy-momentum tensor

$$\tilde{\bar{h}}_{ij}(\omega, \vec{x}) = 4G_N \frac{e^{i\omega r}}{r} \int \tilde{T}_{ij}(\omega, \vec{y}) d^3y$$

This can be simplified even further by employing energy-momentum conservation:

$$\partial^\mu T_{\mu\nu} = 0 = -\partial_0 T_{0\nu} + \partial_i T_{i\nu} \qquad \rightarrow \qquad \partial_i T_{i\nu} = \partial_0 T_{0\nu}$$

Fourier transforming both sides of the above expression gives

$$\begin{aligned}
\partial_0 T_{0\nu} &= \partial_i T_{i\nu} \\
\partial_0 \int T_{0\nu}(\omega, \vec{x}) e^{i\omega t} \frac{dt}{\sqrt{2\pi}} &= \partial_i T_{i\nu}(t, \vec{x}) \\
i\omega \int T_{0\nu}(\omega, \vec{x}) e^{i\omega t} \frac{dt}{\sqrt{2\pi}} &= \partial_i T_{i\nu}(t, \vec{x}) \\
i\omega T_{0\nu} &= \partial_i T_{i\nu}
\end{aligned}$$

As before, this may be broken up into temporal and spatial pieces

$$i\omega T_{00} = \partial_i T_{i0} \qquad\qquad i\omega T_{0j} = \partial_i T_{ij}$$

Taking the divergence of the second expression and simplifying using the first yields

$$\partial_i \partial_j T_{ij} = i\omega \partial_j T_{0j} = -\frac{1}{2}\omega^2 T_{00}$$

where the factor of $\frac{1}{2}$ appears due to the need to symmetrize the LHS. If we can somehow get the expression on the LHS to appear in our solution, we can substitute for the energy density T_{00} which makes the integral almost trivial. We can do this by integrating by parts:

$$\begin{aligned}
\int \tilde{T}_{ij}(\omega, \vec{y}) d^3y &= \int \tilde{T}_{ik} \delta_j^k d^3y \\
&= \int \tilde{T}_{ik} \partial^k y_j d^3y \\
&= \int \partial^k (\tilde{T}_{ik} y_j) d^3y - \int y_j \partial^k \tilde{T}_{ik} d^3y \\
&= -\int y_j \partial^k \tilde{T}_{mk} \partial^m y_i d^3y \\
&= -\int \partial^m (y_i y_j \partial^k \tilde{T}_{mk}) d^3y + \int y_i y_j \partial^m \partial^k \tilde{T}_{mk} d^3y \\
&= -\frac{\omega^2}{2} \int y_i y_j \tilde{T}_{00}(\omega, \vec{y}) d^3y \\
&= -\frac{\omega^2}{2} \tilde{I}_{ij}(\omega)
\end{aligned}$$

where, in the last step, we've introduced what's known as the quadrupole moment tensor. Thus we find our solution in Fourier space is

$$\tilde{\bar{h}}_{ij}(\omega, \vec{x}) = -2G_N \frac{1}{r} \omega^2 e^{i\omega r} I_{ij}(\omega)$$

This may be Fourier transformed back to give the final solution

$$\bar{h}_{ij}(t, \vec{x}) = \frac{2G_N}{r} \left[\frac{d^2 I_{ij}(t)}{dt^2} \right]_{t=t_r}$$

This is the quadrupole formula for gravitational radiation. Think of it as the analog to electric dipole radiation fields (note the similar $\frac{1}{r}$ spatial dependence).

