## Contents

Chapter 1. Preliminaries 3  
1. Special relativity 3  
2. Differential geometry: Mathematicians vs physicists 7  
3. General relativity 10  
4. Exercises 12  

Chapter 2. Exact solutions 13  
1. Minkowski spacetime 13  
2. Penrose diagrams 14  
3. The Schwarzschild solution 17  
4. Friedmann-Lemaître-Robertson-Walker models 23  
5. Matching 33  
6. Oppenheimer-Snyder collapse 35  
7. Exercises 38  

Chapter 3. Causality 47  
1. Past and future 47  
2. Causality conditions 51  
3. Exercises 52  

Appendix: Mathematical concepts for physicists 55  
Toplogy 55  
Metric spaces 56  
Hopf-Rinow theorem 56  
Differential forms 56  
Lie derivative 58  
Cartan structure equations 60  

Bibliography 63
CHAPTER 1

Preliminaries

1. Special relativity

Consider an inertial frame $S'$ moving with velocity $v$ with respect to another inertial frame $S$ along their common $x$-axis (Figure 1). According to classical mechanics, coordinate $x'$ of a point $P$ on the frame $S'$ is related to its $x$ coordinate on the frame $S$ by

$$x' = x - vt.$$ 

Moreover, a clock in $S'$ initially synchronized with a clock in $S$ is assumed to keep the same time:

$$t' = t.$$ 

Thus the spacetime coordinates of events are related by a so-called Galileo transformation

$$\begin{cases} 
  x' = x - vt \\
  t' = t 
\end{cases}.$$ 

![Figure 1. Galileo transformation.](image)

If the point $P$ is moving, its velocity in $S'$ is related to its velocity in $S$ by

$$\frac{dx'}{dt'} = \frac{dx}{dt} - v = \frac{dx}{dt} - v.$$
This is in conflict with the experimental fact that the speed of light is the same in every inertial frame, indicating that classical mechanics is not correct. Einstein solved this problem in 1905 by replacing the Galileo transformation by the so-called Lorentz transformation:

\[
\begin{align*}
    x' &= \gamma(x - vt) \\
    t' &= \gamma(t - vx)
\end{align*}
\]

Here

\[
\gamma = \frac{1}{\sqrt{1 - v^2}},
\]

and we are using units such that the speed of light is \(c = 1\) (for example measuring time in years and distance in light-years). Note that if \(|v|\) is much smaller than the speed of light, \(|v| \ll 1\), then \(\gamma \approx 1\), and we retrieve the Galileo transformation (assuming \(|v| \ll 1\)).

Under the Lorentz transformation velocities transform as

\[
\frac{dx'}{dt'} = \frac{\gamma(dx - vdt)}{\gamma(dt - vdx)} = \frac{dx}{dt} - \frac{v}{1 - v^2},
\]

In particular,

\[
\frac{dx}{dt} = 1 \Rightarrow \frac{dx'}{dt'} = \frac{1 - v}{1 - v} = 1,
\]

that is, the speed of light is the same in the two inertial frames.

In 1908, Minkowski noticed that

\[-(dt')^2 + (dx')^2 = -\gamma^2(dt - vdx)^2 + \gamma^2(dx - vdt)^2 = -dt^2 + dx^2,
\]

that is, the Lorentz transformations could be seen as isometries of \(\mathbb{R}^4\) with the indefinite metric

\[ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 = -dt \otimes dt + dx \otimes dx + dy \otimes dy + dz \otimes dz.\]

**Definition 1.1.** The pseudo-Riemannian manifold \((\mathbb{R}^4, ds^2) \equiv (\mathbb{R}^4, \langle \cdot, \cdot \rangle)\) is called the Minkowski spacetime.

Note that the set of vectors with zero square form a cone (the so-called light cone):

\[\langle v, v \rangle = 0 \iff -(v^0)^2 + (v^1)^2 + (v^2)^2 + (v^3)^2 = 0.\]

**Definition 1.2.** A vector \(v \in \mathbb{R}^4\) is said to be:

1. timelike if \(\langle v, v \rangle < 0\);
2. spacelike if \(\langle v, v \rangle > 0\);
3. lightlike, or null, if \(\langle v, v \rangle = 0\).
4. causal if it is timelike or null;
5. future-pointing if it is causal and \(\langle v, \frac{\partial}{\partial t} \rangle < 0\).

The same classification applies to (smooth) curves \(c : [a, b] \to \mathbb{R}^4\) according to its tangent vector.
timelike future-pointing vector

null vector

spacelike vector

\[ \frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \]

Figure 2. Minkowski geometry (traditionally represented with the \( t \)-axis pointing upwards).

The length \(|\langle v, v \rangle|^{\frac{1}{2}}\) of a timelike (resp. spacelike) vector \( v \in \mathbb{R}^4 \) represents the time (resp. distance) measured between two events \( p \) and \( p + v \) in the inertial frame where these events happen in the same location (resp. are simultaneous). If \( c : [a, b] \to \mathbb{R}^4 \) is a timelike curve then its length

\[
\tau(c) = \int_a^b |\langle \dot{c}(s), \dot{c}(s) \rangle|^{\frac{1}{2}} ds
\]

represents the \textbf{proper time} measured by the particle between events \( c(a) \) and \( c(b) \). We have:

**Proposition 1.3.** (Twin paradox) Of all timelike curves connecting two events \( p, q \in \mathbb{R}^4 \), the curve with \textbf{maximal} length is the line segment (representing inertial motion).

**Proof.** We may assume \( p = (0, 0, 0, 0) \) and \( q = (T, 0, 0, 0) \) on some inertial frame, and parameterize any timelike curve connecting \( p \) to \( q \) by the time coordinate:

\[
c(t) = (t, x(t), y(t), z(t)).
\]

Therefore

\[
\tau(c) = \int_0^T \sqrt{-1 + \dot{x}^2 + \dot{y}^2 + \dot{z}^2} \, dt = \int_0^T \left(1 - \dot{x}^2 - \dot{y}^2 - \dot{z}^2\right)^{\frac{1}{2}} \, dt \leq \int_0^T 1 \, dt = T.
\]

Most problems in special relativity can be recast as questions about the geometry of the Minkowski spacetime.
PROPOSITION 1.4. (Doppler effect) An observer moving with velocity $v$ away from a source of light of period $T$ measures the period to be

$$T' = T \sqrt{\frac{1 + v}{1 - v}}.$$ 

PROOF. Figure 3 represents two light signals emitted by an observer at rest at $x = 0$ with a time difference $T$. These signals are detected by an observer moving with velocity $v$, who measures a time difference $T'$ between them. Now, if the first signal is emitted at $t = t_0$, its history is the line $t = t_0 + x$. Consequently, the moving observer detects the signal at the event with coordinates

$$\begin{align*}
\begin{cases}
t = t_0 + x \\
x = vt
\end{cases} \quad \Leftrightarrow \quad 
\begin{cases}
t = \frac{t_0}{1 - v} \\
x = \frac{vt_0}{1 - v}
\end{cases}
\end{align*}$$

