Chapter 2

Lorentzian geometry

2.1 The metric and causal geometry

In the previous chapter, we studied Minkowski space, motivated by the wave equation. In that chapter, the metric η (or equivalently, the matrix of coefficients of the wave equation) was constant. We could pick a basis $\{\vec{e}_\mu\}$ of constant vectors (i.e. vectors whose components are independent of x^μ) such that $\eta_{\mu\nu}$ took constant values on $\mathbb{E}^{1,3}$. In this chapter we shall allow the metric to vary from place to place. The correct setting in which to do that is the differentiable manifold.

Manifolds

I'll start by recapping the geometric tools that we shall rely on. We start with an $n-$dimensional C^k-manifold, $\mathcal{M}$, where $k > 3$ can be an integer if the manifold has finite regularity, $k = \infty$ if the manifold is smooth or $k = \omega$ if the manifold is analytic. From the differentiable structure we can define the tangent bundle $T\mathcal{M}$ and the cotangent bundle $T^*\mathcal{M}$, and the bundle of $(p, q)-$tensors, $T^p{}_q\mathcal{M}$. We also define the space $\mathfrak{X}_r(\mathcal{M})$ of C^r-sections of the tangent bundle for $r < k$ (i.e. C^r-smooth vector fields) and the space $\mathfrak{X}_r^*(\mathcal{M})$ of C^r-sections of the co-tangent bundle (i.e. C^r-smooth one-forms). You will need to be familiar with these objects. For a brief introduction to these concepts, see §A.3 or else the MA3H5 Manifolds course.

Definition 9. Let $\mathcal{M}$ be an orientable, $n-$dimensional C^k-manifold and let $r < k$. A C^r-*pseudo-Riemannian metric* on $\mathcal{M}$ is a C^r-smooth $(0, 2)-$tensor field g, such that for each $p \in \mathcal{M}$, $g|_p$ is a symmetric, non-degenerate, billinear form on $T_p\mathcal{M}$. We say that $\mathcal{M}$ equipped with g is a pseudo-Riemannian manifold. The tensor g is called the *metric tensor*.

1.) If $g|_p$ has signature $(+ \ldots +)$ we say that $\mathcal{M}$ equipped with g is a *Riemannian manifold.*

2.) If $g|_p$ has signature $(- + \ldots +)$ we say that $\mathcal{M}$ equipped with g is a *Lorentzian manifold.*

A *spacetime* is a four dimensional Lorentzian manifold.

To unpack this definition a little, suppose that we have a local basis of vector fields $\{e_\mu\}_{i=1,\dots,n}$ defined on some open set $\mathcal{U}$. In other words, we have $e_\mu \in \mathfrak{X}_{k-1}(\mathcal{U})$, such that $\{e_\mu|_p\}$ span $T_p\mathcal{M}$ for every $p \in \mathcal{U}$. There is a natural basis of one-form fields dual to $\{e_\mu\}_{i=1,\dots,n}$, which we write as $\{e^\mu\}_{i=1,\dots,n}$, and is defined by

$$e^\nu[e_\mu] = \delta^\nu{}_\mu$$

With respect to this basis, we write

$$g = g_{\mu\nu} e^\mu \otimes e^\nu.$$

Where for each μ, ν we have $g_{\mu\nu} \in C^r(\mathcal{M}; \mathbb{R})$. The condition that g is symmetric implies that $g_{\mu\nu} = g_{\nu\mu}$, and the non-degeneracy condition implies that the $n \times n$ matrix with components $g_{\mu\nu}$ is invertible. The components of the inverse of this matrix, $g^{\mu\nu}$, which satisfy:

$$g_{\mu\nu} g^{\mu\sigma} = \delta_\nu{}^\sigma.$$

give the components of a symmetric $(2,0)-$tensor, called the co-metic:

$$g^{-1} = g^{\mu\nu} e_\mu \otimes e_\nu.$$

The metric at each point p gives us a bilinear form on $T_p\mathcal{M}$, where we have:

$$g(X, Y) = g_{\mu\nu} e^\mu(X) e^\nu(Y) = g_{\mu\nu} X^\mu Y^\nu$$

For any $X, Y \in T_p\mathcal{M}$ Similarly, the cometric gives a bilinear form on $T^*_p\mathcal{M}$, where:

$$g^{-1}(\omega, \eta) = g^{\mu\nu} \omega(e_\mu) \eta(e_\nu) = g^{\mu\nu} \omega_\mu \eta_\nu.$$

We can use the metric and co-metric to identify $T_p\mathcal{M}$ and $T^*_p\mathcal{M}$. Suppose $X \in T_p\mathcal{M}$, we define the co-vector $X^\flat$ by:

$$X^\flat[Y] = g(X, Y), \qquad \forall \; Y \in T_p\mathcal{M}$$

similarly, if $\omega \in T^*_p\mathcal{M}$, we define the vector $\omega^\sharp$ by:

$$\eta[\omega^\sharp] = g^{-1}(\omega, \eta), \qquad \forall \; \eta \in T^*_p\mathcal{M}$$

We can check that in a local basis, we have $X^\flat = X^\mu g_{\mu\nu} e^\nu$ and $\omega^\sharp = \omega_\mu g^{\mu\nu} e_\nu$. In this case, we usually write $g_{\mu\nu} X^\mu = X_\nu$ and $\omega_\mu g^{\mu\nu} = \omega^\nu$. In a similar way, we can identify all of the spaces of $(p, q)-$tensors with the same value of $p + q$. Indices are 'raised' with $g^{\mu\nu}$ and 'lowered' with $g_{\mu\nu}$.

Now, to define the signature we use the fact that a non-degenerate bilinear form can be diagonalised. In other words, at each point we can find a basis $\{e_\mu\}$ such that

$$g_{\mu\nu} = g(e_\mu, e_\nu) = \begin{cases} -1 & \mu = \nu < r \\ 1 & \mu = \nu \geq r \\ 0 & \mu \neq \nu \end{cases}$$

for some $r \in \{1, \dots, n\}$. The number r, which tells us how many positive and how many negative signs appear is independent of the point p at which we diagonalise the metric.

For a Riemannian manifold, the metric is positive definite, each point $p \in \mathcal{M}$ the tangent space $T_p\mathcal{M}$ can be identified with Euclidean space. For a spacetime, at each point $p \in \mathcal{M}$ the tangent space $T_p\mathcal{M}$ can be identified with $\mathbb{E}^{1,3}$. A Lorentzian manifold 'locally looks like Minkowski space', whereas a Riemannian manifold 'locally looks like Euclidean space'.

2.1.1 Examples of pseudo-Riemannian manifolds

We will go through a few examples that will prove useful later in the course.

Example 3. The simplest Riemannian manifold is a $n-$dimensional real vector space V equipped with a positive definite inner product $\langle , \rangle$. A real, finite dimensional vector space is naturally a real analytic manifold, which can be covered by a single coordinate chart as follows. Set $\mathcal{U} = V$. Pick an orthonormal basis $\{\boldsymbol{e}_i\}_{i=1,\dots n}$ for V with respect to the inner product and define:

$$\begin{array}{rcl} \varphi \ : \ \mathcal{U} & \to & \mathbb{R}^n \\ \boldsymbol{x} & \mapsto & (x^i)_{i=1,\dots,n} \end{array}$$

where $\boldsymbol{x} = x^i \boldsymbol{e}_i$.

An element v of $T_{\boldsymbol{x}}V$ can be identified with an element of V in a natural way. Pick a curve $\gamma : (-\epsilon, \epsilon) \to V$ with $\gamma(0) = \boldsymbol{x}$, $\dot\gamma(0) = v$. We know from the definition of v that

$$v^i := \frac{d}{dt}\varphi^i(\gamma(t))\bigg|_{t=0}$$

are independent of the choice of representative curve, where φ^i is the projection of φ onto the i^{th} component, which is a smooth map from V to $\mathbb{R}$. We identify the tangent vector $v \in T_{\boldsymbol{x}}V$ with the vector $\boldsymbol{v} = v^i \boldsymbol{e}_i \in V$. This identification is independent of the choice of basis, so in a natural way we have identified $T_{\boldsymbol{x}}V \simeq V$. Under this identification, we have

$$\frac{\partial}{\partial x^i} \simeq \boldsymbol{e}_i.$$

This identification allows us to define a non-degenerate, symmetric $(0, 2)-$tensor field from the inner product $\langle , \rangle$. Recall that a $(0, 2)-$tensor at a point $\boldsymbol{x}$ is an element of $T^*_{\boldsymbol{x}}V \otimes T^*_{\boldsymbol{x}}V$, or equivalently a bilinear form defined on $T_{\boldsymbol{x}}V$. We define

$$g(v, w) = \langle \boldsymbol{v}, \boldsymbol{w} \rangle$$

where $v, w \in T_{\boldsymbol{x}}V$ are identified with $\boldsymbol{v}, \boldsymbol{w}$.

With respect to the basis $\{\partial_i\}$, we have

$$g\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) = \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = \delta_{ij}.$$

In order to write g concisely, we can use the dual basis to $\{\partial_i\}$ is denoted $\{dx^i\}$, where

$$dx^i\left(\frac{\partial}{\partial x^j}\right) = \delta^i{}_j.$$

We also introduce the symmetric product of two one-forms, which is an element of $(T_{\boldsymbol{x}})^0{}_2V$, i.e. a $(0,2)-$tensor defined by

$$\omega_1\omega_2 := \frac{1}{2}\left(\omega_1 \otimes \omega_2 + \omega_2 \otimes \omega_1\right)$$

for $\omega_1, \omega_2 \in T^*_{\boldsymbol{x}}V$. This acts on a pair of vectors $v_1, v_2 \in T_{\boldsymbol{x}}V$ as:

$$\omega_1\omega_2(v_1, v_2) = \frac{1}{2}\left[\omega_1(v_1)\omega_2(v_2) + \omega_2(v_1)\omega_1(v_2)\right]$$

With this notation, the metric g can be written:

$$g = \delta_{ij}dx^i dx^j$$

Example 4. A more interesting example of a Riemannian manifold is the unit $n-$sphere:

$$S^n = \left\{\boldsymbol{X} \in \mathbb{R}^{n+1} : \langle \boldsymbol{X}, \boldsymbol{X}\rangle = 1\right\}$$

with $\langle , \rangle$ the canonical inner product on $\mathbb{R}^{n+1}$. The unit $n-$sphere is naturally a real analytic Riemannian manifold (in fact it inherits these properties from $\mathbb{R}^{n+1}$). To cover the sphere, we require two coordinate charts. We'll pick an orthonormal basis $\{\boldsymbol{E}_a\}_{a=1,\dots n+1}$ for $\mathbb{R}^{n+1}$ and we define

$$\mathcal{U}_\pm = S^n \setminus \{\pm\boldsymbol{E}_{n+1}\}.$$

On each patch, we use stereographic projection to map to $\mathbb{R}^n$, which we endow with the orthonormal basis $\{\boldsymbol{e}_i\}$.

$$\begin{array}{cccccc} \varphi_\pm & : & \mathcal{U}_\pm & \to & \mathbb{R}^n \\ & & X^a\boldsymbol{E}_a & \mapsto & \frac{1}{1\mp X^{n+1}}X^i\boldsymbol{e}_i =: \boldsymbol{x}_\pm \end{array}$$

We can invert the transformation by noting that $X^iX_i + (X^{n+1})^2 = 1$ and

$$|\boldsymbol{x}_\pm|^2 = \frac{X^iX^j\delta_{ij}}{(1\mp X^{n+1})^2} = \frac{1-(X^{n+1})^2}{(1\mp X^{n+1})^2} = \frac{1\pm X^{n+1}}{1\mp X^{n+1}}$$

Thus

$$\varphi_\pm^{-1}(\boldsymbol{x}) = \frac{2x^i}{1+|\boldsymbol{x}|^2}\boldsymbol{E}_i \mp \frac{1-|\boldsymbol{x}|^2}{1+|\boldsymbol{x}|^2}\boldsymbol{E}_{n+1} \tag{2.1}$$

From here we conclude:

$$\begin{array}{ccccc} \varphi_\mp \circ \varphi_\pm^{-1} & : & \mathbb{R}^n\setminus\{\boldsymbol{0}\} & \to & \mathbb{R}^n\setminus\{\boldsymbol{0}\} \\ & & \boldsymbol{x} & \mapsto & \frac{\boldsymbol{x}}{|\boldsymbol{x}|^2} \end{array}$$

so that the transition functions are indeed real analytic, and S^n is a real analytic manifold.

The tangent space $T_{\boldsymbol{X}}S^n$ can be identified in a natural way with the set:

$$T_{\boldsymbol{X}}S^n \simeq \left\{\boldsymbol{V} \in \mathbb{R}^{n+1} : \langle \boldsymbol{V}, \boldsymbol{X}\rangle = 0\right\}.$$

To see this, consider a tangent vector $V \in T_{\boldsymbol{X}}S^n$ and choose a curve $\gamma : (\epsilon, \epsilon) \to S^n$ with $\gamma(0) = \boldsymbol{X}$, $\dot\gamma(0) = V$. Thinking of γ as a curve in $\mathbb{R}^{n+1}$, we can identify its initial tangent vector with $\boldsymbol{V} \in \mathbb{R}^{n+1}$ as in the previous example. We know that the map $f : \mathbb{R}^{n+1} \to \mathbb{R}$ which takes $\boldsymbol{X}$ to $|\boldsymbol{X}|^2$ is smooth, and constant on S^n. By applying the vector $\dot\gamma(0)$ to this map we deduce that

$$\begin{aligned} 0 = \dot\gamma(0)[f] &= \frac{d}{dt}\left\langle \gamma(t), \gamma(t)\right\rangle\bigg|_{t=0} \\ &= 2\left\langle \gamma(0), \dot\gamma(0)\right\rangle = \left\langle \boldsymbol{V}, \boldsymbol{X}\right\rangle. \end{aligned}$$

This identification of the tangent space of S^n (an abstract manifold construction) with a plane in $\mathbb{R}^{n+1}$ allows us to visualise the two local bases:

$$\left\{\frac{\partial}{\partial x^i_\pm}\right\}_{i=1\ldots n}$$

that are defined via the coordinate charts. Recall that the vector ∂_i at a point $\boldsymbol{X}$ is defined to be the tangent vector to the i^{th} coordinate axis passing through $\boldsymbol{X}$. Accordingly, we can find the bases by differentiating the expression (2.1). We obtain:

$$\begin{aligned} \frac{\partial}{\partial x^i_\pm}\bigg|_{\boldsymbol{X}=\varphi_\pm^{-1}(\boldsymbol{x})} &\simeq \left(\frac{2}{1+|\boldsymbol{x}|^2}\delta_i{}^j - \frac{4x_ix^j}{\left(1+|\boldsymbol{x}|^2\right)^2}\right)\boldsymbol{E}_j \pm \frac{4x_i}{\left(1+|\boldsymbol{x}|^2\right)^2}\boldsymbol{E}_{n+1} \\ &= \left[(1\mp X^{n+1})\delta_i{}^j - X_iX^j\right]\boldsymbol{E}_j \pm (1\mp X^{n+1})X_i\boldsymbol{E}_{n+1} \end{aligned}$$

We see that $\frac{\partial}{\partial x^i_\pm}$ span $T_{\boldsymbol{X}}S^n$ for all $\boldsymbol{X} \in \mathcal{U}_\pm$.

The identification of $T_{\boldsymbol{X}}S^n$ with a subspace of $\mathbb{R}^{n+1}$ gives us a natural way to define a metric on S^n using the inner product of $\mathbb{R}^{n+1}$:

$$g(v, w) = \langle \boldsymbol{V}, \boldsymbol{W}\rangle$$

where $v, w \in T_{\boldsymbol{X}}S^n$ are identified with $\boldsymbol{V}, \boldsymbol{W} \in \mathbb{R}^{n+1}$. We calculate

$$\begin{aligned} g\left(\frac{\partial}{\partial x^i_\pm}, \frac{\partial}{\partial x^j_\pm}\right) &= \left[(1\mp X^{n+1})\delta_i{}^k - X_iX^k\right]\left[(1\mp X^{n+1})\delta_{jk} - X_jX_k\right] \\ &\quad + (1\mp X^{n+1})^2X_iX^i \\ &= (1\mp X^{n+1})^2\delta_{ij} - 2X_iX_j(1\mp X^{n+1}) + X_iX_jX^kX_k \\ &\quad + (1\mp X^{n+1})^2X_iX^i \\ &= (1\mp X^{n+1})^2\delta_{ij} \\ &= \frac{4\delta_{ij}}{\left(1+|\boldsymbol{x}_\pm|^2\right)^2} \end{aligned}$$

We thus have the result that in each stereographic coordinate patch, the standard metric on the unit sphere is given by

$$g = \frac{4\delta_{ij}}{\left(1 + |\boldsymbol{x}_\pm|^2\right)^2} dx^i_\pm dx^j_\pm.$$

Example 5. Let's return to the example of $\mathbb{R}^{n+1}$ equipped with an inner product and an orthonormal basis $\{\boldsymbol{E}_a\}_{a=1,\ldots n+1}$. Suppose we have a point $\boldsymbol{P} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$. We can uniquely write $\boldsymbol{P} = r\boldsymbol{X}$, where $r > 0$ is a real number and $\boldsymbol{X} \in S^n$. In this way we can identify $\mathbb{R}^n \setminus \{\boldsymbol{0}\} \simeq (0, \infty) \times S^n$.