10.3 Gravitational waves from binary stars

Let's consider a simple application of the quadrupole formula. Recall that this was derived assuming a large distance from a non-relativistic source. As such, consider a binary star system confined to the xy plane, orbiting at a distance R from their common center, where the rate of rotation is much less than the speed of light. We can thus treat this using ordinary Newtonian mechanics; the rate of rotation is determined by matching the centripetal force needed to hold the star in place by the Newtonian gravitational force:

$$\frac{G_N M^2}{(2R)^2} = \frac{Mv^2}{R} \qquad \Rightarrow \qquad v = \sqrt{\frac{G_N M}{4R}}$$

This allows us to determine an orbital frequency Ω:

$$\Omega = \frac{v}{R} = \sqrt{\frac{G_N M}{4R^3}}$$

To determine $I_{ij}(t)$, all we need is the energy density T_{00}; we can treat the stars as point masses and since we're assuming a non-relativistic limit, the rest mass energy far exceeds any kinetic terms. Thus the energy density is simply

$$T_{00}(t, \vec{x}) = M\delta(x_3)\left[\delta(x_1 - R\cos\Omega t)\delta(x_2 - R\sin\Omega t) + \delta(x_1 + R\cos\Omega t)\delta(x_2 + R\sin\Omega t)\right]$$

Thanks to all the δ-functions, computing I_{ij} is relatively easy. Let's begin with I_{11}:

$$\begin{aligned}
I_{11}(t, \vec{x}) &= \int y_1 y_1 \tilde{T}_{00}(t, \vec{y}) d^3y \\
&= M \int \delta(y_3) dy_3 \left[\int y_1^2 \delta(y_1 - R\cos\Omega t) dy_1 \int \delta(y_2 - R\sin\Omega t) dy_2 \right.\\
&\qquad \left. + \int y_1^2 \delta(y_1 + R\cos\Omega t) dy_1 \int \delta(y_2 + R\sin\Omega t) dy_2\right] \\
&= M(1)\left[R^2\cos^2\Omega t(1) + R^2\cos^2\Omega t(1)\right] \\
&= 2MR^2\cos^2\Omega t \\
&= MR^2(1 + \cos 2\Omega t)
\end{aligned}$$

By similar computation,

$$I_{22}(t, \vec{x}) = mR^2(1 - \cos 2\Omega t)$$

Now compute $I_{12} = I_{21}$:

$$\begin{aligned}
I_{11}(t, \vec{x}) &= \int y_1 y_2 \tilde{T}_{00}(t, \vec{y}) d^3y \\
&= M \int \delta(y_3) dy_3 \left[\int y_1 \delta(y_1 - R\cos\Omega t) dy_1 \int y_2 \delta(y_2 - R\sin\Omega t) dy_2 \right.\\
&\qquad \left. + \int y_1 \delta(y_1 + R\cos\Omega t) dy_1 \int y_2 \delta(y_2 + R\sin\Omega) dy_2\right] \\
&= M(1)\left[R^2\cos\Omega t\sin\Omega t + (-R\cos\Omega t)(-R\sin\Omega t)\right] \\
&= 2MR^2\cos\Omega t\sin\Omega t \\
&= MR^2\sin 2\Omega t
\end{aligned}$$

Lastly, noting that

$$\int y_3 \delta(y_3) dy_3 = 0$$

we get $I_{i3} = 0$. Taking the second time-derivative of these expressions and evaluating at $t = t_r$ gives $\bar{h}_{ij}$:

$$\bar{h}_{ij} = \frac{8MR^2\Omega^2 G_N}{r} \begin{pmatrix} -\cos 2\Omega t_r & -\sin 2\Omega t_r & 0 \\ -\sin 2\Omega t_r & \cos 2\Omega t_r & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

The rest of $\bar{h}_{\mu\nu}$ could be recovered by employing the gauge relations

$$\bar{h}^{00} = \frac{i}{\Omega}\partial_i \bar{h}^{i0} \qquad \bar{h}^{0j} = \frac{i}{\Omega}\partial_i \bar{h}^{ij}$$

10.4 Energy-momentum of gravitational waves

11 The Schwarzschild Solution

11.1 Deriving the metric

We will explore the Schwarzschild metric, one of the few exact vacuum solutions to the EFEs, that describes spacetime in the near vicinity of a non-rotating, chargeless massive body of radius R and mass M. Since this is a vacuum solution, the metric $g_{\mu\nu}$ must satisfy the vacuum EFEs, $G_{\mu\nu} = 0$. Furthermore, the metric should be asymptotically flat; for $r \to \infty$, the metric should simply reduce to that of special relativity in spherical coordinates. It should also be flat in the limit that $M \to 0$. Thus we seek a generalization of this metric of the form

$$ds^2 = -A(t,r,\theta,\phi)\mathrm{d}t^2 + B(t,r,\theta,\phi)\mathrm{d}r^2 + C(t,r,\theta,\phi)r^2\mathrm{d}\theta^2 + D(t,r,\theta,\phi)r^2\sin^2\theta\mathrm{d}\phi^2$$