Similarly, the second light signal is emitted at $t = t_0 + T$, its history is the line $t = t_0 + T + x$, and it is detected by the moving observer at the event with coordinates

$$\begin{align*}
\begin{cases}
t = \frac{t_0 + T}{1 - v} \\
x = \frac{v(t_0 + T)}{1 - v}
\end{cases}
\end{align*}$$
Therefore the time difference between the signals as measured by the moving observer is

\[
T' = \sqrt{\left(\frac{t_0 + T}{1 - v} - \frac{t_0}{1 - v}\right)^2 - \left(\frac{v(t_0 + T)}{1 - v} - \frac{vt_0}{1 - v}\right)^2} \nonumber
\]

\[
= \sqrt{\frac{T^2}{(1 - v)^2} - \frac{v^2T^2}{(1 - v)^2}} = T \sqrt{\frac{1 - v^2}{(1 - v)^2}} = T \sqrt{\frac{1 + v}{1 - v}}.
\]

\[
\square
\]

In particular, two observers at rest in an inertial frame measure the same frequency for a light signal (Figure 4). However, because the gravitational field couples to all forms of energy (as \(E = mc^2\)), one expects that a photon climbing in a gravitational field to lose energy, hence frequency. In 1912, Einstein realized that this could be modelled by considering curved spacetime geometries, so that equal line segments in a (flat) spacetime diagram do not necessarily correspond to the same length.

\[\text{Figure 4. Minkowski geometry is incompatible with the gravitational redshift.}\]

2. Differential geometry: Mathematicians vs physicists

Einstein’s idea to incorporate gravitation into relativity was to replace the Minkowski spacetime \((\mathbb{R}^4, \langle \cdot, \cdot \rangle)\) by a curved four-dimensional Lorentzian manifold \((M, g) \equiv (\mathbb{R}^4, \langle \cdot, \cdot \rangle)\). Here \(g\) is a Lorentzian metric, that is, a symmetric 2-tensor field such that at each tangent space \(g = \text{diag}(-1, 1, 1, 1)\) in an appropriate basis. Just like in Riemannian geometry, \(g\) determines a
Levi-Civita connection, the unique connection $\nabla$ which is symmetric and compatible with $g$:

$$\nabla_X Y - \nabla_Y X = [X, Y];$$
$$X \cdot \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle,$$

for all vector fields $X, Y, Z$. The curvature of this connection is then given by the operator

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

The formulas above were written using the abstract notation usually employed by mathematicians. It is often very useful (especially when dealing with contractions) to use the more explicit notation usually adopted by physicists, which emphasizes the indices of the various objects when written in local coordinates:

<table>
<thead>
<tr>
<th>Object</th>
<th>Mathematicians</th>
<th>Physicists</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vector field</td>
<td>$X$</td>
<td>$X^\mu$</td>
</tr>
<tr>
<td>Tensor product</td>
<td>$X \otimes Y$</td>
<td>$X^\mu Y^\nu$</td>
</tr>
<tr>
<td>Metric</td>
<td>$g \equiv \langle \cdot, \cdot \rangle$</td>
<td>$g_{\mu\nu}$</td>
</tr>
<tr>
<td>Inner product</td>
<td>$g(X, Y) \equiv \langle X, Y \rangle$</td>
<td>$g_{\mu\nu} X^\mu Y^\nu$</td>
</tr>
<tr>
<td>Associated covector</td>
<td>$X^\sharp \equiv g(X, \cdot)$</td>
<td>$X_\nu \equiv g_{\mu\nu} X^\mu$</td>
</tr>
<tr>
<td>Covariant derivative</td>
<td>$\nabla_X$</td>
<td>$\nabla_\mu X^\nu \equiv \partial_\mu X^\nu + \Gamma^\nu_{\mu\alpha} X^\alpha$</td>
</tr>
<tr>
<td>Covariant derivative tensor</td>
<td>$\nabla X$</td>
<td></td>
</tr>
</tbody>
</table>

The covariant derivative tensor of a vector field $X$, not always emphasized in differential geometry courses for mathematicians, is simply the $(1, 1)$-tensor field defined by

$$\nabla X(Y) = \nabla_Y X.$$

Also not always emphasized in differential geometry courses for mathematicians is the fact that any connection can be naturally extended to act on tensor fields (via the Leibnitz rule). For instance, if $\omega$ is a covector field (1-form) then one defines

$$(\nabla_X \omega)(Y) = X \cdot [\omega(Y)] - \omega(\nabla_X Y).$$

In local coordinates, this is

$$(X^\mu \nabla_\mu \omega_\nu)Y^\nu = X^\mu \partial_\mu (\omega_\nu Y^\nu) - \omega_\nu (X^\mu \nabla_\mu Y^\nu)$$
$$= X^\mu (\partial_\mu \omega_\nu Y^\nu + X^\mu \omega_\nu \partial_\nu Y^\nu) - \omega_\nu (X^\mu \partial_\mu Y^\nu + X^\mu \Gamma^\nu_{\mu\alpha} Y^\alpha)$$
$$= (\partial_\mu \omega_\nu - \Gamma^\nu_{\mu\alpha} \omega_\alpha) X^\mu Y^\nu,$$

that is,

$$\nabla_\mu \omega_\nu = \partial_\mu \omega_\nu - \Gamma^\alpha_{\mu\alpha} \omega_\nu.$$

The generalization for higher rank tensors is obvious: for instance, if $T$ is a $(2, 1)$-tensor then

$$\nabla_\alpha T^\beta_{\mu\nu} = \partial_\alpha T^\beta_{\mu\nu} + \Gamma^\beta_{\alpha\gamma} T^\gamma_{\mu\nu} - \Gamma^\gamma_{\alpha\mu} T^\beta_{\gamma\nu} - \Gamma^\gamma_{\alpha\nu} T^\beta_{\mu\gamma}.$$
2. DIFFERENTIAL GEOMETRY: MATHEMATICIANS VS PHYSICISTS

Note that the condition of compatibility of the Levi-Civita connection with the metric is simply
\[ \nabla g = 0. \]
In particular, the operations of raising and lowering indices commute with covariant differentiation.