We use this observation to cover $\mathbb{R}^n \setminus \{\boldsymbol{0}\}$ with two coordinate patches:

$$\mathcal{U}_\pm = \mathbb{R}^{n+1} \setminus \{\pm r\boldsymbol{E}_{n+1} : r \geq 0\}$$

and define the maps

$$\begin{array}{rcrcl} \varphi_\pm & : & \mathcal{U}_\pm & \to & (0,\infty) \times \mathbb{R}^n \\ & & P^a \boldsymbol{E}_a & \mapsto & \left(|\boldsymbol{P}|, \frac{1}{|\boldsymbol{P}| \mp P^{n+1}} P^i \boldsymbol{e}_i\right). \end{array}$$

In other words, $\varphi_\pm$ maps the point $r\boldsymbol{X}$ to $(r, \boldsymbol{x}_\pm)$, where $\boldsymbol{x}_\pm$ is the stereographic projection of the point $\boldsymbol{X} \in S^n$. The inverse map is given by

$$\varphi^{-1}_\pm (r, \boldsymbol{x}) = r \left(\frac{2x^i}{1 + |\boldsymbol{x}|^2} \boldsymbol{E}_i \mp \frac{1 - |\boldsymbol{x}|^2}{1 + |\boldsymbol{x}|^2} \boldsymbol{E}_{n+1} \right) \tag{2.2}$$

As before, we can identify the tangent vectors with vectors in $\mathbb{R}^{n+1}$. A simple calculation shows that

$$\frac{\partial}{\partial r} \simeq \boldsymbol{X} = \frac{\boldsymbol{P}}{|\boldsymbol{P}|}$$

and

$$\frac{\partial}{\partial n^i_\pm} \simeq r \left\{ \left[(1 \mp X^{n+1}) \delta_i{}^j - X_i X^j \right] \boldsymbol{E}_j \pm (1 \mp X^{n+1}) X_i \boldsymbol{E}_{n+1} \right\}$$

We can calculate

$$g\left(\frac{\partial}{\partial r}, \frac{\partial}{\partial r}\right) = 1, \qquad g\left(\frac{\partial}{\partial r}, \frac{\partial}{\partial n^i_\pm}\right) = 0$$

$$g\left(\frac{\partial}{\partial x^i_\pm}, \frac{\partial}{\partial x^j_\pm}\right) = \frac{4\delta_{ij}}{\left(1 + |\boldsymbol{x}_\pm|^2\right)^2}$$

so that in each patch we have

$$g = dr^2 + r^2 \frac{4\delta_{ij}}{\left(1 + |\boldsymbol{x}_\pm|^2\right)^2} dx^i_\pm dx^j_\pm$$

which is the metric on $\mathbb{R}^{n+1}\setminus\{\mathbf{0}\}$ written in spherical polar coordinates, with stereographic coordinates on the sphere. We have shown that $\mathbb{R}^{n+1}\setminus\{\mathbf{0}\}$ is equivalent, as a Riemannian manifold, to $(0,\infty)\times S^n$ endowed with the metric

$$g = dr^2 + r^2 g_{S^n}.$$

Often the fact that this describes $\mathbb{R}^{n+1}$ with the origin removed is elided, and one will talk of the metric on $\mathbb{R}^{n+1}$ written in polar coordinates.

Example 6. Take $\mathcal{M} = \mathbb{R}^4$ with coordinates $(x^0, x^i) = (x^\mu)_{\mu=0,\ldots,3}$. As previously, this gives us a global basis of vector fields $\partial_\mu := \frac{\partial}{\partial x^\mu}$. We can endow $\mathbb{R}^4$ with a Lorentzian metric by:

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

This is of course the Minkowski spacetime. It is clearly Lorentzian, if we take the basis $e_\mu = \partial_\mu$, we have

$$g(e_\mu, e_\nu) = \eta_{\mu\nu}.$$

Example 7. We take $\mathcal{M} = \mathbb{R}\times(0,\infty)\times S^2$. This can be given the structure of a real analytic manifold covered by the coordinate patches

$$\mathcal{U}_\pm = \mathbb{R}\times(0,\infty)\times\mathcal{U}_\pm^{S^2}$$

where $\mathcal{U}_\pm^{S^2}$ are the stereographic coordinate patches on S^2 previously defined. The coordinate charts are given by

$$\begin{array}{rccl} \varphi_\pm & : & \mathcal{U}_\pm & \to & \mathbb{R}\times(0,\infty)\times\mathbb{R}^2 \\ & & (t, r, \boldsymbol{X}) & \mapsto & (t, r, \boldsymbol{x}_\pm). \end{array}$$

with the notation as in Example 4. We endow $\mathcal{M}$ with the metric

$$g = -\left(1 - \frac{2m}{r}\right) dt^2 + 2\sqrt{\frac{2m}{r}} dt dr + dr^2 + r^2 \frac{4\delta_{AB} dx_\pm^A dx_\pm^B}{\left(1 + |\boldsymbol{x}_\pm|^2\right)^2},$$

where $A, B = 2, 3$. More concisely, we write

$$g = -\left(1 - \frac{2m}{r}\right) dt^2 + 2\sqrt{\frac{2m}{r}} dt dr + dr^2 + r^2 g_{S^2}.$$

This is a Lorentzian metric, as we can see by exhibiting the following local bases:

$$\begin{aligned} e_0 &= \frac{\partial}{\partial t} - \sqrt{\frac{2m}{r}} \frac{\partial}{\partial r} \\ e_1 &= \frac{\partial}{\partial r} \\ e_A^\pm &= \frac{1 + |\boldsymbol{x}_\pm|^2}{2r} \frac{\partial}{\partial x_\pm^A}. \end{aligned}$$

A short calculation shows that

$$g(e^{\pm}_{\mu}, e^{\pm}_{\nu}) = \eta_{\mu\nu},$$

so that in a neighbourhood of every point in $\mathcal{M}$ we can find a local basis with respect to which the metric is diagonal with entries $(- + ++)$.

This metric is the Painlevé-Gullstrand form of the Schwarzschild spacetime. It represents a spacetime containing a black hole region (we shall see what this means shortly).

2.1.2 Causal geometry for Lorentzian manifolds

We will now update some definitions from the previous Chapter. For the most part, things go through in a very straightforward fashion.

Definition 10. A non-zero vector $X \in T_p\mathcal{M}$ is

i) *Timelike* if $g(X, X) < 0$,

ii) *Null*, or *lightlike* if $g(X, X) = 0$,

iii) *Spacelike* if $g(X, X) > 0$.

The definitions of timelike/null/spacelike curves and surfaces generalise from the Minkowski case in the obvious fashion.

We have to adapt a little the definition of a time orientation in the manifold case.

Definition 11. The manifold $\mathcal{M}$ is *time orientable* if there exists a smooth, nowhere vanishing timelike vector field. For a time orientable manifold, a choice of time orientation is a choice of equivalence class of nowhere vanishing timelike vector fields under the equivalence relation

$$T_1 \sim T_2 \quad \Longleftrightarrow \quad g(T_1, T_2) < 0 \text{ on } \mathcal{M}.$$

With a time orientation $[T]_\sim$, we can define what it means for a vector $X \in T_p\mathcal{M}$ to be future directed. We say that X is future directed if $g(X, T)|_p < 0$ for any $T \in [T]_\sim$.

For a time oriented manifold we can define the notion chronological and causal future/past and domain of dependence as for the case of Minkowski spacetime.

Example 8. Let us consider the Painlevé-Gullstrand form of the Schwarzschild spacetime, as in Example 7 above. The surface

$$\Sigma_t = \{t\} \times (0, \infty) \times S^2$$

is everywhere spacelike: local bases for the vectors tangent to Σ_t are given by $\{e^{\pm}_1, e^{\pm}_2, e^{\pm}_3\}$, which are all spacelike.

The spacetime is time orientable: the vector field

$$e_0 = \frac{\partial}{\partial t} - \sqrt{\frac{2m}{r}} \frac{\partial}{\partial r}$$

is defined everywhere and timelike, we pick the time orientation defined by $[e_0]_\sim$. In fact, e_0 is the future directed unit normal to Σ_t.

Consider now a future directed causal curve $\gamma : (-\epsilon, \epsilon) \to \mathcal{M}$. We can write

$$\gamma(s) = (t(s), r(s), \boldsymbol{X}(s)) \in \mathbb{R} \times (0, \infty) \times S^2$$

and the tangent vector is given by:

$$\dot{\gamma}(s) = \dot{t}(s)\frac{\partial}{\partial t} + \dot{r}(s)\frac{\partial}{\partial r} + \dot{\boldsymbol{X}}(s)$$

where we understand $\dot{\boldsymbol{X}}(s) \in T_{\boldsymbol{X}(s)}S^2$. We can calculate:

$$g(\dot{\gamma}(s), e_0) = -\dot{t}(s) < 0$$

by the fact that γ is future directed. Now consider

$$g(\dot{\gamma}(s), \dot{\gamma}(s)) = -\left(1 - \frac{2m}{r(s)}\right)\dot{t}(s)^2 + 2\sqrt{\frac{2m}{r(s)}}\dot{t}(s)\dot{r}(s) + \dot{r}(s)^2 + r^2 g_{S^2}\left(\dot{\boldsymbol{X}}(s), \dot{\boldsymbol{X}}(s)\right),$$

Now, in order that γ is causal, we must have $g(\dot{\gamma}(s), \dot{\gamma}(s)) \leq 0$. Suppose now that $r(s) \leq 2m$ for some s. Then the only way that we can have $g(\dot{\gamma}(s), \dot{\gamma}(s)) \leq 0$ is if $\dot{r}(s) \leq 0$. Thus no future directed causal curve can escape from the region $r \leq 2m$. This region is known as a *black hole region* of the spacetime. We have shown that

$$J^+\left(\{0\} \times (0, 2m] \times S^2\right) \subset [0, \infty) \times (0, 2m] \times S^2$$

In fact, it is fairly straightforward to show that the reverse inclusion holds and that the two sets are equal.

Exercise 2.3. Consider the infinite cylinder $\mathbb{R} \times S^1$ and take as coordinates (x, θ) where $\theta \sim \theta + 2\pi$.

a) Show that when $\mathcal{M}$ is equipped with the Lorentzian metric

$$g = -dx^2 + d\theta^2,$$

it is time-orientable.

b) Now consider the metric

$$g = -\cos\theta dx^2 + 2\sin\theta dx d\theta + \cos\theta d\theta^2$$

i) Show that the vector fields defined for $\theta \in [0, 2\pi)$ by

$$X_0 = \cos\frac{\theta}{2}\partial_x - \sin\frac{\theta}{2}\partial_\theta, \quad X_1 = \sin\frac{\theta}{2}\partial_x + \cos\frac{\theta}{2}\partial_\theta.$$

satisfy

$$g(X_0, X_0) = -1, \quad g(X_0, X_1) = 0, \quad g(X_1, X_1) = 1.$$

Deduce that g is a Lorentzian metric.

ii) Let us denote the point $x = 0, \theta = 0$ by p. Suppose that there exists a nowhere vanishing timelike field T, and without loss of generality assume that $g(X_0, T)|_p < 0$. Show that if $\gamma : [0, 1) \to \mathbb{R} \times [0, 2\pi)$ is any smooth curve with $\gamma(0) = p$ then $g(X_0, T)|_\gamma < 0$.

iii) By considering the curve $\gamma : s \mapsto (0, 2\pi s)$, deduce that $\mathcal{M}$ is not time orientable.

2.2 Affine Connections

We are going to require a means to differentiate vectors (and other tensors). This is a subtle point, and is not completely straightforward. Suppose $\gamma : (-\epsilon, \epsilon) \to \mathcal{M}$ is a curve and $f \in C^1(\mathcal{M})$ is a differentiable function, and we wish to differentiate f along γ. We define

$$\dot{\gamma}(0)[f] = \lim_{h \to 0} \frac{f(\gamma(h)) - f(\gamma(0))}{h}.$$

Now suppose that we have a vector field $X \in \mathfrak{X}_1(\mathcal{M})$. We might try and define the derivative of X along γ by:

$$\dot{\gamma}(0)[X] \stackrel{?}{=} \lim_{h \to 0} \frac{X|_{\gamma(h)} - X|_{\gamma(0)}}{h},$$

however, there is a problem with this. The two vectors whose difference we want to form do not belong to the same vector space: one belongs to $T_{\gamma(h)}\mathcal{M}$ and the other to $T_{\gamma(0)}\mathcal{M}$. We need to decide how to identify the tangent spaces along γ before we can define a quotient such as this. The tool to do this is provided by an 'affine connection'.

Definition 12. Suppose $\mathcal{M}$ is a C^k manifold. A C^r-affine connection ∇, where $r \leq k-2$, is a map from $\mathfrak{X}_r(\mathcal{M}) \times \mathfrak{X}_{r+1}(\mathcal{M}) \to \mathfrak{X}_r(\mathcal{M})$ satisfying:

i) For every constant $\lambda \in \mathbb{R}$, $X_i \in \mathfrak{X}_r(\mathcal{M})$, $Y_i \in \mathfrak{X}_{r+1}(\mathcal{M})$ we have

$$\nabla_{X_1}(Y_1 + \lambda Y_2) = \nabla_{X_1} Y_1 + \lambda \nabla_{X_1} Y_2,$$

and

$$\nabla_{X_1 + \lambda X_2} Y_1 = \nabla_{X_1} Y_1 + \lambda \nabla_{X_2} Y_2.$$

ii) If $f \in C^r(\mathcal{M}; \mathbb{R})$, $X \in \mathfrak{X}_r(\mathcal{M})$, $Y \in \mathfrak{X}_{r+1}(\mathcal{M})$, then

$$\nabla_{fX} Y = f \nabla_X Y$$

iii) If $f \in C^{r+1}(\mathcal{M}; \mathbb{R})$, $X \in \mathfrak{X}_r(\mathcal{M})$, $Y \in \mathfrak{X}_{r+1}(\mathcal{M})$, then

$$\nabla_X (fY) = X[f]Y + f \nabla_X Y$$

We call $\nabla_X Y$ the covariant derivative of Y in the direction X. Suppose we are given a basis, $\{e_\mu\}$ for $\mathfrak{X}_{k-1}(\mathcal{U})$, where $\mathcal{U} \subset \mathcal{M}$ is open. Then we have the following result

Lemma 2.1. *i) An C^r-affine connection ∇ is determined on $\mathcal{U}$ by its components with respect to the basis $\{e_\mu\}$, written ${\Gamma^\mu}_{\nu\sigma} \in C^r(\mathcal{U}; \mathbb{R})$ which are defined by*

$$\nabla_{e_\nu} e_\sigma = {\Gamma^\mu}_{\nu\sigma} e_\mu$$

ii) If $\{e'_\mu\}$ is a new basis for $\mathfrak{X}_{k-1}(\mathcal{U})$, related to the old one by:

$$e_\mu = \Lambda^\nu{}_\mu e'_\nu, \qquad \Lambda^\nu{}_\mu \in C^{k-1}(\mathcal{U};\mathbb{R})$$

then the components of ∇ with respect to the old basis are related to those with respect to the new basis by

$$\Gamma^\rho{}_{\nu\mu} = \Lambda^\tau{}_\nu \Lambda^\sigma{}_\mu \Gamma'^\kappa{}_{\tau\sigma}(\Lambda^{-1})^\rho{}_\kappa + \Lambda^\tau{}_\nu e'_\tau \left[\Lambda^\kappa{}_\mu\right](\Lambda^{-1})^\rho{}_\kappa$$

where $(\Lambda^{-1})^\rho{}_\kappa \Lambda^\kappa{}_\tau = \delta^\rho{}_\tau$.