We can employ the symmetries of the problem to greatly simplify the general form of the metric. The system is time-independent and thus A, B, C, D must also be time-independent. Since the body is non-rotating, there exists no preferred direction of observation and therefore A, B, C, D cannot depend on θ or ϕ. Furthermore, spherical symmetry implies that θ and ϕ are essentially interchangeable (what we call the z-axis is arbitrary) so that $C = D$; without loss of generality, it is possible to assume that $C = D = 1$. Thus we seek a metric of the form

$$ds^2 = -A(r)\mathrm{d}t^2 + B(r)\mathrm{d}r^2 + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\mathrm{d}\phi^2$$

or in matrix form

$$g_{\mu\nu} = \begin{pmatrix} -A(r) & 0 & 0 & 0 \\ 0 & B(r) & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2\sin^2\theta \end{pmatrix}$$

In the interest of brevity, we can use *Mathematica* to compute the Einstein tensor, which fortunately is diagonal. Together with the vacuum condition, this yields four differential equations:

$$0 = \frac{A(r)[(B(r)-1)B(r) + rB'(r)]}{r^2B(r)^2}$$

$$0 = \frac{-B(r)A(r) + A(r) + rA'(r)}{r^2A(r)}$$

$$0 = \frac{r[-2B'(r)A(r)^2 + A(r)(2B(r)(A'(r) + rA''(r)) - rA'(r)B'(r)) - rB(r)A'(r)^2]}{rA(r)^2B(r)^2}$$

$$0 = \frac{r\sin^2\theta[-2B'(r)A(r)^2 + A(r)(2B(r)(A'(r) + rA''(r)) - rA'(r)B'(r)) - rB(r)A'(r)^2]}{4A(r)^2B(r)^2}$$

It is possible to simplify these somewhat. The first equation is satisfied if the factor in square brackets vanishes identically. The same is true of the third equation, which can be seen to be identical to the fourth equation. Thus we can write

$$\begin{aligned}
0 &= \frac{(B(r)-1)B(r)+rB'(r)}{r^2B(r)^2} \\
&= \frac{1}{r^2} - \frac{1}{r^2B(r)} + \frac{1}{r}\frac{B'(r)}{B(r)^2} \\
&= \frac{1}{r}\frac{B'(r)}{B(r)^2} + \frac{1}{r^2}\left(1-\frac{1}{B(r)}\right)
\end{aligned}$$

$$\begin{aligned}
0 &= \frac{-B(r)A(r)+A(r)+rA'(r)}{r^2A(r)} \\
&= -\frac{1}{r^2}B(r) + \frac{1}{r^2} + \frac{1}{r}\frac{A'(r)}{A(r)} \\
&= \frac{1}{r}\frac{A'(r)}{A(r)B(r)} - \frac{1}{r^2}\left(1-\frac{1}{B(r)}\right)
\end{aligned}$$

$$\begin{aligned}
0 &= \frac{-2B'(r)A(r)^2 + A(r)(2B(r)(A'(r)+rA''(r)) - rA'(r)B'(r)) - rB(r)A'(r)^2}{rA(r)^2B(r)^2} \\
&= -\frac{2}{r}\frac{B'(r)}{B(r)^2} + \frac{2}{r}\frac{A'(r)}{A(r)B(r)} + \frac{2A''(r)}{A(r)B(r)} - \frac{A'(r)B'(r)}{A(r)B(r)^2} - \frac{A'(r)^2}{A(r)^2B(r)} \\
&= -\frac{B'(r)}{B(r)} + \frac{A'(r)}{A(r)} + \frac{1}{r}\frac{A''(r)}{A(r)} - \frac{r}{2}\frac{A'(r)}{A(r)}\frac{B'(r)}{B(r)} - \frac{r}{2}\frac{A'(r)^2}{A(r)^2}
\end{aligned}$$