As an exercise in index gymnastics, we will now derive a series of identities involving the Riemann curvature tensor. We start by rewriting its definition in the notation of the physicists:

\[
R_{\alpha\beta\mu\nu}^\gamma X^\alpha Y^\beta Z^\mu =
\]
\[
= X^\alpha \nabla_\alpha (Y^\beta \nabla_\beta Z^\mu) - Y^\alpha \nabla_\alpha (X^\beta \nabla_\beta Z^\mu) - (X^\alpha \nabla_\alpha Y^\beta - Y^\alpha \nabla_\alpha X^\beta) \nabla_\beta Z^\mu
\]
\[
= (X^\alpha \nabla_\alpha Y^\beta) (\nabla_\beta Z^\mu) + X^\alpha Y^\beta \nabla_\alpha Z^\mu - (Y^\alpha \nabla_\alpha X^\beta) (\nabla_\beta Z^\mu)
\]
\[
- Y^\alpha X^\beta \nabla_\alpha \nabla_\beta Z^\mu - (X^\alpha \nabla_\alpha Y^\beta) \nabla_\beta Z^\mu + (Y^\alpha \nabla_\alpha X^\beta) \nabla_\beta Z^\mu
\]
\[
= X^\alpha Y^\beta (\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha) Z^\mu.
\]
In other words,
\[
R_{\alpha\beta\mu\nu}^\gamma Z^\nu = (\nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha) Z^\mu,
\]
or, equivalently,
\[
2 \nabla_{[\alpha} \nabla_{\beta]} Z^\mu = R_{\alpha\beta\mu\nu} Z^\nu,
\]
where the square brackets indicate anti-symmetrization. This is readily generalized for arbitrary tensors: from
\[
2 \nabla_{[\alpha} \nabla_{\beta]} (Z^\mu W^\nu) = (2 \nabla_{[\alpha} \nabla_{\beta]} Z^\mu) W^\nu + (2 \nabla_{[\alpha} \nabla_{\beta]} W^\nu) Z^\mu
\]
\[
= R_{\alpha\beta\mu\sigma} Z^\sigma W^\nu + R_{\alpha\beta\nu\sigma} W^\sigma Z^\mu
\]
one readily concludes that
\[
2 \nabla_{[\alpha} \nabla_{\beta]} T_{\mu\nu} = R_{\alpha\beta\mu\sigma} T^\sigma_{\nu} + R_{\alpha\beta\nu\sigma} T^\sigma_{\mu}.
\]
Let us choose
\[
Z^\mu = \nabla_\mu f \equiv \partial_\mu f
\]in equation (1). We obtain
\[
R_{[\alpha\beta\mu],\nu} Z^\nu = 2 \nabla_{[\alpha} \nabla_{\beta]} Z^\mu = 2 \nabla_{[\alpha} \nabla_{\beta]} Z^\mu = 0,
\]
because
\[
\nabla_{[\mu} Z^\nu = \partial_{[\mu} Z^\nu - \Gamma^\nu_{\mu\nu} Z_\alpha = \partial_{[\mu} \partial_\nu f = 0.
\]
Since we can choose \( Z \) arbitrarily at a given point, it follows that
\[
R_{[\alpha\beta\mu],\nu} = 0 \iff R_{\alpha\beta\mu\nu} + R_{\beta\mu\alpha\nu} + R_{\mu\alpha\beta\nu} = 0.
\]
This is the so-called first Bianchi identity, and is key for obtaining the full set of symmetries of the Riemann curvature tensor:
\[
R_{\alpha\beta\mu\nu} = - R_{\beta\alpha\mu\nu} = - R_{\alpha\beta\nu\mu} = R_{\mu\alpha\beta}.
\]
In the notation of the mathematicians, it is written as
\[
R(X, Y) Z + R(Y, Z) X + R(Z, X) Y = 0.
\]
for all vector fields $X, Y, Z$.

Let us now take the covariant derivative of equation (1):

$$\nabla_\gamma R_{\alpha\beta\mu\nu} Z^\nu + R_{\alpha\beta\mu\nu} \nabla_\gamma Z^\nu = 2 \nabla_\gamma (\nabla_\alpha \nabla_\beta) Z_\mu.$$

At any given point we can choose $Z$ such that

$$\nabla_\gamma Z^\nu \equiv \partial_\gamma Z^\nu + \Gamma^\nu_{\gamma\delta} Z^\delta = 0.$$

Assuming this, we then obtain

$$\nabla_\gamma \left[R_{\alpha\beta} Z_\mu\right] = 2 \nabla_\gamma \left[\nabla_\alpha \nabla_\beta\right] Z_\mu = 2 \nabla_\gamma (\nabla_\alpha \nabla_\beta) Z_\mu = 0.$$

Since we can choose $Z$ arbitrarily at a given point, it follows that

$$(2) \quad \nabla_{[\alpha} R_{\beta\gamma\mu\nu]} = 0 \iff \nabla_\alpha R_{\beta\gamma\mu\nu} + \nabla_\beta R_{\gamma\alpha\mu\nu} + \nabla_\gamma R_{\alpha\beta\mu\nu} = 0.$$

This is the so-called **second Bianchi identity**. In the notation of the mathematicians, it is written as

$$\nabla \left[R(X, Y, Z, \cdot, \cdot) + \nabla R(Y, Z, X, \cdot, \cdot) + \nabla R(Z, X, Y, \cdot, \cdot) = 0\right]$$

for all vector fields $X, Y, Z$.

Recall that the Riemann curvature tensor has only one independent contraction, called the **Ricci tensor**:

$$R_{\mu\nu} = R_{\alpha\mu} \alpha^{\nu}.$$

The trace of the Ricci tensor, in turn, is known as the **scalar curvature**:

$$R = g^{\mu\nu} R_{\mu\nu}.$$

These quantities satisfy the so-called **contracted Bianchi identity**, which is obtained from (2) by contracting the pairs of indices $(\beta, \mu)$ and $(\gamma, \nu)$:

$$\nabla_\alpha R - \nabla^\beta R_{\alpha\beta} - \nabla^\gamma R_{\alpha\gamma} = 0 \iff \nabla^\beta R_{\alpha\beta} - \frac{1}{2} \nabla_\alpha R = 0 \iff \nabla^\beta \left(R_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta}\right) = 0.$$

The contracted Bianchi identity is equivalent to the statement that the **Einstein tensor**

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

is divergenceless:

$$\nabla^\mu G_{\mu\nu} = 0.$$

3. **General relativity**

Newtonian gravity is described by a scalar function $\phi$, called the **gravitational potential**. The equation of motion for a free-falling particle of mass $m$ in Cartesian coordinates is

$$m \frac{d^2 x^i}{dt^2} = -m \partial_i \phi \iff \frac{d^2 x^i}{dt^2} = -\partial_i \phi.$$
Note that all free-falling particles describe the same trajectories (an observation dating back to Galileo). The gravitational potential is determined from the matter mass density $\rho$ by the Poisson equation
\[ \Delta \phi = 4\pi \rho \]
(using units such that Newton’s gravitational constant is $G = 1$; this choice, together with $c = 1$, defines the so-called geometrized units, where lengths, time intervals and masses all have the same dimensions).