Proof. i) Suppose $X \in \mathfrak{X}_r(\mathcal{M})$, $Y \in \mathfrak{X}_{r+1}(\mathcal{M})$. In $\mathcal{U}$ we can uniquely write $X = X^\mu e_\mu$ and $Y = Y^\mu e_\mu$ for $X^\mu \in C^r(\mathcal{U};\mathbb{R})$, $Y^\mu \in C^{r+1}(\mathcal{U};\mathbb{R})$. Then in $\mathcal{U}$

$$\begin{aligned} \nabla_X Y &= \nabla_{X^\nu e_\nu}\left(Y^\sigma e_\sigma\right) \\ &= X^\nu \nabla_{e_\mu}\left(Y^\sigma e_\sigma\right) \\ &= X^\nu e_\nu(Y^\sigma) e_\sigma + X^\nu Y^\sigma \nabla_{e_\nu} e_\sigma \\ &= \left[X^\nu e_\nu(Y^\mu) + X^\nu Y^\sigma \Gamma^\mu{}_{\nu\sigma}\right] e_\mu \end{aligned}$$

which uniquely determines $\nabla_X Y$.

ii) We calculate

$$\begin{aligned} \nabla_{e_\nu} e_\mu &= \nabla_{\Lambda^\tau{}_\nu e'_\tau}\left(\Lambda^\sigma{}_\mu e'_\sigma\right) \\ &= \Lambda^\tau{}_\nu \Lambda^\sigma{}_\mu \nabla_{e'_\tau} e'_\sigma + \Lambda^\tau{}_\nu e'_\tau \left[\Lambda^\sigma{}_\mu\right] e'_\sigma \\ &= \left(\Lambda^\tau{}_\nu \Lambda^\sigma{}_\mu \Gamma'^\kappa{}_{\tau\sigma} + \Lambda^\tau{}_\nu e'_\tau \left[\Lambda^\kappa{}_\mu\right]\right) e'_\kappa \\ &= \left(\Lambda^\tau{}_\nu \Lambda^\sigma{}_\mu \Gamma'^\kappa{}_{\tau\sigma}(\Lambda^{-1})^\rho{}_\kappa + \Lambda^\tau{}_\nu e'_\tau \left[\Lambda^\kappa{}_\mu\right](\Lambda^{-1})^\rho{}_\kappa\right) e_\rho \end{aligned}$$

where in the last line we use the fact that $e'_\kappa = (\Lambda^{-1})^\rho{}_\kappa e_\rho$.

□

Notice that the second part of the Lemma justifies our restriction $r \leq k-2$, since this is the best regularity for the functions $\Gamma^\cdot{}_{\cdot\cdot}$ that is consistent with the regularity of the underlying manifold. It is also important to note that the components $\Gamma^\mu{}_{\nu\sigma}$ do *not* transform as the components of some $(1,2)-$tensor field on the manifold.

A connection allows us to transport a vector along a curve. In order to do this, we require the following result:

Lemma 2.2. *Let $\gamma : [0,1] \to \mathcal{M}$ be a C^2-curve with $\gamma(0) = p$, $\gamma(1) = q$. Suppose that ∇ is a C^r-affine connection, with $r \geq 1$. Given $X_p \in T_p\mathcal{M}$, there is a unique vector field X defined along γ such that*

$$\nabla_{\dot\gamma(t)} X = 0, \qquad X|_{t=0} = X_p. \tag{2.3}$$

We say that X is the parallel transport of X_p along γ.

Proof. Let us suppose that $\gamma([0,1]) \subset \mathcal{U}$, for some open set $\mathcal{U}$ on which we can pick a basis $\{e_\mu\}$ for $\mathfrak{X}_{k-1}(\mathcal{U})$. If this is not the case, we can split γ into a finite number of curves for which this is true. If we write $X = X^\mu e_\mu$ and $X_p = X^\mu_p e_\mu$, then the condition (2.3) becomes

$$\begin{aligned} 0 &= \dot\gamma(t)[X^\mu] + X^\nu \dot\gamma^\sigma \Gamma^\mu{}_{\sigma\nu}, \\ &= \dot X^\mu(t) + X^\nu(t)\dot\gamma^\sigma \Gamma^\mu{}_{\sigma\nu} \end{aligned}$$

together with the initial condition $X^\mu(0) = X^\mu_p(0)$. This is a linear ODE for the coefficients $X^\mu(t)$ with coefficients in C^1, hence a unique solution exists for $t \in [0,1]$. □

Definition 13. A *geodesic* or *auto-parallel curve* with respect to the connection ∇ is a C^2-curve $\gamma : (a,b) \to \mathcal{M}$ satisfying

$$\nabla_{\dot\gamma(t)}\dot\gamma(t) = \lambda(t)\dot\gamma(t)$$

for some $\lambda : (a,b) \to \mathbb{R}$. By a change of parameterisation it is possible to arrange that $\lambda = 0$, in this case, the parameterisation is called an *affine parameterisation.*

In a local coordinate chart, for an affinely parameterised geodesic we can write $\varphi \circ \gamma(t) = (x^\mu(t)) \in \mathbb{R}^n$, so that $\dot\gamma(t) = \dot x^\mu \frac{\partial}{\partial x^\mu}$. Inserting this into the expression for a parallel transported geodesic, we find that the functions $x^\mu(t)$ satisfy:

$$\ddot x^\mu + \dot x^\nu \dot x^\sigma \Gamma^\mu{}_{\nu\sigma} = 0. \tag{2.4}$$

where $\Gamma^\mu{}_{\nu\sigma}$ are the components of the connection with respect to the coordinate induced basis. From basic ODE theory (Picard-Lindelöf theorem), we have:

Lemma 2.3. *Suppose $\mathcal{M}$ is a C^k-manifold, where $k > 2$. Suppose $\mathcal{M}$ is equipped with a C^2-affine connection ∇. Given $p \in \mathcal{M}$, $X_p \in T_p\mathcal{M}$, there exists an $\epsilon > 0$ and an affinely parameterised geodesic curve $\gamma : (-\epsilon, \epsilon) \to \mathcal{M}$ such that*

$$\gamma(0) = p, \qquad \dot\gamma(0) = X_p.$$

Moreover, γ is unique up to extension: If $\gamma' : (-\epsilon', \epsilon') \to \mathcal{M}$ is another such affinely parameterised geodesic where (without loss of generality) $\epsilon' > \epsilon$, then $\gamma'(t) = \gamma(t)$ for $t \in (-\epsilon, \epsilon)$.

Note: we only have local existence in t because the ODE is not linear (the $\Gamma^\mu{}_{\nu\sigma}$ depend on $x(t)$ implicitly).

Exercise 2.4. Consider $\mathcal{M} = \mathbb{R}^3$, with a choice of orthonormal basis $\{\boldsymbol{e}_i\}$ with respect to the Euclidean metric $g_{ij} = \delta_{ij}$. Define a smooth connection on vectors by

$${}^{(\tau)}\nabla_{\boldsymbol{e}_i}\boldsymbol{e}_j = \tau\epsilon^k{}_{ij}\boldsymbol{e}_k$$

where ϵ_{ijk} is totally anti-symmetric with $\epsilon_{123} = 1$ and $\tau \in \mathbb{R}$ is a constant. Consider the curve $\gamma : (-1,1) \to \mathbb{R}^3$ given by $\gamma(t) = t\boldsymbol{e}_3$. Show that the vector fields

$$\begin{aligned} \boldsymbol{X}_1 &= \boldsymbol{e}_1\cos(\tau x^3) - \boldsymbol{e}_2\sin(\tau x^3) \\ \boldsymbol{X}_2 &= \boldsymbol{e}_1\sin(\tau x^3) + \boldsymbol{e}_2\cos(\tau x^3) \\ \boldsymbol{X}_3 &= \boldsymbol{e}_3 \end{aligned}$$

are all parallely transported by ${}^{(\tau)}\nabla$ along γ.

An important property of an affine connection is the *torsion*. This is an anti-symmetric C^r−regular $(1,2)$−tensor field defined by

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

where we recall that the commutator of two vector fields is a vector field which acts on functions as:

$$[X,Y](f) = X(Yf) - Y(Xf).$$

If $T(X,Y)$ vanishes, we say that the affine connection ∇ is symmetric.

Exercise 2.5. a) Suppose that $f \in C^1(\mathcal{M};\mathbb{R})$, show that

$$T(fX,Y) = fT(X,Y), \qquad T(X,Y) = -T(Y,X).$$

Deduce that if $\{e_\mu\}$ is locally a basis with $X = X^\mu e_\mu$, $Y = Y^\mu e_\mu$ then:

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

for some C^r−functions $T^\sigma{}_{\mu\nu} := e^\sigma\left[T(e_\mu, e_\nu)\right]$.

b) Show that for the connection defined in Exercise 2.4, the torsion is given by:

$$T^i{}_{jk} = 2\tau\epsilon^i{}_{jk}.$$

Given a connection ∇, we can define a connection on one-forms. Given $X \in \mathfrak{X}_r(\mathcal{M})$ and $\omega \in \mathfrak{X}^*_{r+1}(\mathcal{M})$, we define $\nabla_X\omega$ by requiring:

$$(\nabla_X\omega)[Y] = X(\omega[Y]) - \omega[\nabla_X Y]\,. \tag{2.5}$$

for any $Y \in \mathfrak{X}_{k-1}(\mathcal{M})$.

Exercise 2.6. Show that if $f \in C^{k-1}(\mathcal{M})$, then (2.5) implies

$$(\nabla_X\omega)[fY] = f(\nabla_X\omega)[Y]$$

for any $Y \in \mathfrak{X}_{k-1}(\mathcal{M})$. Deduce that $\nabla_X\omega \in \mathfrak{X}^*_r(\mathcal{M})$.

We can locally express the covariant derivative of a one-form using the same component functions $\Gamma^\mu{}_{\nu\sigma}$ as we used to express the covariant derivative of a vector:

Lemma 2.4. *Suppose that $\{e_\mu\}$ is a basis for $\mathfrak{X}_{k-1}(\mathcal{U})$, where $\mathcal{U} \subset \mathcal{M}$ is open, and that $\{e^\mu\}$ is a dual basis. Then in $\mathcal{U}$ we have:*

$$\nabla_{e_\nu} e^\mu = -\Gamma^\mu{}_{\nu\sigma} e^\sigma$$

where $\Gamma^\mu{}_{\nu\sigma}$ are defined by

$$\nabla_{e_\nu} e_\sigma = \Gamma^\mu{}_{\nu\sigma} e_\mu.$$

Proof. Consider

$$\begin{aligned}(\nabla_{e_\nu} e^\mu)[e_\tau] &= e_\nu (e^\mu[e_\tau]) - e^\mu [\nabla_{e_\nu} e_\tau] \\ &= e_\nu (\delta^\mu{}_\tau) - e^\mu [\Gamma^\kappa{}_{\nu\tau} e_\kappa] \\ &= 0 - \Gamma^\kappa{}_{\nu\tau} e^\mu [e_\kappa] \\ &= -\Gamma^\mu{}_{\nu\tau}\end{aligned}$$

Since a one-form is completely determined by it's action on a basis of vector fields, this suffices to show the result. □

Finally, we can extend the connection to act on arbitrary tensor fields by requiring that the Leibniz rule applies to tensor products. For example:

$$\nabla_X (Y \otimes \omega) = (\nabla_X Y) \otimes \omega + Y \otimes (\nabla_X \omega),$$

with obvious generalisations to arbitrary tensor products of vector fields and one-forms.

Lemma 2.5. *Suppose S is a C^{r+1} regular (p,q) tensor field, which is written with respect to a local basis $e_\mu \in \mathfrak{X}_{k-1}(\mathcal{U})$ of vector fields as*

$$S = S^{\mu_1,\dots\mu_p}{}_{\nu_1\dots\nu_q} e_{\mu_1} \otimes \cdots \otimes e_{\mu_p} \otimes e^{\nu_1} \otimes \cdots \otimes e^{\nu_q} \tag{2.6}$$

Where $S^{\mu_1,\dots\mu_p}{}_{\nu_1\dots\nu_q} \in C^{r+1}(\mathcal{U};\mathbb{R})$. Then $\nabla_X S$ is a C^r regular (p,q) tensor field, with components given in the local basis by:

$$\begin{aligned}(\nabla_X S)^{\mu_1,\dots\mu_p}{}_{\nu_1\dots\nu_q} &= X^\sigma e_\sigma \left[S^{\mu_1\dots\mu_p}{}_{\nu_1,\dots\nu_q}\right] \\ &+ X^\sigma \Gamma^{\mu_1}{}_{\sigma\tau} S^{\tau\mu_2,\dots\mu_p}{}_{\nu_1,\dots\nu_q} + \ldots + X^\sigma \Gamma^{\mu_j}{}_{\sigma\tau} S^{\mu_1\dots,\mu_{j-1}\tau\mu_{j+1}\dots\mu_p}{}_{\nu_1,\dots\nu_q} + \ldots \\ &- X^\sigma \Gamma^\tau{}_{\sigma\nu_1} S^{\mu_1\mu_2\dots\mu_p}{}_{\tau\dots\nu_q} - \ldots - X^\sigma \Gamma^\tau{}_{\sigma\nu_j} S^{\mu_1\dots\mu_p}{}_{\nu_1\dots\nu_{j-1}\tau\nu_{j+1}\dots\nu_q} - \ldots \\ &=: X^\sigma \nabla_\sigma S^{\mu_1,\dots\mu_p}{}_{\nu_1\dots\nu_q}\end{aligned}$$

We define ∇S to be the $(p,q+1)-$tensor field with components $\nabla_\sigma S^{\mu_1,\dots\mu_p}{}_{\nu_1\dots\nu_q}$.

Proof. We simply apply the definition of the covariant derivative to (2.6) and then express the right hand side in terms of the basis vectors. □

2.2.1 The Levi-Civita connection

There are many connections on a given manifold. We want to single out a single connection, and it turns out that this is possible when we require the connection to be compatible with our metric and symmetric.

Theorem 2.1 (Fundamental theorem of (pseudo-)Riemannian geometry)**.** *Suppose $\mathcal{M}$ is a $C^{k+1}-$regular pseudo-Riemannian manifold which carries a C^k-regular metric g, where $k \geq 1$. Then there exists a unique symmetric $C^{k-1}-$affine connection, ∇, such that $\nabla g = 0$. This connection is known as the* Levi-Civita connection.

Proof. We prove the statement by an explicit construction. Suppose that $X, Y, Z \in \mathfrak{X}_k(\mathcal{M})$. First note that the symmetry of the connection implies

$$\nabla_X Y - \nabla_Y X = [X, Y] \tag{2.7}$$

Suppose that we have a symmetric connection satisfying $\nabla_X g = 0$, then we have

$$\begin{aligned} X[g(Y,Z)] &= \nabla_X (g(Y,X)) \\ &= (\nabla_X g)(X,Y) + g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \\ &= g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \end{aligned} \tag{2.8}$$

similarly

$$Y[g(Z,X)] = g(\nabla_Y Z, X) + g(Z, \nabla_Y X) \tag{2.9}$$

$$Z[g(X,Y)] = g(\nabla_Z X, Y) + g(X, \nabla_Z Y) \tag{2.10}$$

Taking (2.8)+(2.9)−(2.10), and using (2.7) we have:

$$\begin{aligned} X[g(Y,Z)] + Y[g(Z,X)] - Z[g(X,Y)] = 2g(\nabla_X Y, Z) - g([X,Y], Z) \\ + g([Y,Z], X) - g([Z,X], Y) \end{aligned}$$

so that, after re-arranging we have

$$\begin{aligned} 2g(\nabla_X Y, Z) = X[g(Y,Z)] + Y[g(Z,X)] - Z[g(X,Y)] \\ + g([X,Y], Z) + g([Z,Y], X) + g([Z,X], Y) \end{aligned} \tag{2.11}$$

We can calculate the right hand side explicitly in terms of g, X, Y, without needing to know ∇. Since Z is arbitrary, this uniquely determines $\nabla_X Y$ by the non-degeneracy of g. Thus if the connection exists, it is unique.