The first of these equations depends entirely on $B(r)$ so it can be solved directly:

$$\begin{aligned}
0 &= \frac{1}{r}\frac{B'(r)}{B(r)^2} + \frac{1}{r^2}\left(1-\frac{1}{B(r)}\right) \\
\frac{B'(r)}{B(r)^2} &= -\frac{1}{r}\left(1-\frac{1}{B(r)}\right) \\
-B'(r) &= \frac{1}{r}\left(B(r)^2 - B(r)\right) \\
\Rightarrow -\frac{dr}{r} &= \frac{dB}{B(r)(B(r)-1)} \\
-\log r + c_1 &= -\log\left(\frac{B(r)-1}{B(r)}\right) \\
\frac{c_1}{r} &= \frac{B(r)-1}{B(r)} \\
\frac{c_1}{r}B(r) &= B(r)-1 \\
B(r)\left(\frac{c_1}{r}-1\right) &= -1 \\
B(r) &= \frac{1}{1-\frac{c_1}{r}}
\end{aligned}$$

where we're left with an integration constant c_1 to be determined. We can now plug this back into one of our two other differential equations to solve for $A(r)$; let's do so with the second equation:

$$\begin{aligned}
0 &= \frac{1}{r}\frac{A'(r)}{A(r)B(r)} - \frac{1}{r^2}\left(1 - \frac{1}{B(r)}\right) \\
&= \frac{1}{r}\frac{A'(r)}{A(r)}\left(1 - \frac{c_1}{r}\right) - \frac{1}{r^2}\left(1 - \left(1 - \frac{c_1}{r}\right)\right) \\
&= \frac{A'(r)}{A(r)}(r - c_1) - \frac{c_1}{r} \\
A'(r) &= \frac{c_1 A(r)}{r(r - c_1)} \\
\frac{dA}{A} &= c_1 \frac{dr}{r(r - c_1)} \\
\log A(r) &= c_1\left(\frac{1}{c_1}\right)\log\left(\frac{r - c_1}{r}\right) \\
A(r) &= 1 - \frac{c_1}{r}
\end{aligned}$$

Thus our metric is

$$ds^2 = -\left(1 - \frac{c_1}{r}\right)\mathrm{d}t^2 + \frac{1}{1 - \frac{c_1}{r}}\mathrm{d}r^2 + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\mathrm{d}\phi^2$$

We still need to determine c_1. For $r \gg c_1$, the Schwarzschild metric looks approximately like the Minkowski metric with a perturbation $h_{00} = -\frac{c_1}{r}$. In section 7.2, we found that this term may be interpreted as the Newtonian potential $h_{00} = -2\Phi = -\frac{2G_N M}{r}$. Thus in order to recover Newtonian gravity, $c_1 = 2G_N M$. Thus the full Schwarzschild metric is

$$ds^2 = -\left(1 - \frac{2G_N M}{r}\right)\mathrm{d}t^2 + \frac{1}{1 - \frac{2G_N M}{r}}\mathrm{d}r^2 + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\mathrm{d}\phi^2$$

The constant $c_1 = 2G_N M$ must have units of length and may therefore be interpreted as a radius, known as the Schwarzschild radius $R_S = 2G_N M$.

11.2 The Birkhoff theorem

11.3 Geodesics

11.4 Orbital mechanics

11.5 Singularities

11.6 Maximally extended Schwarzschild solution

12 The Kerr Solution

12.1 Spacetime near a rotating black hole

12.2 Symmetries

12.3 Frame-dragging

12.4 Geodesics

13 Cosmology

13.1 FLRW metric

13.2 The Friedmann equation

13.3 Scale factor evolution

13.4 Inflation

Part IV
Special topics

14 Gravity & Effective Field Theory

15 Gravity & spin

16 String theory

16.1 Nambu-Goto action

16.2 Polyakov action

16.3 AdS-CFT duality

17 Kaluza-Klein theory

18 Twistors