To implement his idea of describing gravity via a curved four-dimensional Lorentzian manifold $(M, g)$, Einstein had to specify $(i)$ how free-falling particles would move on this manifold, and $(ii)$ how to determine the curved metric $g$. Since free particles move along straight lines in the Minkowski spacetime, Einstein proposed that free falling particles should move along timelike geodesics. In other words, he suggested replacing the Newtonian equation of motion by the geodesic equation
\[ \ddot{x}^\mu + \Gamma^\mu_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta = 0. \]
Moreover, Einstein knew that it is possible to define the energy-momentum tensor $T_{\mu\nu}$ of the matter content of the Minkowski spacetime, so that the conservation of energy and momentum is equivalent to the vanishing of its divergence:
\[ \nabla_\mu T_{\mu\nu} = 0. \]
This inspired Einstein to propose that $g$ should satisfy the so-called Einstein field equation:
\[ G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi T_{\mu\nu}. \]
Here $\Lambda$ is a constant, known as the cosmological constant. Note that Einstein field equation implies, via the contracted Bianchi identity, that the energy-momentum tensor is divergenceless.

As a simple example, we consider a pressureless perfect fluid, known as dust. Its energy-momentum tensor is
\[ T_{\mu\nu} = \rho U_\mu U_\nu, \]
where $\rho$ is the dust rest density and $U$ is a unit timelike vector field tangent to the histories of the dust particles. The equations of motion for the dust can be found from
\[ \nabla^\nu T_{\mu\nu} = 0 \Leftrightarrow \left[ \nabla^\nu (\rho U_\mu) \right] U_\nu + \rho U_\mu \nabla^\nu U_\nu = 0 \]
\[ \Leftrightarrow \text{div}(\rho U)U + \rho \nabla U U = 0. \]
Since $U$ and $\nabla U U$ are orthogonal (because $\langle U, U \rangle = -1$), we find
\[ \begin{cases} \text{div}(\rho U) = 0 \\ \nabla U U = 0 \end{cases} \]
in the support of $\rho$. These are, respectively, the equation of conservation of mass and the geodesic equation. Thus the fact that free-falling particles
move along geodesics can be seen as a consequence of the Einstein field equation (at least in this model).

4. Exercises

(1) **Twin paradox:** Two twins, Alice and Bob, are separated on their 20th birthday. While Alice remains on Earth (which is an inertial frame to a very good approximation), Bob departs at 80% of the speed of light towards Planet X, 8 light-years away from Earth. Therefore Bob reaches his destination 10 years later (as measured on the Earth’s frame). After a short stay, he returns to Earth, again at 80% of the speed of light. Consequently Alice is 40 years old when she sees Bob again.

(a) How old is Bob when they meet again?
(b) How can the asymmetry in the twins’ ages be explained? Notice that from Bob’s point of view he is at rest in his spaceship and it is the Earth which moves away and then back again.
(c) Imagine that each twin watches the other through a very powerful telescope. What do they see? In particular, how much time do they experience as they see one year elapse for their twin?

(2) A particularly simple matter model is that of a smooth **massless scalar field** $\phi : M \to \mathbb{R}$, whose energy-momentum tensor is

$$T_{\mu\nu} = \nabla_\mu \phi \nabla_\nu \phi - \frac{1}{2} (\nabla_\alpha \phi \nabla^\alpha \phi) g_{\mu\nu}.$$ 

Show that if the Lorentzian manifold $(M, g)$ satisfies the Einstein equation with this matter model then $\phi$ satisfies the wave equation

$$\Box \phi = 0 \iff \nabla^\mu \nabla_\mu \phi = 0.$$ 

(3) The energy-momentum tensor for **perfect fluid** is

$$T_{\mu\nu} = (\rho + p) U_\mu U_\nu + pg_{\mu\nu},$$

where $\rho$ is the fluid’s rest density, $p$ is the fluid’s rest pressure, and $U$ is a unit timelike vector field tangent to the histories of the fluid particles. Show that:

(a) $(T_{\mu\nu}) = \text{diag}(\rho, p, p, p)$ in any orthonormal frame including $U$;
(b) The motion equations for the perfect fluid are

$$\begin{align*}
\text{div}(\rho U) + p \text{div} U &= 0 \\
(p + p) \nabla_\nu U &= -(\text{grad} \ p)^\perp,
\end{align*}$$

where $^\perp$ represents the orthogonal projection on the spacelike hyperplane orthogonal to $U$. 

CHAPTER 2

**Exact solutions**

1. Minkowski spacetime

The simplest solution of the Einstein field equation with zero cosmological constant in vacuum (i.e. with vanishing energy-momentum tensor) is the Minkowski spacetime, that is, $\mathbb{R}^4$ with the metric

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2.$$  

Since this metric is flat, its curvature vanishes, and so do its Ricci and Einstein tensors. It represents a universe where there is no gravity whatsoever. Transforming the Cartesian coordinates $(x, y, z)$ to spherical coordinates $(r, \theta, \varphi)$ yields

$$ds^2 = -dt^2 + dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\varphi^2).$$

Performing the additional change of coordinates

$$\begin{cases}
  u = t - r \quad \text{(retarded time)} \\
  v = t + r \quad \text{(advanced time)}
\end{cases}$$

we obtain

$$ds^2 = -du dv + r^2 (d\theta^2 + \sin^2 \theta d\varphi^2),$$

where

$$r(u, v) = \frac{1}{2} (v - u).$$

The coordinates $(u, v)$ are called **null coordinates**: their level sets are null cones formed by outgoing/ingoing null geodesics emanating from the center. Note that they are subject to the constraint

$$r \geq 0 \Leftrightarrow v \geq u.$$  

Finally, the coordinate change

(3)

$$\begin{cases}
  \tilde{u} = \tanh u \\
  \tilde{v} = \tanh v
\end{cases} \Leftrightarrow \begin{cases}
  u = \text{arctanh} \, \tilde{u} \\
  v = \text{arctanh} \, \tilde{v}
\end{cases}$$

brings the metric into the form

$$ds^2 = -\frac{1}{(1 - \tilde{u}^2)(1 - \tilde{v}^2)} d\tilde{u} d\tilde{v} + r^2 (d\theta^2 + \sin^2 \theta d\varphi^2),$$

where now

$$r (\tilde{u}, \tilde{v}) = \frac{1}{2} (\text{arctanh} \, \tilde{v} - \text{arctanh} \, \tilde{u})$$
and
\[ -1 < \tilde{u} \leq \tilde{v} < 1. \]
Because \((\tilde{u}, \tilde{v})\) are also null coordinates, it is common to represent their axes tilted by 45°. The plane region defined by (4) is then represented in Figure 1.