On the other hand, we can use (2.11) to define a connection. It is straightforward to check that if we do this, that the connection is symmetric and $\nabla g = 0$. □

Exercise 2.7. Consider the connection defined in Exercise 2.4. Show that if we define a Riemannian metric on $\mathbb{R}^3$ by $g(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{v} \cdot \boldsymbol{w} = v^i w^j \delta_{ij}$, then the connection ${}^{(\tau)}\nabla$ satisfies

$$\boldsymbol{x}[g(\boldsymbol{y}, \boldsymbol{z})] = g\left({}^{(\tau)}\nabla_{\boldsymbol{x}} \boldsymbol{y}, \boldsymbol{z}\right) + g\left(\boldsymbol{y}, {}^{(\tau)}\nabla_{\boldsymbol{x}} \boldsymbol{z}\right).$$

Deduce that ${}^{(0)}\nabla$ is the Levi-Civita connection of g.

Suppose that we have a local basis of vector fields $\{e_\mu\}$. We know that for each μ, ν, the commutator $[e_\mu, e_\nu]$ is a vector field, so it can be written in terms of the basis vector fields:

$$[e_\mu, e_\nu] = C^\sigma{}_{\mu\nu} e_\sigma,$$

where $C^\sigma{}_{\mu\nu} \in C^{k-1}(\mathcal{M}; \mathbb{R})$ are uniquely determined. Now, let us consider (2.11) applied to the vector fields $X = e_\mu, Y = e_\nu, Z = e_\sigma$, and recall that $g_{\mu\nu} := g(e_\mu, e_\nu)$. We find:

$$\begin{aligned} 2\Gamma^\tau{}_{\mu\nu} g_{\tau\sigma} = e_\mu(g_{\nu\sigma}) + e_\nu(g_{\sigma\mu}) - e_\sigma(g_{\mu,\nu}) \\ + C^\tau{}_{\mu\nu} g_{\tau\sigma} + C^\tau{}_{\sigma\mu} g_{\tau\nu} + C^\tau{}_{\sigma\nu} g_{\tau\mu} \end{aligned}$$

so that multiplying by $g^{\sigma\kappa}$ (the co-metric) we deduce:

$$\Gamma^{\kappa}{}_{\mu\nu} = \frac{1}{2} g^{\sigma\kappa} \Big[e_\mu(g_{\nu\sigma}) + e_\nu(g_{\sigma\mu}) - e_\sigma(g_{\mu\nu}) + C_{\sigma\mu\nu} + C_{\nu\sigma\mu} + C_{\mu\sigma\nu} \Big]$$

This expression simplifies in two useful cases:

1. If $\{e_\mu\}$ is a local coordinate basis, so that $e_\mu = \frac{\partial}{\partial x^\mu}$, then we have $C^{\mu}{}_{\nu\mu} \equiv 0$ and:

$$\Gamma^{\kappa}{}_{\mu\nu} = \frac{1}{2} g^{\sigma\kappa} \left(\frac{\partial g_{\nu\sigma}}{\partial x^\mu} + \frac{\partial g_{\sigma\mu}}{\partial x^\nu} - \frac{\partial g_{\mu\nu}}{\partial x^\sigma} \right) \tag{2.12}$$

2. If $\{e_\mu\}$ is an orthonormal basis, so that $g_{\mu\nu}$ is a set of constants equal to ± 1 or 0. In this case $e_\mu(g_{\nu\sigma}) \equiv 0$ and we have

$$\Gamma^{\kappa}{}_{\mu\nu} = \frac{1}{2} g^{\sigma\kappa} \left(C_{\sigma\mu\nu} + C_{\nu\sigma\mu} + C_{\mu\sigma\nu} \right)$$

Example 9. Recall the Schwarzschild metric in Painlevé-Gullstrand coordinates (Example 7) is a Lorentzian metric on $\mathcal{M} = \mathbb{R} \times (0, \infty) \times S^2$ given by:

$$g = -\left(1 - \frac{2m}{r}\right) dt^2 + 2\sqrt{\frac{2m}{r}} dt dr + dr^2 + r^2 g_{S^2}.$$

Recall also that we have an orthonormal basis for this matrix given locally by:

$$\begin{aligned} e_0 &= \frac{\partial}{\partial t} - \sqrt{\frac{2m}{r}} \frac{\partial}{\partial r} \\ e_1 &= \frac{\partial}{\partial r} \\ e_A &= \frac{1}{r} b_A, \end{aligned}$$

where $\{b_A\}_{A=2,3}$ is a local orthonormal basis for S^2 (the unit round sphere), such that $[b_A, b_B] = \underline{C}^{E}{}_{AB} b_E$ for some functions $\underline{C}^{E}{}_{AB} \in C^\infty(S^2; \mathbb{R})$. We calculate:

$$\begin{aligned} [e_0, e_1] &= -\frac{1}{2r} \sqrt{\frac{2m}{r}} \frac{\partial}{\partial r}, \\ [e_0, e_A] &= \frac{1}{r^2} \sqrt{\frac{2m}{r}} b_A, \\ [e_1, e_A] &= -\frac{1}{r^2} b_A, \\ [e_A, e_B] &= \frac{1}{r^2} \underline{C}^{E}{}_{AB} b_E. \end{aligned}$$

From here, we deduce that:

$$
\begin{aligned}
C^1{}_{01} &= -C^1{}_{10} = -\frac{1}{2r}\sqrt{\frac{2m}{r}}, \\
C^A{}_{0B} &= -C^A{}_{B0} = \frac{1}{r}\sqrt{\frac{2m}{r}}\delta^A{}_B, \\
C^A{}_{1B} &= -C^A{}_{B1} = -\frac{1}{r}\delta^A{}_B, \\
C^E{}_{AB} &= -C^E{}_{BA} = \frac{1}{r}\underline{C}^E{}_{AB}.
\end{aligned}
$$

and all other components of $C^\mu{}_{\nu\sigma}$ vanish. Since $\{e_\mu\}$ is an orthonormal basis, we can raise and lower indices as in the Minkowski case: we change sign when raising or lowering a 0 index, but not otherwise. We calculate (for example):

$$
\begin{aligned}
\Gamma^1{}_{10} &= \frac{1}{2}\left(C^1{}_{10} - C^0{}_{11} + C^1{}_{10}\right) \\
&= \frac{1}{2r}\sqrt{\frac{2m}{r}}.
\end{aligned}
$$

Similar calculations lead us to conclude that the non-vanishing components of Γ with respect to the basis $\{e_\mu\}$ are:

$$
\begin{aligned}
\Gamma^1{}_{10} &= \frac{1}{2r}\sqrt{\frac{2m}{r}}, & \Gamma^0{}_{11} &= \frac{1}{2r}\sqrt{\frac{2m}{r}}, & \Gamma^A{}_{B0} &= -\frac{1}{r}\sqrt{\frac{2m}{r}}\delta^A{}_B, \\
\Gamma^A{}_{B1} &= \frac{1}{r}\delta^A{}_B, & \Gamma^0{}_{AB} &= -\frac{1}{r}\sqrt{\frac{2m}{r}}\delta_{AB}, & \Gamma^1{}_{AB} &= \frac{1}{r}\delta_{AB}, \\
& & \Gamma^C{}_{AB} &= \frac{1}{r}\underline{\Gamma}^C{}_{AB}, & &
\end{aligned}
$$

where $\underline{\Gamma}^C{}_{AB}$ are the connection coefficients of the Levi-Civita connection of the unit metric on S^2 with respect to the basis $\{b_A\}$.

We can use these coefficients to directly give the action of the Levi-Civita connection of g on the basis $\{e_\mu\}$:

$$
\begin{aligned}
\nabla_{e_0} e_0 &= 0, & \nabla_{e_1} e_0 &= \frac{1}{2r}\sqrt{\frac{2m}{r}}e_1, & \nabla_{e_A} e_0 &= -\frac{1}{r}\sqrt{\frac{2m}{r}}e_A, \\
\nabla_{e_0} e_1 &= 0, & \nabla_{e_1} e_1 &= \frac{1}{2r}\sqrt{\frac{2m}{r}}e_0, & \nabla_{e_A} e_1 &= \frac{1}{r}e_A,
\end{aligned} \tag{2.13}
$$

and

$$
\nabla_{e_0} e_A = 0, \qquad \nabla_{e_1} e_A = 0, \qquad \nabla_{e_A} e_B = -\delta_{AB}\left(\frac{1}{r}\sqrt{\frac{2m}{r}}e_0 + \frac{1}{r}e_1\right) + \frac{1}{r}\underline{\Gamma}^C{}_{AB}e_C.
$$

Exercise(∗). Repeat for yourself the calculations of the previous example.

Definition 14. A *geodesic* of the metric g, is a geodesic of the Levi-Civita connection of g, i.e. a curve $\gamma : (-\epsilon, \epsilon) \to \mathcal{M}$ satisfying

$$\nabla_{\dot{\gamma}(t)}\dot{\gamma}(t) = \lambda(t)\dot{\gamma}(t)$$

For a suitable choice of parameterisation, the function $\lambda(t)$ may be chosen to vanish, in which case we talk of an affinely parameterised geodesic.

The geodesic equations can also be derived from a variational principle. In Riemannian geometry this is especially useful as geodesics can be shown to locally minimise the distance between two points. For Lorentzian geometries, the geodesics typically extremise the length, but do not necessarily minimise it.

Exercise(∗) (If you have studied Lagrangian mechanics)**.** Suppose we work in a coordinate chart $\mathcal{U} \subset \mathcal{M}$ in which the metric takes the form:

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

Show that the affinely parameterised geodesic equations

$$\ddot{x}^\mu + \dot{x}^\nu\dot{x}^\sigma\Gamma^\mu{}_{\nu\sigma} = 0,$$

where

$$\Gamma^\kappa{}_{\mu\nu} = \frac{1}{2}g^{\sigma\kappa}\left(\frac{\partial g_{\nu\sigma}}{\partial x^\mu} + \frac{\partial g_{\sigma\mu}}{\partial x^\nu} - \frac{\partial g_{\mu\nu}}{\partial x^\sigma}\right)$$

can be derived from the Lagrangian:

$$L = \frac{1}{2}g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu.$$

Exercise 3.1. Suppose that S is a $(1,1)-$tensor and R is a $(0,2)-$tensor, both at least C^1-regular, which with respect to a local basis $\{e_\mu\}$ may be written:

$$S = S^\mu{}_\nu e_\mu \otimes e^\nu, \qquad R = R_{\sigma\tau}e^\sigma \otimes e^\tau.$$

Suppose ∇ is the Levi-Civita connection.

i) With the notation of Lemma 2.5, show that:

$$\nabla_\kappa\left(S^\mu{}_\nu R_{\sigma\tau}\right) = \left(\nabla_\kappa S^\mu{}_\nu\right)R_{\sigma\tau} + S^\mu{}_\nu\left(\nabla_\kappa R_{\sigma\tau}\right).$$

ii) how also that

$$\nabla_\kappa\delta^\mu{}_\sigma = 0, \qquad \nabla_\kappa g_{\mu\nu} = 0, \qquad \nabla_\kappa g^{\mu\nu}.$$

Conclude that ∇, thought of as a map from $(p,q)-$tensors to $(p,q+1)-$tensors commutes with contractions and raising and lowering of indices.

2.2.2 The wave equation

Suppose that $\mathcal{M}$ is a C^k-manifold with a $C^{k-1}-$regular Riemannian or Lorentzian metric g, which has an associated $C^{k-2}-$regular Levi-Civita connection ∇. Given $f \in C^k(\mathcal{M};\mathbb{R})$, and $X, Y \in \mathfrak{X}_{k-1}(\mathcal{M})$ we define:

$$H_f(X,Y) := X(Yf) - (\nabla_X Y)\, f.$$

Lemma 2.6. *H_f is a symmetric, $C^{k-2}-$regular $(0,2)-$tensor.*

Proof. To show that H_f is symmetric, we calculate:

$$\begin{aligned} H_f(X,Y) - H_f(Y,X) &= X(Yf) - Y(Xf) - [(\nabla_X Y) - (\nabla_Y X)]\, f \\ &= [X,Y]f - [(\nabla_X Y) - (\nabla_Y X)]\, f \\ &= T(X,Y)f = 0. \end{aligned}$$

To show that H_f is a tensor, we need to show that for any $g \in C^{k-1}(\mathcal{M})$ we have:

$$H_f(gX,Y) = gH_f(X,Y) = H_f(X,gY).$$

By the symmetry we can just show the first equality, which follows from

$$H_f(gX,Y) = (gX)[Yf] - (g\nabla_X Y)f = gH_f(X,Y).$$

The regularity of H_f follows from the assumptions on the connection and on f. □

The tensor H_f is called the Hessian of f and is sometimes written $\nabla\nabla f$, so that if $\{e_\mu\}$ is a local basis, we write:

$$\nabla_\mu \nabla_\nu f = H_f(e_\mu, e_\nu).$$

Note carefully that in general

$$\nabla_\mu \nabla_\nu f \neq \nabla_{e_\mu} \nabla_{e_\nu} f = e_\mu(e_\nu f).$$

Exercise 3.2. In a local coordinate basis, $\{e_\mu = \frac{\partial}{\partial x^\mu}\}$, show that we can write

$$H_f(X,Y) = X^\mu Y^\nu \left(\frac{\partial^2 f}{\partial x^\mu \partial x^\nu} - \Gamma^\sigma{}_{\mu\nu} \frac{\partial f}{\partial x^\sigma} \right),$$

Deduce that for a local coordinate basis

$$\nabla_\mu \nabla_\nu f = \frac{\partial^2 f}{\partial x^\mu \partial x^\nu} - \Gamma^\sigma{}_{\mu\nu} \frac{\partial f}{\partial x^\sigma}$$

Definition 15. Suppose $f \in C^k(\mathcal{M};\mathbb{R})$.

i) If g is *Lorentzian*, we define the wave operator $\Box_g$ acting on f, written $\Box_g f \in C^{k-2}(\mathcal{M};\mathbb{R})$ is defined to be the trace of H_f with respect to the metric g. With respect to a local coordinate basis $\{\partial/\partial x^\mu\}$, this is given by:

$$\Box_g f := g^{\mu\nu} \nabla_\mu \nabla_\nu f = \nabla^\mu \nabla_\mu f.$$

ii) If g is *Riemannian*, we define the Laplace operator Δ_g acting on f, written $\Delta_g f \in C^{k-2}(\mathcal{M};\mathbb{R})$ to be the trace of H_f with respect to the metric g. With respect to a local coordinate basis $\{\partial/\partial x^i\}$, this is given by:

$$\Delta_g f := g^{ij}\nabla_i\nabla_i f = \nabla^i\nabla_i f.$$

In a local coordinate basis, we have

$$\Box_g f = g^{\mu\nu}\frac{\partial^2 f}{\partial x^\mu \partial x^\nu} - g^{\mu\nu}\Gamma^\sigma{}_{\mu\nu}\frac{\partial f}{\partial x^\sigma}, \tag{2.14}$$

with a similar expression for the Laplace operator:

$$\Delta_g f = g^{ij}\frac{\partial^2 f}{\partial x^i \partial x^j} - g^{ij}\Gamma^k{}_{ij}\frac{\partial f}{\partial x^k}, \tag{2.15}$$

When g is *Lorentzian*, we recognise $\Box_g$ as a hyperbolic differential operator. In the case where g is Riemannian, the operator Δ_g is rather an elliptic differential operator.