![Figure 1. Range of the coordinates \((\tilde{u}, \tilde{v})\).](image)

This region is usually called the **Penrose diagram** for the Minkowski spacetime. If we take each point in the diagram to represent a sphere \(S^2\) of radius \(r(\tilde{u}, \tilde{v})\), the diagram itself represents the full spacetime manifold, in a way that makes causality relations apparent: any causal curve is represented in the diagram by a curve with tangent at most 45° from the vertical. In Figure 2 we represent some level hypersurfaces of \(t\) and \(r\) in the Penrose diagram. The former approach the point \(i^0\) in the boundary of the diagram, called the **spacelike infinity**, whereas the later go from the boundary point \(i^-\) (**past timelike infinity**) to the boundary point \(i^+\) (**future timelike infinity**). Finally, null geodesics start at the null boundary line \(I^-\) (**past null infinity**) and end at the null boundary line \(I^+\) (**future null infinity**). These boundary points and lines represent ideal points at infinity, and do not correspond to actual points in the Minkowski spacetime.

### 2. Penrose diagrams

The concept of Penrose diagram can be easily generalized for any spherically symmetric space-time. Such spacetimes have metric
\[
\begin{align*}
&ds^2 = g_{AB}dx^A dx^B + r^2 \left( d\theta^2 + \sin^2 \theta d\phi^2 \right), \\
&\quad r = r(x^0, x^1),
\end{align*}
\]
where \(g_{AB}dx^A dx^B\) is a Lorentzian metric on a 2-dimensional quotient manifold with boundary (which we assume to be diffeomorphic to a region of the plane). It turns out that any such metric is conformal to the Minkowski metric:
\[
\begin{align*}
&g_{AB}dx^A dx^B = -\Omega^2 du dv, \\
&\quad \Omega = \Omega(u, v).
\end{align*}
\]
This can be seen locally as follows: choose a spacelike line $\Sigma$, a coordinate $u$ along it, and a coordinate $w$ along a family of null geodesics emanating from $\Sigma$, so that $\Sigma$ corresponds to $w = 0$ (Figure 3). Then near $\Sigma$ we have

$$g_{uu} = \left\langle \frac{\partial}{\partial u}, \frac{\partial}{\partial u} \right\rangle > 0$$

and

$$g_{ww} = \left\langle \frac{\partial}{\partial w}, \frac{\partial}{\partial w} \right\rangle = 0.$$
As for any 1-form in a 2-dimensional manifold, we have
\[
du + \frac{2g_{uw}}{g_{uu}}dw = f dv
\]
for suitable functions \(f\) and \(v\). Note that \(f\) cannot vanish, because \((u, w)\) are local coordinates. Moreover, we can assume \(f > 0\) by replacing \(v\) with \(-v\) if necessary. Choosing \(\Omega^2 = fg_{uu}\) then yields (5).

We then see that any spherically symmetric metric can be written as
\[
ds^2 = -\Omega^2 du dv + r^2 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right)
\]
with \(\Omega = \Omega(u, v)\) and \(r = r(u, v)\). By rescaling \(u\) and \(v\) if necessary, we can assume that the range of \((u, v)\) is bounded, and hence obtain a Penrose diagram depicting the causal geometry. As we will see, this is extremely helpful in more complicated spherical symmetric solutions of the Einstein field equation.

**Remark 2.1.** From
\[
(g_{AB}) = \begin{pmatrix} 0 & -\frac{\Omega^2}{2} \\ -\frac{\Omega^2}{2} & 0 \end{pmatrix} \Rightarrow (g^{AB}) = \begin{pmatrix} 0 & -\frac{\Omega^2}{2} \\ -\frac{\Omega^2}{2} & 0 \end{pmatrix}
\]

it is easily seen that
\[
\partial_A \left( \sqrt{\text{det}(-g_{CD})} g^{AB} \partial_B u \right) = 0 \Leftrightarrow \nabla_A \nabla^A u = 0.
\]

and similarly for \(v\). In other words, the null coordinates \(u\) and \(v\) are solutions of the wave equation in the 2-dimensional Lorentzian manifold:
\[
\Box u = \Box v = 0.
\]

This is the Lorentzian analogue of the so-called **isothermal coordinates** for Riemannian surfaces. The proof that the later exist locally is however slightly more complicated: given a point \(p\) on the surface, one chooses a local harmonic function with nonvanishing derivative,
\[
\Delta u = 0, \quad (du)_p \neq 0,
\]
and considers the equation
\[
dv = \star du.
\]

Here \(\star\) is the Hodge star, which for generic orientable \(n\)-dimensional pseudo-Riemannian manifolds is defined as follows: if \(\{\omega^1, \ldots, \omega^n\}\) is any positively oriented orthonormal coframe then
\[
\star(\omega^1 \wedge \cdots \wedge \omega^k) = \langle \omega^1, \omega^1 \rangle \cdots \langle \omega^k, \omega^k \rangle \omega^{k+1} \wedge \cdots \wedge \omega^n.
\]

By the Poincaré Lemma, equation (7) can be locally solved, since
\[
d \star du = \star \star d \star du = \star (\Delta u) = 0.
\]

Moreover, \(v\) is itself harmonic, because
\[
\Delta v = \star d \star dv = \star d \star du = \star d(-du) = 0.
\]
Finally,
\[ \|du\| = \|dv\| = \frac{1}{\Omega} \]
for some local function \( \Omega > 0 \), and so the metric is written in these coordinates as
\[ ds^2 = \Omega^2 (du^2 + dv^2) . \]

3. The Schwarzschild solution

If we try to solve the vacuum Einstein field equation with zero cosmological constant for a spherically symmetric Lorentzian metric, we obtain, after suitably rescaling the time coordinate, the \textbf{Schwarzschild solution}
\[
ds^2 = - \left(1 - \frac{2M}{r}\right)dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\phi^2)
\]
(where \( M \in \mathbb{R} \) is a constant). Note that for \( M = 0 \) we retrieve the Minkowski metric in spherical coordinates. Note also that if \( M > 0 \) then the metric is defined in two disconnected domains of coordinates, corresponding to \( r \in (0, 2M) \) and \( r \in (2M, +\infty) \).