It's often useful to have a more concrete expression for this operator. We can find one by recalling that the Levi-Civita connection has components that are given in a local coordinate basis by (2.12):

$$\Gamma^\kappa{}_{\mu\nu} = \frac{1}{2}g^{\sigma\kappa}\left(\frac{\partial g_{\nu\sigma}}{\partial x^\mu} + \frac{\partial g_{\sigma\mu}}{\partial x^\nu} - \frac{\partial g_{\mu\nu}}{\partial x^\sigma}\right)$$

Contracting with the co-metric, we have:

$$g^{\mu\nu}\Gamma^\kappa{}_{\mu\nu} = g^{\sigma\kappa}\frac{\partial g_{\nu\sigma}}{\partial x^\mu}g^{\mu\nu} - \frac{1}{2}g^{\sigma\kappa}\frac{\partial g_{\mu\nu}}{\partial x^\sigma}g^{\mu\nu}$$

Now, recall the following facts regarding differentiating matrices:

$$\frac{dA^{-1}}{dt}(t) = -A^{-1}\frac{dA}{dt}(t)A^{-1}(t).$$

$$\frac{d}{dt}\det A(t) = [\det A(t)]\,\mathrm{Tr}\left(\frac{dA}{dt}(t)A^{-1}(t)\right).$$

We recognise that we can write

$$g^{\mu\nu}\Gamma^\kappa{}_{\mu\nu} = -\frac{\partial}{\partial x^\mu}g^{\mu\kappa} - \frac{g^{\sigma\kappa}}{\sqrt{|g|}}\frac{\partial}{\partial x^\sigma}\sqrt{|g|}$$

where we have defined $|g| := |\det(g_{\mu\nu})|$. Returning to (2.14), we deduce that in local coordinates:

$$\Box_g f = \frac{1}{\sqrt{|g|}}\frac{\partial}{\partial x^\mu}\left(\sqrt{|g|}g^{\mu\nu}\frac{\partial f}{\partial x^\nu}\right).$$

A similar expression of course holds in the case that g is Riemannian:

$$\Delta_g f = \frac{1}{\sqrt{|g|}}\frac{\partial}{\partial x^i}\left(\sqrt{|g|}g^{ij}\frac{\partial f}{\partial x^j}\right).$$

An alternative way to define the Hessian and wave/Laplace operator is instead to recall that given a function $f \in C^r(\mathcal{M};\mathbb{R})$, there is a natural one-form $df \in \mathfrak{X}^*_{r-1}(\mathcal{M})$ defined by:

$$df(X) = Xf,$$

for any $X \in \mathfrak{X}(\mathcal{M})$. With respect to a local basis:

$$df = (e_\mu f)e^\mu = (\nabla_\mu f)e^\mu.$$

Since we have a metric g we can associate df with a vector field, sometimes called $\text{grad}_g f$, which acts on $h \in C^1(\mathcal{M};\mathbb{R})$ by:

$$[\text{grad}_g f](h) = g^{-1}(df, dh).$$

In a local basis:

$$\text{grad}_g f = g^{\mu\nu}(df)_\nu e_\mu = g^{\mu\nu}(e_\nu f)e_\mu$$

To define the wave/Laplace operator, we introduce the divergence of a vector field as follows. If $V \in \mathfrak{X}_r(\mathcal{M})$, then the divergence of V, written $\text{div}_g V \in C^{r-1}(\mathcal{M};\mathbb{R})$ is given by:

$$\text{div}_g V = e^\mu \left[\nabla_{e_\mu} V\right] = \nabla_\mu V^\mu.$$

Clearly, we have

$$\Box_g f = \text{div}_g \left(\text{grad}_g f\right).$$

Exercise 3.3. For $V \in \mathfrak{X}_r(\mathcal{M})$, the divergence of V, written $\text{div}_g V \in C^{r-1}(\mathcal{M};\mathbb{R})$ is defined with respect to a local basis by:

$$\text{div}_g V := \nabla_\mu V^\mu = e^\mu \left(\nabla_{e_\mu} V\right).$$

Show that with respect to a local coordinate basis we have:

$$\text{div}_g V = \frac{1}{\sqrt{|g}} \frac{\partial}{\partial x^\mu}\left(\sqrt{|g|} V^\mu\right).$$

Deduce the expression for the wave/Laplace operator with respect to local coordinates, using:

$$\Box_g f = \text{div}_g \left(\text{grad}_g f\right).$$

We can in fact generalise the definition of the wave/Laplace operator to arbitrary tensors. For example, if S is a C^2-regular $(1,1)-$tensor, we can define (in a local basis):

$$\Box_g S = \left(\nabla_\mu \nabla^\mu S^\nu{}_\tau\right) e_\nu \otimes e^\tau.$$

Example 10. Consider the sphere S^2, with the stereographic coordinate charts $(\mathcal{U}^\pm, \varphi_\pm)$ that we previously defined. Recall that the metric on each patch is given by:

$$g_{S^2} = \Omega^2 \delta_{ij} dx^i dx^j.$$

Where

$$\Omega = \frac{2}{1 + |\boldsymbol{x}|^2}$$

As a matrix, the components g_{ij} are

$$[g_{ij}] = \Omega^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

so that $|g| = \Omega^4$ and

$$[g^{ij}] = \Omega^{-2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

so that $\sqrt{|g|}g^{ij} = \delta^{ij}$. As a consequence, we have

$$\Delta_{S^2} f = \Omega^{-2} \Delta_{\mathbb{R}^2} f$$

where

$$\Delta_{\mathbb{R}^2} f = \delta^{ij} \frac{\partial^2 f}{\partial x^i x^j}$$

is the standard Laplacian on $\mathbb{R}^2$.

Consider the following integral:

$$I := \int_{|\boldsymbol{x}|<R} f \Delta_{S^2} f \sqrt{|g|} dx = \int_{|\boldsymbol{x}|<R} f \Delta_{\mathbb{R}^2} f \ dx.$$

Now, we can use the standard divergence theorem on $\mathbb{R}^2$ to deduce

$$I = \int_{|\boldsymbol{x}|=R} f \frac{\partial f}{\partial \nu} d\sigma - \int_{|\boldsymbol{x}|<R} \delta^{ij} \frac{\partial f}{\partial x^i} \frac{\partial f}{\partial x^j} dx.$$

I claim that if f is a smooth function on S^2, then the first term vanishes as $R \to \infty$, and we can conclude:

$$\int_{\mathbb{R}^2} f \Delta_{S^2} f \, \Omega^2 dx = - \int_{\mathbb{R}^2} \delta^{ij} \frac{\partial f}{\partial x^i} \frac{\partial f}{\partial x^j} dx.$$

We immediately conclude that a harmonic function on S^2 (i.e. a function $f \in C^2(S^2; \mathbb{R})$ satisfying $\Delta_{S^2} f = 0$) must necessarily be constant.

To check the vanishing of the boundary term, by considering the transition function between the two stereographic patches, it's simple to check that if f is a smooth function on S^2 which is given in one stereographic patch by $f(\boldsymbol{x})$, then the function $\tilde{f}(\boldsymbol{x}) = f\left(\frac{\boldsymbol{x}}{|\boldsymbol{x}|^2}\right)$ should be smooth at the origin. Consider

$$\frac{\partial \tilde{f}}{\partial x^i}(\boldsymbol{x}) = \frac{\delta^{ij} |\boldsymbol{x}|^2 - 2x^i x^j}{|\boldsymbol{x}|^4} \frac{\partial f}{\partial x^j} \left(\frac{\boldsymbol{x}}{|\boldsymbol{x}|^2} \right)$$

Setting $\boldsymbol{y} = \frac{\boldsymbol{x}}{|\boldsymbol{x}|^2}$, and noting that $\frac{\partial \tilde{f}}{\partial x^i}(\boldsymbol{x})$ is bounded as $\boldsymbol{x} \to 0$, we conclude that for large $|y|$, we have

$$\left| \frac{\partial f}{\partial x^i}(\boldsymbol{y}) \right| \leq \frac{C}{|\boldsymbol{y}^2|}$$

from which the result easily follows.

Exercise 3.4. Consider the sphere $S^2 = \{\boldsymbol{x} \in \mathbb{R}^3 : |\boldsymbol{x}| = 1\}$, and set

$$\mathcal{U} = S^2 \setminus \{(-\sqrt{1-t^2}, 0, t) : t \in [-1, 1]\},$$

i.e. the sphere with a line of longitude from north to south pole removed. We define a coordinate chart $(\mathcal{U}, \varphi)$, where φ is defined in terms of its inverse by:

$$\begin{array}{rccc} \varphi^{-1} & : & (0,\pi)\times(-\pi,\pi) & \to & \mathcal{U} \\ & & (\theta,\phi) & \mapsto & \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}. \end{array}$$

a) Show that with the standard identification of tangent vectors in S^2 with vectors in $\mathbb{R}^3$ that we have

$$\left.\frac{\partial}{\partial\theta}\right|_{\varphi^{-1}(\theta,\phi)} \simeq \boldsymbol{e}_\theta := \begin{pmatrix} \cos\theta\cos\phi \\ \cos\theta\sin\phi \\ \sin\theta \end{pmatrix}, \qquad \left.\frac{\partial}{\partial\phi}\right|_{\varphi^{-1}(\theta,\phi)} \simeq \boldsymbol{e}_\phi := \begin{pmatrix} -\sin\theta\sin\phi \\ \sin\theta\cos\phi \\ 0 \end{pmatrix}.$$

b) Show that

$$\begin{aligned} \langle \boldsymbol{e}_\theta, \boldsymbol{e}_\theta \rangle|_{\varphi^{-1}(\theta,\phi)} &= 1, \\ \langle \boldsymbol{e}_\theta, \boldsymbol{e}_\phi \rangle|_{\varphi^{-1}(\theta,\phi)} &= 0, \\ \langle \boldsymbol{e}_\phi, \boldsymbol{e}_\phi \rangle|_{\varphi^{-1}(\theta,\phi)} &= \sin^2\theta. \end{aligned}$$

where $\langle,\rangle$ is the standard inner product on $\mathbb{R}^3$.

c) Deduce that in the θ, ϕ coordinates:

$$g_{S^2} = d\theta^2 + \sin^2\theta d\phi^2.$$

d) Show that in these coordinates:

$$\Delta_{g_{S^2}} f = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 f}{\partial\phi^2}$$

e) Suppose f is a smooth function on S^2. Show that:

$$\int_{\theta=0}^{\pi}\int_{\phi=-\pi}^{\pi} f\Delta_{g_{S^2}} f \sin\theta d\theta d\phi = -\int_{\theta=0}^{\pi}\int_{\phi=-\pi}^{\pi}\left[\left(\frac{\partial f}{\partial\theta}\right)^2 + \frac{1}{\sin^2\theta}\left(\frac{\partial f}{\partial\phi}\right)^2\right]\sin\theta d\theta d\phi$$

Example 11. A more interesting example is furnished by the Painlevé-Gullstrand form of the Schwarzschild Black Hole. Recall that in a coordinate patch we have

$$g = -\left(1 - \frac{2m}{r}\right)dt^2 + 2\sqrt{\frac{2m}{r}}dtdr + dr^2 + r^2\Omega^2\delta_{ij}dx^i dx^j,$$

with Ω as above. Writing the components of the metric as a matrix, we have

$$[g_{\mu\nu}] = \begin{pmatrix} -\left(1-\frac{2m}{r}\right) & \sqrt{\frac{2m}{r}} & 0 & 0 \\ \sqrt{\frac{2m}{r}} & 1 & 0 & 0 \\ 0 & 0 & \Omega^2 & 0 \\ 0 & 0 & 0 & \Omega^2 \end{pmatrix}$$

so that $|g| = r^4\Omega^4$, and

$$[g^{\mu\nu}] = \begin{pmatrix} -1 & \sqrt{\frac{2m}{r}} & 0 & 0 \\ \sqrt{\frac{2m}{r}} & \left(1-\frac{2m}{r}\right) & 0 & 0 \\ 0 & 0 & \Omega^{-2} & 0 \\ 0 & 0 & 0 & \Omega^{-2} \end{pmatrix}$$

We find

$$\Box_g f = -\frac{\partial^2 f}{\partial t^2} + \sqrt{\frac{2m}{r}}\frac{\partial^2 f}{\partial t \partial r} + \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\sqrt{\frac{2m}{r}}\frac{\partial f}{\partial t}\right) + \frac{1}{r^2}\frac{\partial}{\partial r}\left([r^2 - 2mr]\frac{\partial f}{\partial r}\right) + \frac{1}{r^2}\Delta_{S^2} f. \tag{2.16}$$

In the same way that studying the wave equation in Minkowski space helped us to understand the causal properties (such as signal propagation) for that spacetime, solutions of the $\Box_g u = 0$ are of great interest in studying the spacetime $(\mathcal{M}, g)$. We will give a taste of results in this direction, but the study of wave equations on Lorentzian manifolds is a large field of study with many fascinating recent advances.

Theorem 2.2. *Let $(\mathcal{M}, g)$ be the Painlevé-Gullstrand form of the Schwarzschild spacetime, as in Example 7 above. Suppose that $f \in C^2(\mathcal{M}; \mathbb{R})$ vanishes for sufficiently large r at every fixed time t. Define:*

$$E_f[t] := \frac{1}{2}\int_{(2m,\infty)\times\mathbb{R}^2} \left[\left(\frac{\partial f}{\partial t}\right)^2 + \left(1-\frac{2m}{r}\right)\left(\frac{\partial f}{\partial r}\right)^2 + \left|\not\nabla f\right|^2\right] r^2\Omega^2 dr dx$$

where

$$\Omega^2 := \frac{4}{\left(1+|\boldsymbol{x}|^2\right)^2}, \qquad \not\nabla_A f = \frac{1}{r\Omega}\frac{\partial}{\partial x^A}.$$

Then we have

$$\frac{dE_f}{dt} + 4m^2\int_{\mathbb{R}^2}\left|\frac{\partial f}{\partial t}\right|^2_{r=2m}\Omega^2 dx = -\int_{(2m,\infty)\times\mathbb{R}^2}\Box_g f\frac{\partial f}{\partial t}\, r^2\Omega^2 dr dx.$$

Proof. We multiply the expression (2.16) by:

$$-\frac{\partial f}{\partial t} r^2\Omega^2$$

and integrate over $(2m, \infty) \times \mathbb{R}^2$ with the measure $drdx$. The first term gives

$$\int_{(2m,\infty)\times\mathbb{R}^2} \frac{\partial^2 f}{\partial t^2}\frac{\partial f}{\partial t} r^2 \Omega^2 drdx = \frac{d}{dt}\frac{1}{2}\int_{(2m,\infty)\times\mathbb{R}^2} \left(\frac{\partial f}{\partial t}\right)^2 r^2 \Omega^2 drdx$$

The second term we integrate by parts in r to give:

$$\begin{aligned} -\int_{(2m,\infty)\times\mathbb{R}^2} \sqrt{\frac{2m}{r}}\frac{\partial^2 f}{\partial t \partial r}\frac{\partial f}{\partial t} r^2 \Omega^2 drdx &= +\int_{(2m,\infty)\times\mathbb{R}^2} \frac{\partial}{\partial r}\left(r^2\sqrt{\frac{2m}{r}}\frac{\partial f}{\partial t}\right)\frac{\partial f}{\partial t}\Omega^2 drdx \\ &\quad + 4m^2 \int_{\mathbb{R}^2} \left.\frac{\partial f}{\partial t}\right|^2_{r=2m} \Omega^2 dx \end{aligned}$$

where we use the fact that f vanishes for sufficiently large r. This term will cancel with the other mixed derivative term, leaving only the boundary term. Next consider the term involving two radial derivatives. We have

$$\begin{aligned} -\int_{(2m,\infty)\times\mathbb{R}^2} & \frac{1}{r^2}\frac{\partial}{\partial r}\left([r^2 - 2mr]\frac{\partial f}{\partial r}\right)\frac{\partial f}{\partial t} r^2 \Omega^2 drdx \\ &= \int_{(2m,\infty)\times\mathbb{R}^2} [r^2 - 2mr]\frac{\partial^2 f}{\partial t \partial r}\frac{\partial f}{\partial r}\Omega^2 drdx \\ &= \frac{d}{dt}\frac{1}{2}\int_{(2m,\infty)\times\mathbb{R}^2} \left[1 - \frac{2m}{r}\right]\left(\frac{\partial f}{\partial r}\right)^2 r^2 \Omega^2 drdx \end{aligned}$$

Where we use the fact that f vanishes for sufficiently large r, and that the factor $r^2 - 2mr$ vanishes at $r = 2m$ to discard boundary terms. Finally, we use the result

$$\int_{\mathbb{R}^2} f \Delta_{S^2} f \, \Omega^2 dx = -\int_{\mathbb{R}^2} \delta^{ij}\frac{\partial f}{\partial x^i}\frac{\partial f}{\partial x^j} dx.$$

to integrate the last term by parts in x and we're done. □

Corollary 2.3. *Suppose $u_1, u_2 \in C^2(\mathcal{M})$ are two solutions of the wave equation on the Schwarzschild spacetime, vanishing for sufficiently large r at each fixed t, such that:*

$$u_1 = u_2, \qquad \frac{\partial}{\partial t}u_1 = \frac{\partial}{\partial t}u_2, \quad \text{on } \{0\} \times (2m, \infty) \times S^2$$

Then $u_1 = u_2$ in the region $\mathcal{R} = [0, \infty) \times (2m, \infty) \times S^2$.

Proof. We apply Theorem 2.2 to $u_1 - u_2$. We immediately conclude that $E_{u_1-u_2}[t]$ is monotone decreasing and positive, however, it is initially zero thus $E_{u_1-u_2}[t] = 0$ for $t \geq 0$. We conclude that $u_1 - u_2 \equiv 0$ in the region $\mathcal{R}$. □

In fact, this result follows from a more abstract result that says that a solution of the wave equation on a Lorentzian manifold is determined in $D^+(\Sigma)$ by its Cauchy data on Σ.

Proposition 1. *Let $(\mathcal{M}, g)$ be a smooth, orientable, time orientable, Lorentzian manifold. Suppose that $\mathcal{U} \subset \mathcal{M}$ is a open set with compact closure, whose boundary consists of two smooth compact components: Σ, Σ', which are both spacelike and such that $\mathcal{U} \subset D^+(\Sigma)$. Suppose that $\psi \in C^2(\overline{\mathcal{U}})$ solves the equation:*

$$\Box_g \psi = 0 \tag{2.17}$$

in $\mathcal{U}$. Then if

$$\psi|_\Sigma = 0, \qquad N_\Sigma \psi|_\Sigma = 0,$$

Where N_Σ is the future unit normal of Σ, we have that $\psi = 0$ in $\mathcal{U}$.

The proof of this result will mirror (and extend) the proof of Theorem 1.8, the analogous result for Minkowski space, and will require us to reintroduce some machinery from the Minkowski proof, now adapted to the manifold case.

Definition 16. Suppose $\mathcal{U} \subset \mathcal{M}$ is open. Given $\psi \in C^2(\mathcal{U})$, we define a symmetric $(0,2)-$tensor, the *energy momentum tensor* by

$$T[\psi] := d\psi \otimes d\psi - \frac{1}{2} |d\psi|_g \, g$$

or, with respect to a local basis $\{e_\mu\}$:

$$T[\psi] = \left((e_\mu \psi)(e_\nu \psi) - \frac{1}{2} g_{\mu\nu} (e_\sigma \psi)(e_\tau \psi) g^{\sigma\tau} \right) e^\mu \otimes e^\nu \tag{2.18}$$

$$= \left(\nabla_\mu \psi \nabla_\nu \psi - \frac{1}{2} g_{\mu\nu} \nabla_\sigma \psi \nabla^\sigma \psi \right) e^\mu \otimes e^\nu \tag{2.19}$$

Exercise 3.5. In the case that $(\mathcal{M}, g)$ is the Minkowski spacetime and that $\{e_\mu\}$ is an inertial frame, show that (2.18) agrees with the previous defninition of the energy momentum tensor.

Theorem 2.4. *The energy momentum tensor has the following properties:*

1. *We have a formula for the divergence:*

$$div_g T[\psi] := \nabla_\mu T^\mu{}_\nu[\psi] e^\nu = (\Box_g \psi) \, d\psi$$

2. *Suppose $V \in T_p\mathcal{M}$ is a unit timelike vector. Then at p, we have $T[\psi](V, V) \geq 0$, with equality iff $d\psi$ vanishes at p. If $\{e_\mu\}$ is any basis for $T_p\mathcal{M}$, there exists a constant $C > 0$ depending on V and $\{e_\mu\}$ such that:*

$$\frac{1}{C} \sum_\mu (e_\mu \psi)^2 \leq T[\psi](V, V) \leq C \sum_\mu (e_\mu \psi)^2 .$$

3. *If $W \in T_p\mathcal{M}$ is an unit timelike vector, then:*

$$\frac{1}{4\, |g(V, W)|} T[\psi](V, V) \leq T[\psi](V, W) \leq |g(V, W)| \, T[\psi](V, V).$$

Proof. 1. We work in a local basis. We have:

$$
\begin{aligned}
\nabla_\mu T^\mu{}_\nu &= \nabla_\mu \left(\nabla^\mu \psi \nabla_\nu \psi - \frac{1}{2} \delta^\mu{}_\nu \nabla_\sigma \psi \nabla^\sigma \psi \right) \\
&= \Box_g \psi \nabla_\nu \psi + \nabla^\mu \psi \nabla_\mu \nabla_\nu \psi - \nabla_\nu \nabla_\sigma \psi \nabla^\sigma \psi \\
&= \Box_g \psi \nabla_\nu \psi
\end{aligned}
$$

2, 3. The proofs of the weak and dominant energy conditions follow directly from the same proofs in the Minkowski case: We pick an orthonormal basis for $T_p\mathcal{M}$ in which $V = e^0$, and then follow the same calculations as before. Since we work only in the tangent space at a point, all of the same calculations are valid.

□

We shall require a version of the divergence theorem for a Lorentzian manifold:

Lemma 2.7. *Suppose $\mathcal{M}$ is a smooth orientable, time orientable, $(n+1)-$dimensional Lorentzian manifold. Suppose that $V \in \mathfrak{X}_1(\mathcal{M})$ is a vector field and that $\mathcal{U} \subset \mathcal{M}$ is an open set with compact closure, whose boundary $\Sigma = \partial\mathcal{U}$ consists piecewise of smooth embedded submanifolds and can be written as $\Sigma = \Sigma_s \cup \Sigma_t$ where Σ_s is spacelike and Σ_t is timelike. Then we have*

$$
\int_{\mathcal{U}} \mathrm{div}_g V dX = \int_{\Sigma_s} g(V, N) d\sigma - \int_{\Sigma_t} g(V, N) d\sigma,
$$

where vector N is the unit outwards *normal (with respect to g). The volume measure dX and surface measure $d\sigma$ are positive, and on each coordinate chart are equivalent to the $(n+1)-$dimensional (resp. $n-$dimensional) Lebesgue measure.*

The final result that we shall require is a statement about the causal structure of $D^+(\Sigma)$ for a spacelike Σ:

Lemma 2.8. *Suppose $\mathcal{M}$ is a smooth orientable, time orientable, $(n+1)-$dimensional Lorentzian manifold. Suppose $\Sigma \subset \mathcal{M}$ is a smooth embedded $n-$dimensional spacelike submanifold, with $D^+(\Sigma)$ non-empty. Then there exists a function $t \in C^\infty(D^+(\Sigma); \mathbb{R})$, such that:*

i) $\Sigma = \{t = 0\}$.

ii) The level sets of t are spacelike, equivalently $\mathrm{grad}_g t$ *is everywhere causal.*

iii) t increases along any future directed causal curve, equivalently $-\mathrm{grad}_g t$ *is future directed.*

Such a t is called a *time function* for $D^+(\Sigma)$. The proof that such a function exists is quite subtle and is beyond the scope of the course. It is fairly straightforward to produce a C^0-function which is increasing on any future directed causal curve, and whose level sets are achronal (i.e. no two points on a single level set can be connected by a timelike curve). Showing that this C^0−result can be improved to higher regularity is not trivial.[1]

We now have the pieces we require to prove the proposition.

Proof of Proposition 1.

1. Consider the vector field given in a local basis $\{e_\mu\}$ by:

$$^V J[\psi] = T_\mu{}^\nu[\psi]V^\mu e_\nu$$

If ψ solves the wave equation, we have

$$\mathrm{div}_g\left(^V J[\psi]\right) = T^{\mu\nu}[\psi]^V\Pi_{\mu\nu}$$

where the deformation tensor is given in a local basis by:

$$^V\Pi := \nabla_{(\mu}V_{\nu)} = \frac{1}{2}\left(\nabla_\mu V_\nu + \nabla_\nu V_\mu\right)e^\nu \otimes e^\mu.$$

2. By the assumptions of the proposition together with Lemma 2.8, we have on $D^+(\Sigma)$ a smooth time function t, and a vector X which is everywhere timelike. We define

$$V = e^{-\lambda t}X.$$

We apply the divergence theorem to $^V J[\psi]$ on the region $\mathcal{U}$ to find:

$$\int_{\mathcal{U}} \mathrm{div}_g\left(^V J[\psi]\right) dX = \int_\Sigma g\left(^V(J[\psi], N\right) d\sigma - \int_{\Sigma'} g\left(^V(J[\psi], N\right) d\sigma$$

where N is the future directed unit normal vector.

3. We calculate:

$$^V\Pi_{\mu\nu} = -\lambda e^{-\lambda t}X_{(\mu}\nabla_{\nu)}t + {}^X\Pi_{\mu\nu}e^{-\lambda t}$$

so that locally we have:

$$\mathrm{div}_g\left(^V J[\psi]\right) = -\lambda T_{\mu\nu}[\psi]X^\mu\nabla^\nu t + T^{\mu\nu}[\psi]^X\Pi_{\mu\nu}$$

4. By the fact that $-\nabla^\nu t$ is future directed and timelike, we have that $-T_{\mu\nu}[\psi]X^\mu\nabla^\nu t$ is positive, and controls all derivatives of ψ at each point by part 3 of Theorem 2.4. For sufficiently large λ, making use of the fact that $\mathcal{U}$ has compact closure, we conclude that

$$\mathrm{div}_g\left(^V J[\psi]\right) \geq e^{-\lambda t}T[\psi](X,X) \geq C^{-1}T[\psi](X,X)$$

everywhere in $\mathcal{U}$ for some large finite C, by the compactness of $\overline{\mathcal{U}}$.

[1]Those interested can find the C^0−result in the paper:
"Domain of dependence," R. Geroch, J.Math.Phys. **11** (1970) 437-439.
The smooth result is in the paper:
"On Smooth Cauchy hypersurfaces and Geroch's splitting theorem,"paper
A. Bernal, M. Sanchez Commun.Math.Phys. **243** (2003) 461-470

5. Turning to the surface terms, we have

$$g\left({}^V J[\psi], N\right) = T[\psi](V, N)$$

so again using the compactness of Σ, Σ', we have:

$$\int_\Sigma g\left({}^V J[\psi], N\right) d\sigma \leq C \int_\Sigma T[\psi](X, X) d\sigma$$

and

$$\int_{\Sigma'} g\left({}^V J[\psi], N\right) d\sigma \geq C^{-1} \int_{\Sigma'} T[\psi](X, X) d\sigma$$

6. We conclude that there exists a constant C, independent of ψ such that

$$\int_{\mathcal{U}} T[\psi](X, X) dX \leq C \int_\Sigma T[\psi](X, X) d\sigma.$$

Now recall that $T[\psi](X, X)$ vanishes at a point if and only if $d\psi$. We see that if ψ and $N_\Sigma \psi$ vanish on Σ (and hence $d\psi = 0$ on Σ), we must have $df = 0$in $\mathcal{U}$ and hence $\psi = 0$.

□

Notice that in fact we have a stronger statement than in the proposition, we in fact have a quantitative estimate for a solution ψ of the wave equation in terms of initial data. If you attended the advanced PDE course, you will recognise that we have in fact established H^1-control of ψ in the region $\mathcal{U}$ in terms of the H^1-initial data.

2.3 Curvature

In order to consider the dynamical gravitational field, we want to write down some equation satisfied by the metric. There are three main goals we have in arriving at this equation:

1. The equation should be hyperbolic, that is to say it locally has similar properties to the wave equation.
2. The equation should be geometric, it shouldn't depend on a particular choice of basis or coordinate system.
3. The gravitational field should couple to matter in a natural way.

The first condition will ensure that the features we observe in special relativity (such as finite speed of signal propagation) are not affected by the gravitational field. The second condition is a mathematical consequence of the equivalence principle. The third condition, while somewhat vague, is necessary since we expect the gravitational field to interact with matter in the spacetime.

One possible first guess at how we could achieve these goals would be to assume that the metric tensor obeys the wave equation:

$$\Box_g g = F$$

where the wave operator is defined on a symmetric 2−tensor in the obvious fashion $(\Box_g h)_{\mu\nu} = \nabla^\sigma \nabla_\sigma h_{\mu\nu}$ in a local basis, and F should represent some source term for the gravitational field. Unfortunately, it is trivial that $\Box_g g \equiv 0$, so this won't work. It turns out that the correct object to play the role of '$\Box g$' is a term involving the *curvature* of the metric. Before we can write down the Einstein equations then, we need to take a bit of time to define curvatures.

2.3.1 Riemann curvature

Suppose that $\mathcal{M}$ is a C^k-manifold, $k \geq 3$, equipped with a C^r-regular pseudo-Riemannian metric g, where $2 \leq r < k$. The associated $C^{r-1}-$regular Levi-Civita connection is ∇. Given $X, Y \in \mathfrak{X}_{k-1}(\mathcal{M})$ and a C^s-regular $(p,q)-$tensor S, with $s \leq r$, we define the $C^{s-2}-$regular (p,q) tensor $R(X,Y)S$ by:

$$R(X,Y)S := \nabla_X \nabla_Y S - \nabla_Y \nabla_X S - \nabla_{[X,Y]} S. \tag{2.20}$$

Lemma 2.9. *The tensor $R(X,Y)S$ has the following properties:*

i) It is antisymmetric in X, Y:

$$R(X,Y)S = -R(Y,X)S$$

for any $X, Y \in \mathfrak{X}_{k-1}(\mathcal{M})$ and C^s-regular (p,q) tensor S.

ii) It is $f-$linear in the first (and hence second) slot:

$$R(fX_1 + X_2, Y)S = fR(X_1, Y)S + R(X_2, Y)S,$$

for any $f \in C^1(\mathcal{M};\mathbb{R}), X_i, Y \in \mathfrak{X}_{k-1}(\mathcal{M})$ and C^s-regular (p,q) tensor S.

iii) It is $f-$linear in S:

$$R(X,Y)[fS_1 + S_2] = fR(X,Y)S_1 + R(X,Y)S_2,$$

for any $f \in C^2(\mathcal{M};\mathbb{R}), X, Y \in \mathfrak{X}_{k-1}(\mathcal{M})$ and C^s-regular (p,q) tensors S_i.

iv) The following metric compatibility condition holds:

$$g\left(R(X,Y)Z, W\right) + g\left(Z, R(X,Y)W\right) = 0, \qquad \forall \quad X, Y, Z, W \in \mathfrak{X}_{k-1}(\mathcal{M};\mathbb{R}).$$

Proof. i) This is immediate by inspection of the definition (2.20).

ii) First note that it is straightforward to see from the definition that

$$R(X_1 + X_2, Y)S = R(X_1, Y)S + R(X_2, Y)S,$$

which follows from the $\mathbb{R}$-linearity of the connection. It remains to show that $R(fX,Y)S = fR(X,Y)S$. For this we calculate

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

iii) Again, the $\mathbb{R}-$linearity of the connection quickly gives us that

$$R(X,Y)[S_1 + S_2] = R(X,Y)S_1 + R(X,Y)S_2,$$

so it remains to show that $R(X,Y)[fS] = fR(X,Y)S$. We calculate:

$$\begin{aligned} R(X,Y)[fS] &= \nabla_X \nabla_Y (fS) - \nabla_Y \nabla_X (fS), \\ &= \nabla_X (f\nabla_Y S + Y(f)S) - \nabla_Y (f\nabla_X S + X(f)S) - \nabla_{[X,Y]}(fS), \\ &= f\nabla_X \nabla_Y S + X(f)\nabla_Y S + Y(f)\nabla_X S + X(Yf)S \\ &\quad - (f\nabla_Y \nabla_X S + Y(f)\nabla_X S + X(f)\nabla_Y S + Y(Xf)S) \\ &\quad - f\nabla_{[X,Y]}S - ([X,Y]f)\, S \\ &= fR(X,Y)S + (X(Yf) - Y(Xf) - [X,Y]f)\, S \\ &= fR(X,Y)S. \end{aligned}$$

iv) For the final part, we first recall from the definition of the commutator that

$$X(Yf) - Y(Xf) - [X,Y]f = 0 \tag{2.21}$$

for any sufficiently smooth function. We will apply this with $f = g(Z,W)$. Using the fact that ∇ is the Levi-Civita connection, we have:

$$\begin{aligned} X\left(Y\left[g\left(Z,W\right)\right]\right) &= X\left(g\left(\nabla_Y Z, W\right) + g(Z, \nabla_Y W\right) \\ &= g\left(\nabla_X \nabla_Y Z, W\right) + g\left(\nabla_Y Z, \nabla_X W\right) \\ &\quad + g\left(\nabla_X Z, \nabla_Y W\right) + g\left(Z, \nabla_X \nabla_Y W\right) \end{aligned} \tag{2.22}$$

Similarly, we have

$$\begin{aligned} Y\left(X\left[g\left(Z,W\right)\right]\right) &= g\left(\nabla_Y \nabla_X Z, W\right) + g\left(\nabla_X Z, \nabla_Y W\right) \\ &\quad + g\left(\nabla_Y Z, \nabla_X W\right) + g\left(Z, \nabla_Y \nabla_X W\right) \end{aligned} \tag{2.23}$$

and

$$[X,Y]\left(g(Z,W)\right) = g\left(\nabla_{[X,Y]} Z, W\right) + g\left(Z, \nabla_{[X,Y]} W\right) \tag{2.24}$$

Taking (2.22)$-$(2.23)$-$(2.24) and noting a cancellation between terms with one derivative falling on Z and one on W, we arrive at the result.