The physical interpretation of the Schwarzschild solution can be found by considering the proper time of a timelike curve parameterized by the time coordinate:
\[
\tau = \int_{t_0}^{t_1} \left[ \left(1 - \frac{2M}{r}\right) - \left(1 - \frac{2M}{r}\right)^{-1} r^2 - r^2 \dot{\theta}^2 - r^2 \sin^2 \theta \dot{\phi}^2 \right]^{\frac{1}{2}} dt,
\]
where \( \dot{r} = \frac{dr}{dt} \), etc. The integrand \( L_S \) is the Lagrangian for geodesic motion in the Schwarzschild spacetime when parameterized by the time coordinate. Now for motions with speeds much smaller than the speed of light we have \( \dot{r}^2 \ll 1 \), etc. Assuming \( \frac{M}{r} \ll 1 \) as well we have
\[
L_S = \left[ 1 - \frac{2M}{r} - \left(1 - \frac{2M}{r}\right)^{-1} r^2 - r^2 \dot{\theta}^2 - r^2 \sin^2 \theta \dot{\phi}^2 \right]^{\frac{1}{2}}
\]
\[
\simeq 1 - \frac{M}{r} - \frac{1}{2} \left( \dot{r}^2 + r^2 \dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2 \right) = 1 - L_N,
\]
where
\[
L_N = \frac{1}{2} \left( \dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta \dot{\phi}^2 \right) + \frac{M}{r}
\]
is precisely the Newtonian Lagrangian for the motion of a particle in the gravitational field of a point mass \( M \). The Schwarzschild solution should therefore be considered the relativistic analogue of this field.
To write the Schwarzschild metric in the form (6) we note that the quotient metric is

\[ ds^2 = - \left( 1 - \frac{2M}{r} \right) dt^2 + \left( 1 - \frac{2M}{r} \right)^{-1} dr^2 \]

\[ = - \left( 1 - \frac{2M}{r} \right) \left[ dt^2 - \left( 1 - \frac{2M}{r} \right) dr^2 \right] \]

\[ = - \left( 1 - \frac{2M}{r} \right) \left[ dt - \left( 1 - \frac{2M}{r} \right)^{-1} dr \right] \left[ dt + \left( 1 - \frac{2M}{r} \right)^{-1} dr \right] \]

\[ = - \left( 1 - \frac{2M}{r} \right) du \, dv, \]

where we define

\[ u = t - \int \left( 1 - \frac{2M}{r} \right)^{-1} dr = t - r - 2M \log |r - 2M| \]

and

\[ v = t + \int \left( 1 - \frac{2M}{r} \right)^{-1} dr = t + r + 2M \log |r - 2M|. \]

In the domain of coordinates \( r > 2M \) we have \( 1 - \frac{2M}{r} > 0 \), and so the quotient metric is already in the required form. Note however that, unlike what happened in the Minkowski spacetime, we now have \( v - u = 2r + 4M \log |r - 2M| \in (-\infty, +\infty) \).

Consequently, by applying the coordinate rescaling (3) we obtain the full square, instead of a triangle (Figure 4). Besides the infinity points and null boundaries also present in the Penrose diagram for the Minkowski spacetime, there are two new null boundaries, \( \mathcal{H}^- \) (past event horizon) and \( \mathcal{H}^+ \) (future event horizon), where \( r = 2M \).

It seems reasonable to expect that the metric can be extended across the horizons, since \( r \) does not tend to zero nor to infinity there; this expectation is confirmed by calculating the so-called Kretschmann scalar:

\[ R_{\alpha\beta\mu\nu} R^{\alpha\beta\mu\nu} = \frac{48}{r^6}. \]

This is perfectly well behaved as \( r \to 2M \), and seems to indicate that the horizons are mere singularities of the coordinate system \( (t, r) \). To show that this is indeed the case, note that in the \( (u, r) \) coordinate system the quotient metric is written

\[ ds^2 = - \left( 1 - \frac{2M}{r} \right) du^2 - 2du \, dr. \]

Since

\[ \det \begin{pmatrix} -1 & \frac{2M}{r} & -1 \\ -1 & 0 \end{pmatrix} = -1, \]

the metric is well behaved as \( r \to 2M \).
we see that the metric is well defined across \( r = 2M \) in this coordinate system. Moreover, we know that it solves the Einstein equation in the coordinate domains \( r < 2M \) and \( r > 2M \); by continuity, it must solve it in the whole domain \( r < 2M \) and \( r > 2M \). Note that the coordinate domains \( r < 2M \) and \( r > 2M \) are glued along \( r = 2M \) so that the outgoing null geodesics \( u = \text{constant} \) go from \( r = 0 \) to \( r = +\infty \); in other words, the gluing is along the past event horizon \( H^- \).

To obtain the Penrose diagram for the coordinate domain \( r < 2M \) we note that the quotient metric can be written as

\[
\begin{align*}
  ds^2 &= -\left(1 - \frac{2M}{r}\right) du dv = -\left(\frac{2M}{r} - 1\right) du (-dv) = -\left(\frac{2M}{r} - 1\right) du dv',
\end{align*}
\]

where \( v' = -v \). Since in this coordinate domain \( \frac{2M}{r} - 1 > 0 \), the quotient metric is in the required form. Note however that we now have

\[
  u + v' = -2r - 4M \log |r - 2M| \in (-4M \log(2M), +\infty),
\]

and by setting

\[
  v'' = v' + 4M \log(2M)
\]

we obtain

\[
  u + v'' > 0.
\]

Consequently, by applying the coordinate rescaling (3) we obtain a triangle (Figure 5). There is now a spacelike boundary, where \( r = 0 \), and two null boundaries \( H^- \), where \( r = 2M \). The Penrose diagram for the domain of the coordinates \((u, r)\) can be obtained gluing the Penrose diagrams in Figures 4 and 5 along \( H^- \), so that the null geodesics \( u = \text{constant} \) match (Figure 6).
2. EXACT SOLUTIONS

Figure 5. Penrose diagram for a region $r < 2M$ of the Schwarzschild spacetime, with some level hypersurfaces of $t$ and $r$ represented.

If instead we use the $(v, r)$ coordinate system, the quotient metric is written

$$ds^2 = - \left( 1 - \frac{2M}{r} \right) dv^2 + 2dv dr.$$  

Again the metric is well defined across $r = 2M$ in this coordinate system, since

$$\det \begin{pmatrix} -1 + \frac{2M}{r} & 1 \\ 1 & 0 \end{pmatrix} = -1,$$

and solves the Einstein equation in the whole coordinate domain $r \in (0, +\infty)$. The coordinate domains $r < 2M$ and $r > 2M$ are now glued along $r = 2M$ so that the ingoing null geodesics $v = \text{constant}$ go from $r = +\infty$ to $r = 0$; in other words, the gluing is along the future event horizon $\mathcal{H}^+$.  