□

Corollary 2.5. *Suppose $X, Y, Z \in \mathfrak{X}_{k-1}(\mathcal{M})$, and let $\{e_\mu\}$ be a local basis, whose dual basis is $\{e^\mu\}$. Then if $X = X^\mu e_\mu$, $Y = Y^\mu e_\mu$, $Z = Z^\mu e_\mu$, we can write:*

$$R(X,Y)Z = R^\tau{}_{\sigma\mu\nu} X^\mu Y^\nu Z^\sigma e_\tau$$

Where $R^\mu{}_{\nu\sigma\tau} \in C^{r-2}(\mathcal{M};\mathbb{R})$ are the components of a $(1,3)-$tensor, called the Riemann tensor, *given by:*

$$R^\tau{}_{\sigma\mu\nu} = e^\tau\left(R(e_\mu, e_\nu)e_\sigma\right).$$

Proof. We simply use the linearity results established above to write

$$\begin{aligned} R(X,Y)Z &= R(X^\mu e_\mu, Y^\nu e_\nu)\,[Z^\mu e_\nu] \\ &= X^\mu Y^\nu Z^\sigma R(e_\mu, e_\nu)e_\sigma \\ &= X^\mu Y^\nu Z^\sigma e^\tau\,[R(e_\mu, e_\nu)e_\sigma]\,e_\tau \\ &= R^\tau{}_{\sigma\mu\nu} X^\mu Y^\nu Z^\sigma e_\tau. \end{aligned}$$

□

From parts *i*), *iv*) of Lemma 2.9, it is straightforward to show that the Riemann tensor has the following symmetries:

$$R^\tau{}_{\sigma\mu\nu} = -R^\tau{}_{\sigma\nu\mu}, \qquad R_{\tau\sigma\mu\nu} = -R_{\sigma\tau\mu\nu},$$

where we recall that indices are raised and lowered with the metric.

Exercise 3.6. a) Suppose $\omega \in \mathfrak{X}^*_{k-1}(\mathcal{M})$ and $X, Y, Z \in \mathfrak{X}_{k-1}(\mathcal{M})$. By considering (2.21) with $f = \omega[Z]$, and recalling that $X\,(\omega[Z]) = (\nabla_X\omega)[Z] + \omega\,[\nabla_X Z]$, show that:

$$\omega\,[R(X,Y)Z] + (R(X,Y)\omega)\,Z = 0.$$

b) Deduce that in a local basis, the action of $R(X,Y)$ on a one-form is given by:

$$R(X,Y)\omega = -R^\tau{}_{\sigma\mu\nu} X^\mu Y^\nu \omega_\tau e^\sigma$$

c) Show that if S_1, S_2 are $C^{k-1}-$regular tensor fields, then

$$R(X,Y)\,[S_1 \otimes S_2] = [R(X,Y)S_1] \otimes S_2 + S_1 \otimes [R(X,Y)S_2]\,.$$

d) Deduce that in a local basis, the action of $R(X,Y)$ on an arbitrary $(p,q)-$tensor is given by:

$$\begin{aligned} [R(X,Y)S]^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_q} = \Big[& R^{\mu_1}{}_{\sigma\mu\nu} S^{\sigma\mu_2\ldots\mu_p}{}_{\nu_1\ldots\nu_q} + \cdots + R^{\mu_p}{}_{\sigma\mu\nu} S^{\mu_1\ldots\mu_{p-1}\sigma}{}_{\nu_1\ldots\nu_q} \\ & - R^\sigma{}_{\nu_1\mu\nu} S^{\mu_1\ldots\mu_p}{}_{\sigma\nu_2\ldots\nu_q} - \cdots - R^\sigma{}_{\nu_q\mu\nu} S^{\mu_1\ldots\mu_p}{}_{\nu_1\ldots\nu_{q-1}\sigma} \Big] X^\mu Y^\nu \end{aligned}$$

The exercise above shows that the Riemann tensor on its own is sufficient to determine the action of $R(X,Y)$ on an arbitrary tensor.

The Riemann tensor has further symmetries, which are encapsulated in the Bianchi identities. We will simply state these at this stage.

Theorem 2.6. *The following identities hold for any vector fields $X, Y, Z, W \in \mathfrak{X}_{k-1}(\mathcal{M})$:*

$$R(X,Y)Z + R(Y,Z)X + R(Z,X)Y = 0 \qquad \textit{1st Bianchi identity}$$

$$[\nabla_W R]\,(X,Y)Z + [\nabla_X R]\,(Y,W)Z + [\nabla_Y R]\,(W,X)Z = 0 \qquad \textit{2nd Bianchi identity}$$

Proof. The proof is essentially by direct calculation, which can be simplified somewhat by choosing the vector fields cleverly. See for example the book of Do Carmo, "Riemannian Geometry", pp 91, 106. □

Corollary 2.7. *With respect to a local basis, the first Bianchi identity can be written:*

$$R^{\tau}{}_{\sigma\mu\nu} + R^{\tau}{}_{\mu\nu\sigma} + R^{\tau}{}_{\nu\sigma\mu} = 0. \tag{2.25}$$

This, together with the previously established antisymmetry properties implies:

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

With this in mind, we can write the second Bianchi identity is equivalent to:

$$\nabla_\sigma R_{\mu\nu\tau\lambda} + \nabla_\mu R_{\nu\sigma\tau\lambda} + \nabla_\nu R_{\sigma\mu\tau\lambda} = 0. \tag{2.26}$$

Proof. Writing the 1^{st} Bianchi identity with respect to a local basis, we have:

$$R^{\tau}{}_{\sigma\mu\nu} X^\mu Y^\nu Z^\sigma e_\tau + R^{\tau}{}_{\sigma\mu\nu} Y^\mu Z^\nu X^\sigma e_\tau + R^{\tau}{}_{\sigma\mu\nu} Z^\mu X^\nu Y^\sigma e_\tau = 0$$

relabelling indices, and pulling out factors, this is equivalent to:

$$\left(R^{\tau}{}_{\sigma\mu\nu} + R^{\tau}{}_{\mu\nu\sigma} + R^{\tau}{}_{\nu\sigma\mu}\right) X^\mu Y^\nu Z^\sigma e_\tau = 0$$

which gives (2.25) since the vector fields X, Y, Z are arbitrary and $\{e_\mu\}$ is a basis.

Now, consider the same identity written out four times, with indices all lowered:

$$\begin{aligned}
0 &= R_{\tau\sigma\mu\nu} + R_{\tau\mu\nu\sigma} + R_{\tau\nu\sigma\mu} \\
0 &= R_{\sigma\mu\nu\tau} + R_{\sigma\nu\tau\mu} + R_{\sigma\tau\mu\nu} \\
0 &= R_{\mu\nu\tau\sigma} + R_{\mu\tau\sigma\nu} + R_{\mu\sigma\nu\tau} \\
0 &= R_{\nu\tau\sigma\mu} + R_{\nu\sigma\mu\tau} + R_{\nu\mu\tau\sigma}
\end{aligned}$$

Adding these four identities and using the antisymmetry on the first and last pairs of indices, all of the terms in the first column cancel against terms in the last column, and we have:

$$0 = 2R_{\tau\mu\nu\sigma} - 2R_{\nu\sigma\tau\mu}$$

which gives the result.

For the last part, we can write the second Bianchi identity with respect to the local basis as:

$$\left(\nabla_\sigma R^{\tau}{}_{\lambda\mu\nu}\right) W^\sigma X^\mu Y^\nu Z^\lambda e_\tau + \left(\nabla_\sigma R^{\tau}{}_{\lambda\mu\nu}\right) X^\sigma Y^\mu W^\nu Z^\lambda e_\tau + \left(\nabla_\sigma R^{\tau}{}_{\lambda\mu\nu}\right) Y^\sigma W^\mu X^\nu Z^\lambda e_\tau = 0$$

Relabelling, this becomes

$$\left(\nabla_\sigma R^{\tau}{}_{\lambda\mu\nu} + \nabla_\mu R^{\tau}{}_{\lambda\nu\sigma} + \nabla_\nu R^{\tau}{}_{\lambda\sigma\mu}\right) W^\sigma X^\mu Y^\nu Z^\lambda e_\tau = 0$$

Thus the second Bianchi identity is equivalent to the vanishing of the bracket above. Lowering τ and using the interchange symmetry we have:

$$\nabla_\sigma R_{\mu\nu\tau\lambda} + \nabla_\mu R_{\nu\sigma\tau\lambda} + \nabla_\nu R_{\sigma\mu\tau\lambda} = 0.$$

□

2.3.2 Ricci and scalar curvature

From the Riemann tensor it is possible to construct other curvature tensors, which capture some aspect or other of the geometry of the manifold. An important curvature is the Ricci curvature, which is a $C^{r-2}-$regular $(0,2)-$tensor, where for $X, Y \in T_p(\mathcal{M})$, $Ric_g(X,Y)|_p$ is defined as the trace of the endomorphism:

$$\mathscr{R}: \begin{array}{ccc} T_p\mathcal{M} & \to & T_p\mathcal{M} \\ Z & \mapsto & R(Z,X)Y \end{array}$$

In a local basis, we have

$$Ric_g(X,Y) = R^\tau{}_{\mu\tau\nu} X^\mu Y^\nu = R_{\mu\nu} X^\mu Y^\nu,$$

where we define:

$$R_{\mu\nu} := Ric_g(e_\mu, e_\nu).$$

by the interchange symmetry of the Riemann tensor, the Ricci tensor is symmetric, $Ric_g(X,Y) = Ric_g(Y,X)$.

We also define the scalar curvature, sometimes called the Ricci scalar, R_g, which is the trace of Ric_g with respect to the metric. In local coordinates:

$$R_g = g^{\mu\nu} R_{\mu\nu}$$

Lemma 2.10. *The contracted Bianchi identities hold:*

$$\mathrm{div}_g \left(Ric_g - \frac{1}{2} R_g g \right) = 0.$$

Proof. We work in a local basis. Consider the second Bianchi identity in components (2.26):

$$\nabla_\sigma R_{\mu\nu\tau\lambda} + \nabla_\mu R_{\nu\sigma\tau\lambda} + \nabla_\nu R_{\sigma\mu\tau\lambda} = 0.$$

Contracting with $g^{\mu\tau}$, we have:

$$\nabla_\sigma R_{\nu\lambda} + \nabla^\tau R_{\nu\sigma\tau\lambda} - \nabla_\nu R_{\sigma\lambda} = 0.$$

Contracting again with $g^{\sigma\lambda}$, we obtain:

$$2\nabla^\sigma R_{\nu\sigma} - \nabla_\nu R_g = 0$$

Dividing by 2 and re-writing this, we have:

$$\nabla^\sigma \left(R_{\nu\sigma} - \frac{1}{2} R_g g_{\nu\sigma} \right) = 0,$$

which is the result. □

The tensor field $Ric_g - \frac{1}{2} R_g g$ is often referred to as the Einstein tensor, and denoted G.

Example 12. From Example 9 we have all of the information we require to calculate the Riemann tensor for the Schwarzschild metric in Painlevé-Gullstrand coordinates, with respect to the orthonormal basis $\{e_\mu\}$ given in Example 7. We calculate, for example:

$$\begin{aligned}
R(e_0,e_1)e_0 &= \nabla_{e_0}\nabla_{e_1}e_0 - \nabla_{e_1}\nabla_{e_0}e_0 - \nabla_{[e_0,e_1]}e_0 \\
&= \nabla_{e_0}\left(\frac{1}{2r}\sqrt{\frac{2m}{r}}e_1\right) - \nabla_{e_1}(0) - \nabla_{-\frac{1}{2r}\sqrt{\frac{2m}{r}}e_1}e_0 \\
&= -\sqrt{\frac{2m}{r}}\frac{\partial}{\partial r}\left(\frac{1}{2r}\sqrt{\frac{2m}{r}}\right)e_1 + \frac{1}{2r}\sqrt{\frac{2m}{r}}\frac{1}{2r}\sqrt{\frac{2m}{r}}e_1 \\
&= \frac{2m}{r^3}e_1
\end{aligned}$$

In this way, we can calculate:

$$\begin{aligned}
&R(e_0,e_1)e_0 = \frac{2m}{r^3}e_1, \quad && R(e_0,e_1)e_1 = \frac{2m}{r^3}e_0, \quad && R(e_0,e_1)e_A = 0, \\
&R(e_0,e_A)e_0 = -\frac{m}{r^3}e_A, && R(e_0,e_A)e_1 = 0, && R(e_0,e_A)e_B = -\frac{m}{r^3}\delta_{AB}e_0, \\
&R(e_1,e_A)e_0 = 0, && R(e_1,e_A)e_1 = \frac{m}{r^3}e_A, && R(e_1,e_A)e_B = -\frac{m}{r^3}\delta_{AB}e_1, \\
&R(e_A,e_B)e_0 = 0, && R(e_A,e_B)e_1 = 0, && R(e_A,e_B)e_C = \frac{2m}{r^3}\left(\delta_{BC}e_A - \delta_{AC}e_B\right),
\end{aligned}$$

Or in terms of components, we can extract the non-trivial components of the Riemann tensor:

$$\begin{aligned}
R_{0101} &= -\frac{2m}{r^3}, \qquad & R_{0A0B} &= \frac{m}{r^3}\delta_{AB}, \\
R_{1A1B} &= -\frac{m}{r^3}\delta_{AB}, & R_{ABCD} &= \frac{2m}{r^3}\left(\delta_{BD}\delta_{AC} - \delta_{AD}\delta_{BC}\right),
\end{aligned}$$

with all other components either related to these by symmetries of the Riemann tensor, or else vanishing.

To calculate the Ricci tensor, we have to take the trace over the first and third indices of the Riemann tensor. Since we work in an orthonormal basis, this is relatively straightforward, and we find, for example:

$$\begin{aligned}
R_{00} = R^{\mu}{}_{0\mu0} &= R_{1010} + \delta^{AB}R_{A0B0} \\
&= -\frac{2m}{r^3} + \frac{m}{r^3}\delta_{AB}\delta^{AB} = 0.
\end{aligned}$$

Similarly:

$$\begin{aligned}
R_{11} = R^{\mu}{}_{0\mu0} &= -R_{0101} + \delta^{AB}R_{A1B1} \\
&= \frac{2m}{r^3} - \frac{m}{r^3}\delta_{AB}\delta^{AB} = 0.
\end{aligned}$$

and finally:

$$\begin{aligned} R_{AB} &= R^{\mu}{}_{A\mu B} = -R_{0A0B} + R_{1A1B} + R_{CADB}\delta^{CD} \\ &= -\frac{m}{r^3}\delta_{AB} - \frac{m}{r^3}\delta_{AB} + \frac{2m}{r^3}\delta^{CD}\left(\delta_{CD}\delta_{AB} - \delta_{AD}\delta_{BC}\right) \\ &= 0. \end{aligned}$$

We thus conclude that the Ricci tensor (and hence also the scalar curvature) of the Schwarzschild metric vanishes identically!

Note that the full curvature does *not* vanish: in fact, we see that the curvature becomes singular near $r = 0$, which is the location of the black hole singularity.

2.3.3 Local expressions

We can work out an expression for the various curvatures in terms of the connection coefficients, and ultimately the components of the metric. To do this, we simply have to insert expressions for various derivatives.