Figure 6. Penrose diagram for the domain of the coordinates $(u, r)$ in the Schwarzschild spacetime, with some level hypersurfaces of $t$ and $r$ represented.
To obtain the Penrose diagram for the coordinate domain $r < 2M$ we note that the quotient metric can be written as

$$ds^2 = -\left(1 - \frac{2M}{r}\right) du dv = -\left(\frac{2M}{r} - 1\right) (-du) dv = -\left(\frac{2M}{r} - 1\right) du' dv,$$

where $u' = -u$. Since in this coordinate domain $\frac{2M}{r} - 1 > 0$, the quotient metric is in the required form. We have

$$u' + v = 2r + 4M \log |r - 2M| \in (-\infty, 4M \log(2M)),$$

and by setting

$$u'' = u' - 4M \log(2M)$$

we obtain

$$u'' + v < 0.$$

Consequently, by applying the coordinate rescaling (3) we obtain a triangle (Figure 7). Again there is a spacelike boundary, where $r = 0$, and two null boundaries $\mathcal{H}^+$, where $r = 2M$. The Penrose diagram for the domain of the coordinates $(v, r)$ can be obtained gluing the Penrose diagrams in Figures 4 and 7 along $\mathcal{H}^+$, so that the null geodesics $v = \text{constant}$ match (Figure 8).

![Penrose diagram for a region $r < 2M$ of the Schwarzschild spacetime, with some level hypersurfaces of $t$ and $r$ represented.](image)

Both regions $r < 2M$ can of course be glued to the region $r > 2M$ simultaneously. By symmetry, it is also clear that a mirror-reversed copy of the region $r > 2M$ can be glued to the surviving null boundaries $\mathcal{H}^-$ and $\mathcal{H}^+$ (Figure 9). The resulting spacetime, known as the **maximal analytical extension** of the Schwarzschild solution, is a solution of the Einstein equation which cannot be extended any further, since $r \to 0$ or $r \to +\infty$ on the boundary of its Penrose diagram. Note that by continuity the Einstein equation holds at the point where the four Penrose diagrams intersect (known as the **bifurcate sphere**).

Let us now analyze in detail the Penrose diagram for the maximal analytic extension of the Schwarzschild spacetime. There are two **asymptotically flat** regions $r > 2M$, corresponding to two causally disconnected universes, joined by a **wormhole**. There are also two regions where $r < 2M$: a
2. EXACT SOLUTIONS

\[ r = 0 \]

\[ i^+ \quad H^+ \quad I^+ \quad J^+ \quad i^0 \]

\[ i^- \quad H^- \quad I^- \quad J^- \quad i^- \]

**Figure 8.** Penrose diagram for the domain of the coordinates \((v, r)\) in the Schwarzschild spacetime, with some level hypersurfaces of \(t\) and \(r\) represented.

**Figure 9.** Penrose diagram for the maximal analytic extension of the Schwarzschild spacetime, with some level hypersurfaces of \(t\) and \(r\) represented.

**black hole region**, bounded by the future event horizons \(H^+\), from which no causal curve can escape; and a **white hole region**, bounded by the past event horizons \(H^-\), from which every causal curve must escape. Note that the horizons themselves correspond to spheres which are propagating at the speed of light, but whose radius remains constant, \(r = 2M\).

The black hole in the maximal analytic extension of the Schwarzschild spacetime is an **eternal black hole**, that is, a black hole which has always existed (as opposed to having formed by some physical process). We will see shortly how to use the Schwarzschild solution to model physically realistic black holes.
4. Friedmann-Lemaître-Robertson-Walker models

The simplest models of cosmology, the study of the Universe as a whole, are obtained from the assumption that space is homogeneous and isotropic (which is true on average at very large scales). It is well known that the only isotropic 3-dimensional Riemannian metrics are, up to scale, given by

\[ dl^2 = \frac{dr^2}{1 - kr^2} + r^2 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right), \]

where

\[ k = \begin{cases} 
1 & \text{for the standard metric on } S^3 \\
0 & \text{for the standard metric on } \mathbb{R}^3 \\
-1 & \text{for the standard metric on } H^3.
\end{cases} \]

Allowing for a time-dependent scale factor \( a(t) \) (also known as the “radius of the Universe”), we arrive at the Friedmann-Lemaître-Robertson-Walker (FLRW) family of Lorentzian metrics:

\[ ds^2 = -dt^2 + a^2(t) \left[ \frac{dr^2}{1 - kr^2} + r^2 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right) \right]. \]

To interpret these metrics, we consider a general Lorentzian metric of the form

\[ ds^2 = -e^{2\phi} dt^2 + \gamma_{ij} dx^i dx^j = -e^{2\phi} dt^2 + dl^2. \]

The Riemannian metric \( dl^2 \) is readily interpreted as giving the distances measured between nearby observers with fixed space coordinates \( x^i \) in radar experiments: indeed, such observers measure proper time \( \tau \) given by

\[ d\tau^2 = e^{2\phi} dt^2. \]

The null geodesics representing a radar signal bounced by a given observer from a nearby observer (Figure 10) satisfy

\[ ds^2 = 0 \Leftrightarrow e^{2\phi} dt^2 = dl^2 \Leftrightarrow d\tau^2 = dl^2 \Leftrightarrow d\tau = \pm dl. \]

Since the speed of light is \( c = 1 \), the distance traveled between the observers will be half the time between the emission and the reception of the signal:

\[ \frac{(\tau + dl) - (\tau - dl)}{2} = dl. \]

Moreover, the unit timelike covector field tangent to the trajectories of the observers with fixed space coordinates \( x^i \) is

\[ U = -e^{\phi} dt \Leftrightarrow U_\mu = -e^{\phi} \nabla_\mu t. \]
Therefore
\[ \nabla_{U} U^\nu = -U^\nu \nabla_\nu (e^\phi \nabla^\mu t) = (U \cdot \phi) U^\nu - e^\phi U^\nu \nabla_\nu \nabla^\mu t \]
\[ = (U \cdot \phi) U^\nu - e^\phi U^\nu U_\nu \nabla^\mu (e^{-\phi} U^\mu) \]
\[ = (U \cdot \phi) U^\nu - U^\nu U_\nu \nabla^\mu \phi + U^\nu \nabla^\mu U_\nu \]
\[ = \nabla^\mu \phi + (U^\nu \nabla_\nu \phi) U^\mu + \frac{1}{2} \nabla^\mu (U^\nu U_\nu) = \nabla^\mu \phi + (U^\nu \nabla_\nu \phi) U^\mu, \]
since \( U^\nu \nabla_\nu = -1 \). In other words,
\[ \nabla_{U} U = (\text{grad } \phi)^{\perp}, \]
where \( ^\perp \) represents the orthogonal projection on the spacelike hyperplane orthogonal to \( U \).