Lemma 2.11. *Suppose $\{e_\mu\}$ is a local basis, with commutator coefficients given by:*

$$[e_\mu, e_\nu] = C^{\sigma}{}_{\mu\nu} e_\sigma,$$

and connection coefficients given by

$$\nabla_{e_\mu} e_\nu = \Gamma^{\sigma}{}_{\mu\nu} e_\sigma.$$

Then we have the following expression for the components of the Riemann tensor:

$$R^{\tau}{}_{\sigma\mu\nu} = e_\mu\left(\Gamma^{\tau}{}_{\nu\sigma}\right) + \Gamma^{\lambda}{}_{\nu\sigma}\Gamma^{\tau}{}_{\mu\lambda} - e_\nu\left(\Gamma^{\tau}{}_{\mu\sigma}\right) - \Gamma^{\lambda}{}_{\mu\sigma}\Gamma^{\tau}{}_{\nu\lambda} - \Gamma^{\tau}{}_{\lambda\sigma}C^{\lambda}{}_{\mu\nu} \tag{2.27}$$

For the Ricci tensor, we have:

$$R_{\sigma\nu} = e_\mu\left(\Gamma^{\mu}{}_{\nu\sigma}\right) + \Gamma^{\lambda}{}_{\nu\sigma}\Gamma^{\mu}{}_{\mu\lambda} - e_\nu\left(\Gamma^{\mu}{}_{\mu\sigma}\right) - \Gamma^{\lambda}{}_{\mu\sigma}\Gamma^{\mu}{}_{\nu\lambda} - \Gamma^{\mu}{}_{\lambda\sigma}C^{\lambda}{}_{\mu\nu}$$

Proof. We calculate

$$\begin{aligned} \nabla_{e_\mu}\nabla_{e_\nu} e_\sigma &= \nabla_{e_\mu}\left(\Gamma^{\lambda}{}_{\nu\sigma} e_\lambda\right) \\ &= e_\mu\left(\Gamma^{\lambda}{}_{\nu\sigma}\right) e_\lambda + \Gamma^{\lambda}{}_{\nu\sigma}\Gamma^{\tau}{}_{\mu\lambda} e_\tau \\ &= \left[e_\mu\left(\Gamma^{\tau}{}_{\nu\sigma}\right) + \Gamma^{\lambda}{}_{\nu\sigma}\Gamma^{\tau}{}_{\mu\lambda}\right] e_\tau \end{aligned}$$

Similarly,

$$\nabla_{e_\nu}\nabla_{e_\mu} e_\sigma = \left[e_\nu\left(\Gamma^{\tau}{}_{\mu\sigma}\right) + \Gamma^{\lambda}{}_{\mu\sigma}\Gamma^{\tau}{}_{\nu\lambda}\right] e_\tau.$$

We also have

$$\begin{aligned} \nabla_{[e_\mu, e_\nu]} e_\sigma &= \nabla_{C^{\lambda}{}_{\mu\nu} e_\lambda} e_\sigma \\ &= C^{\lambda}{}_{\mu\nu}\nabla_{e_\lambda} e_\sigma \\ &= \Gamma^{\mu}{}_{\lambda\sigma} C^{\lambda}{}_{\mu\nu} e_\tau \end{aligned}$$

Now, recalling that

$$R(e_\mu, e_\nu)e_\sigma = \nabla_{e_\mu}\nabla_{e_\nu}e_\sigma - \nabla_{e_\nu}\nabla_{e_\mu}e_\sigma - \nabla_{[e_\mu,e_\nu]}e_\sigma = R^\tau{}_{\sigma\mu\nu}e_\tau$$

we have the expression above for the Riemann tensor. The expression for the Ricci tensor follows from contracting on τ, μ. □

This expression allows us to calculate the form of the Ricci tensor, thought of as a (nonlinear) partial differential operator acting on g.

Corollary 2.8. *In a local coordinate basis, we can write:*

$$R_{\alpha\sigma\mu\nu} = \frac{1}{2}\left(\frac{\partial^2 g_{\alpha\nu}}{\partial x^\mu \partial x^\sigma} + \frac{\partial^2 g_{\sigma\mu}}{\partial x^\nu \partial x^\alpha} - \frac{\partial^2 g_{\alpha\mu}}{\partial x^\nu \partial x^\sigma} - \frac{\partial^2 g_{\sigma\nu}}{\partial x^\mu \partial x^\alpha}\right) - \Gamma_{\lambda\mu\alpha}\Gamma^\lambda{}_{\nu\sigma} + \Gamma_{\lambda\nu\alpha}\Gamma^\lambda{}_{\mu\sigma}. \quad (2.28)$$

and

$$\begin{aligned} R_{\sigma\nu} = &-\frac{1}{2}g^{\mu\alpha}\frac{\partial^2 g_{\sigma\nu}}{\partial x^\mu \partial x^\alpha} \\ &+ \frac{1}{2}\frac{\partial}{\partial x^\sigma}\Gamma_{\nu\mu}{}^\mu + \frac{1}{2}\frac{\partial}{\partial x^\nu}\Gamma_{\sigma\mu}{}^\mu - \Gamma_{\lambda\mu}{}^\mu\Gamma^\lambda{}_{\nu\sigma} \\ &+ \Gamma_{\tau\lambda\nu}\Gamma^{\tau\lambda}{}_\sigma + \Gamma_{\tau\lambda\nu}\Gamma_\sigma{}^{\tau\lambda} + \Gamma_{\tau\lambda\sigma}\Gamma_\nu{}^{\tau\lambda} \end{aligned} \quad (2.29)$$

Proof. 1. As a short hand, we will use $\partial_\alpha = \frac{\partial}{\partial x^\alpha}$, $\partial^2_{\alpha\beta} = \frac{\partial^2}{\partial x^\alpha \partial x^\beta}$. We have to be careful with the partial derivatives, since unlike covariant derivatives they do *not* commute with the raising and lowering of indices. Now recall:

$$\Gamma_{\alpha\beta\tau} = g_{\alpha\tau}\Gamma^\tau{}_{\beta\tau} = \frac{1}{2}\left(\partial_\beta g_{\alpha\tau} + \partial_\tau g_{\alpha\beta} - \partial_\alpha g_{\beta\tau}\right)$$

so that

$$\partial_\beta g_{\alpha\tau} = \Gamma_{\alpha\beta\tau} + \Gamma_{\tau\beta\alpha} \quad (2.30)$$

As a consequence, we have

$$g_{\alpha\tau}\partial_\beta\left(\Gamma^\tau{}_{\mu\nu}\right) = \partial_\beta\left(\Gamma_{\alpha\mu\nu}\right) - \left(\Gamma_{\alpha\beta\tau} + \Gamma_{\tau\beta\alpha}\right)\Gamma^\tau{}_{\mu\nu}.$$

2. Inserting this into the expression (2.27) and recalling that in a coordinate basis we have $C^\lambda{}_{\mu\nu} \equiv 0$, we find:

$$\begin{aligned} R_{\alpha\sigma\mu\nu} &= g_{\alpha\tau}R^\tau{}_{\sigma\mu\nu} \\ &= g_{\alpha\tau}\partial_\mu\left(\Gamma^\tau{}_{\nu\sigma}\right) - g_{\alpha\tau}\partial_\nu\left(\Gamma^\tau{}_{\mu\sigma}\right) + \Gamma^\lambda{}_{\nu\sigma}\Gamma_{\alpha\mu\lambda} - \Gamma^\lambda{}_{\mu\sigma}\Gamma_{\alpha\nu\lambda} \\ &= \partial_\mu\left(\Gamma_{\alpha\nu\sigma}\right) - \left(\Gamma_{\alpha\mu\tau} + \Gamma_{\tau\mu\alpha}\right)\Gamma^\tau{}_{\nu\sigma} \\ &\quad - \partial_\nu\left(\Gamma_{\alpha\mu\sigma}\right) + \left(\Gamma_{\alpha\nu\tau} + \Gamma_{\tau\nu\alpha}\right)\Gamma^\tau{}_{\mu\sigma} \\ &\quad + \Gamma^\lambda{}_{\nu\sigma}\Gamma_{\alpha\mu\lambda} - \Gamma^\lambda{}_{\mu\sigma}\Gamma_{\alpha\nu\lambda} \\ &= \partial_\mu\left(\Gamma_{\alpha\nu\sigma}\right) - \partial_\nu\left(\Gamma_{\alpha\mu\sigma}\right) - \Gamma_{\tau\mu\alpha}\Gamma^\tau{}_{\nu\sigma} + \Gamma_{\tau\nu\alpha}\Gamma^\tau{}_{\mu\sigma} \end{aligned}$$

Inserting the definition of Γ into the first two terms, we calculate

$$\begin{aligned}R_{\alpha\sigma\mu\nu} &= \frac{1}{2}\left(\partial^2_{\mu\nu}g_{\alpha\sigma} + \partial^2_{\mu\sigma}g_{\alpha\nu} - \partial^2_{\mu\alpha}g_{\nu\sigma} - \partial^2_{\nu\mu}g_{\alpha\sigma} - \partial^2_{\nu\sigma}g_{\alpha\mu} + \partial^2_{\nu\alpha}g_{\mu\sigma}\right)\\ &\quad - \Gamma_{\tau\mu\alpha}\Gamma^{\tau}{}_{\nu\sigma} + \Gamma_{\tau\nu\alpha}\Gamma^{\tau}{}_{\mu\sigma}\\ &= \frac{1}{2}\left(\partial^2_{\mu\sigma}g_{\alpha\nu} + \partial^2_{\nu\alpha}g_{\mu\sigma} - \partial^2_{\mu\alpha}g_{\nu\sigma} - \partial^2_{\nu\sigma}g_{\alpha\mu}\right) - \Gamma_{\tau\mu\alpha}\Gamma^{\tau}{}_{\nu\sigma} + \Gamma_{\tau\nu\alpha}\Gamma^{\tau}{}_{\mu\sigma}\end{aligned}$$

which is (2.28).

3. To find the Ricci tensor, we first differentiate (2.30) to find:

$$\begin{aligned}\partial^2_{\mu\sigma}g_{\alpha\nu} &= \partial_\sigma\left(\Gamma_{\alpha\mu\nu} + \Gamma_{\nu\mu\alpha}\right)\\ \partial^2_{\nu\alpha}g_{\mu\sigma} &= \partial_\nu\left(\Gamma_{\mu\alpha\sigma} + \Gamma_{\sigma\alpha\mu}\right)\\ \partial^2_{\nu\sigma}g_{\alpha\mu} &= \frac{1}{2}\partial_\nu\left(\Gamma_{\alpha\sigma\mu} + \Gamma_{\mu\sigma\alpha}\right) + \frac{1}{2}\partial_\sigma\left(\Gamma_{\alpha\nu\mu} + \Gamma_{\mu\nu\alpha}\right)\end{aligned}$$

taking the sum of the first two equalities and subtracting the third, we have

$$\begin{aligned}\partial^2_{\mu\sigma}g_{\alpha\nu} + \partial^2_{\nu\alpha}g_{\mu\sigma} - \partial^2_{\nu\sigma}g_{\alpha\mu} &= \partial_\sigma\left(\Gamma_{\nu\mu\alpha}\right) + \partial_\nu\left(\Gamma_{\sigma\mu\alpha}\right)\\ &\quad + \frac{1}{2}\left[\partial_\sigma\left(\Gamma_{\alpha\mu\nu}\right) - \partial_\sigma\left(\Gamma_{\mu\alpha\nu}\right) + \partial_\nu\left(\Gamma_{\alpha\mu\sigma}\right) - \partial_\nu\left(\Gamma_{\mu\alpha\sigma}\right)\right]\end{aligned}$$

Notice that the term in square brackets is antisymmetric in α and μ, so that

$$\begin{aligned}g^{\alpha\mu}\left(\partial^2_{\mu\sigma}g_{\alpha\nu} + \partial^2_{\nu\alpha}g_{\mu\sigma} - \partial^2_{\nu\sigma}g_{\alpha\mu}\right) &= g^{\mu\alpha}\left(\partial_\sigma\left(\Gamma_{\nu\mu\alpha}\right) + \partial_\nu\left(\Gamma_{\sigma\mu\alpha}\right)\right)\\ &= \partial_\sigma\left(\Gamma_{\nu\mu}{}^{\mu}\right) + \partial_\nu\left(\Gamma_{\sigma\mu}{}^{\mu}\right)\\ &\quad - \Gamma_{\nu\mu\alpha}\partial_\sigma g^{\mu\alpha} - \Gamma_{\sigma\mu\alpha}\partial_\nu g^{\mu\alpha}\end{aligned}$$

Now, since $\partial_\mu g^{\alpha\beta} = -g^{\lambda\alpha}\left(\partial_\sigma g_{\lambda\tau}\right)g^{\beta\tau}$, we have:

$$\begin{aligned}-\Gamma_{\nu\mu\alpha}\partial_\sigma g^{\mu\alpha} - \Gamma_{\sigma\mu\alpha}\partial_\nu g^{\mu\alpha} &= \Gamma_\nu{}^{\mu\alpha}\partial_\sigma g_{\mu\alpha} + \Gamma_\sigma{}^{\mu\alpha}\partial_\nu g_{\mu\alpha}\\ &= \Gamma_\nu{}^{\mu\alpha}\left(\Gamma_{\mu\sigma\alpha} + \Gamma_{\alpha\sigma\mu}\right) + \Gamma_\sigma{}^{\mu\alpha}\left(\Gamma_{\mu\nu\alpha} + \Gamma_{\alpha\nu\mu}\right)\\ &= 2\Gamma_{\tau\lambda\nu}\Gamma_\sigma{}^{\tau\lambda} + 2\Gamma_{\tau\lambda\sigma}\Gamma_\nu{}^{\tau\lambda}\end{aligned}$$

4. To obtain (2.29) we multiply the expression (2.28) by $g^{\alpha\mu}$ and use the result of part 3. above to rewrite three of the four $\partial^2 g$ terms in terms of Γ

□

The form that we have put the Ricci tensor in might seem a bit strange: we've made some second derivatives of the metric explicit and others we have written in terms of the connection components. To explain why we have done this, at this stage it's useful to introduce a particular choice of coordinates, known as *wave coordinates*. Recall that the wave equation is given in a local coordinate basis by (2.14):

$$0 = \Box_g f = g^{\mu\nu}\frac{\partial^2 f}{\partial x^\mu \partial x^\nu} - g^{\mu\nu}\Gamma^{\sigma}{}_{\mu\nu}\frac{\partial f}{\partial x^\sigma}.$$

We say that we are working in wave coordinates if the coordinate functions x^α are themselves solutions of the wave equation, so that $\Box_g x^\alpha = 0$. Since $\partial_\sigma x^\alpha = \delta^\alpha{}_\sigma$, we get the following condition on the connection components when expressed in a local basis induced by wave coordinates:

$$0 = -g^{\mu\nu}\Gamma^\alpha{}_{\mu\nu} = \Gamma^\alpha{}_\mu{}^\mu. \tag{2.31}$$

Theorem 2.9. *With respect to wave coordinates, the Ricci tensor takes the form:*

$$R_{\sigma\nu} = R^{(w)}_{\sigma\nu} := -\frac{1}{2}g^{\mu\alpha}\frac{\partial^2 g_{\sigma\nu}}{\partial x^\mu \partial x^\alpha} + P_{\sigma\nu}(g, \partial g) \tag{2.32}$$

Here P is a homogeneous quadratic form in the first derivatives,

$$P_{\sigma\nu}(g, \partial g) = P^{\alpha\beta\gamma,\lambda\tau\mu}_{\sigma\nu}(g)\partial_\alpha g_{\beta\gamma}\partial_\lambda g_{\tau\mu},$$

where $P^{\alpha\beta\gamma,\lambda\tau\mu}_{\sigma\nu}(g) = P^{\lambda\tau\mu,\alpha\beta\gamma}_{\sigma\nu}(g)$ is determined from:

$$P^{\alpha\beta\gamma,\lambda\tau\mu}_{\sigma\nu}(g)\partial_\alpha g_{\beta\gamma}\partial_\lambda g_{\tau\mu} = \Gamma_{\tau\lambda\nu}\Gamma^{\tau\lambda}{}_\sigma + \Gamma_{\tau\lambda\nu}\Gamma_\sigma{}^{\tau\lambda} + \Gamma_{\tau\lambda\sigma}\Gamma_\nu{}^{\tau\lambda}. \tag{2.33}$$

Proof. We simply insert the condition $0 = \Gamma^\alpha{}_\mu{}^\mu$ into (2.29), which implies that all of the terms on the second line vanish. The fact that P is a homogeneous quadratic form in the first derivatives follows from the fact that $\Gamma_{\alpha\beta\gamma}$ is linear in the first derivatives, and that the right hand side of (2.33) is a symmetric quadratic form in $\Gamma_{\alpha\beta\gamma}$. □

This shows that the Ricci tensor (in appropriate coordinates) can be though of as a quasilinear wave operator acting on the metric tensor.