Therefore the observers with fixed space coordinates in the FLRW models have zero acceleration, that is, they are free-falling (by opposition to the corresponding observers in the Schwarzschild spacetime, who must accelerate to remain at fixed \( r > 2M \)). Moreover, the distance between two such observers varies as
\[ d(t) = a(t) \frac{d_0}{a_0} \Rightarrow \frac{\dot{d}}{a} = \frac{\dot{a}}{a} = \frac{\ddot{a}}{a}. \]
This relation, known as the Hubble law, is often written as
\[ v = H d, \]
where \( v \) is the relative velocity and
\[ H = \frac{\dot{a}}{a} \]
is the so-called Hubble constant (for historical reasons, since it actually varies in time).
We will model the matter content of the universe as an uniform dust of galaxies placed at fixed space coordinates (hence free-falling):

\[ T = \rho(t)dt \otimes dt. \]

Plugging the metric (8) and the energy-momentum tensor (9) into the Einstein equation, and integrating once, results in the so-called Friedmann equations

\[
\begin{align*}
\frac{1}{2} \ddot{a}^2 - \frac{\alpha}{a} - \frac{\Lambda}{6} a^2 &= -\frac{k}{2} \\
\frac{4\pi}{3} \rho a^3 &= \alpha
\end{align*}
\]

(where \( \alpha \) is an integration constant). The first Friedmann equation is a first order ODE for \( a(t) \); it can be seen as the equation of conservation of energy for a particle moving in the 1-dimensional effective potential

\[ V(a) = -\frac{\alpha}{a} - \frac{\Lambda}{6} a^2 \]

with energy \(-\frac{k}{2}\). Once this equation has been solved, the second Friedmann equation yields \( \rho(t) \) from \( a(t) \). We now examine in details the FLRW models arising from the solutions of these equations.

### 4.1. Milne universe.

If we set \( \alpha = \Lambda = 0 \) then the first Friedmann equation becomes

\[ \ddot{a}^2 = -k. \]

Therefore either \( k = 0 \) and \( \dot{a} = 0 \), which corresponds to the Minkowski spacetime, or \( k = -1 \) and \( \dot{a}^2 = 1 \), that is

\[ ds^2 = -dt^2 + t^2 d\tilde{H}^3, \]

where \( d\tilde{H}^3 \) represents the metric of the unit hyperbolic 3-space; this is the so-called Milne universe. It turns out that the Milne universe is isometric to an open region of the Minkowski spacetime, namely the region limited by the future (or past) light cone of the origin. This region is foliated by hyperboloids \( \Sigma_t \) of the form

\[ u^2 - x^2 - y^2 - z^2 = t^2, \]

whose induced metric is that of a hyperbolic space of radius \( t \) (Figure 11). Note that the light cone corresponds to \( a(t) = t = 0 \), that is, the Big Bang of the Milne universe.

### 4.2. de Sitter universe.

If \( \alpha = 0 \) and \( \Lambda > 0 \) we can choose units such that \( \Lambda = 3 \). The first Friedmann equation then becomes

\[ \dot{a}^2 - a^2 = -k. \]
In this case all three values $k = 1$, $k = 0$ and $k = -1$ are possible; the corresponding metrics are, respectively,

$$ds^2 = -dt^2 + \cosh^2 t \, dl_{S^3}^2;$$
$$ds^2 = -dt^2 + e^{2t} \, dl_{\mathbb{R}^3}^2;$$
$$ds^2 = -dt^2 + \sinh^2 t \, dl_{H^3}^2,$$

where $dl_{S^3}^2$, $dl_{\mathbb{R}^3}^2$ and $dl_{H^3}^2$ represent the metric of the unit 3-sphere, the Euclidean 3-space and the unit hyperbolic 3-space.

It turns out that the last two models correspond to open regions of the first, which is then called the de Sitter universe. It represents a spherical universe which contracts to a minimum radius (1 in our units) and then re-expands. It is easily seen to be isometric to the unit hyperboloid

$$-u^2 + x^2 + y^2 + z^2 + w^2 = 1$$

in the Minkowski 5-dimensional spacetime (Figure 12).

To obtain the Penrose diagram for the de Sitter universe we write its metric as

$$ds^2 = -dt^2 + \cosh^2 t \left[ d\psi^2 + \sin^2 \psi \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right) \right]$$
$$= \cosh^2 t \left[ -d\tau^2 + d\psi^2 + \sin^2 \psi \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right) \right]$$
$$= \cosh^2 t \left[ -d\tau^2 + d\psi^2 \right] + \rho^2 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right),$$

where $\psi \in [0, \pi]$,

$$\tau = \int_{-\infty}^{t} \frac{dt}{\cosh t}.$$
4. FRIEDMANN-LEMAÎTRE-ROBERTSON-WALKER MODELS

Figure 12. de Sitter universe.

and

\[ r = \cosh t \sin \psi. \]

Since

\[ \int_{-\infty}^{+\infty} \frac{dt}{\cosh t} = \pi, \]

we see that the quotient metric is conformal to the square \((0, \pi) \times [0, \pi]\) of the Minkowski 2-dimensional spacetime, and so the Penrose diagram is as depicted in Figure 12. Note that there are two lines where \(r = 0\), corresponding to two antipodal points of the 3-sphere. A light ray emitted from one of these points at \(t = -\infty\) has just enough time to reach the other at \(t = +\infty\) (dashed line in the diagram). Note also that in this case \(\mathcal{I}^-\) and \(\mathcal{I}^+\) (defined as the past and future boundary points approached by null geodesics along which \(r \to +\infty\)) are spacelike boundaries.

4.3. Anti-de Sitter universe. If \(\alpha = 0\) and \(\Lambda < 0\) we can choose units such that \(\Lambda = -3\). The first Friedmann equation then becomes

\[ \dot{a}^2 + a^2 = -k. \]

In this case only \(k = -1\) is possible; the corresponding metric is

\[ ds^2 = -dt^2 + \cos^2 t \, d\ell_{H^3}^2. \]

It turns out (as we shall see in Chapter ??) that this model is an open region of the spacetime with metric

\[ ds^2 = -\cosh^2 \psi \, dt^2 + d\psi^2 + \sinh^2 \psi \left( d\theta^2 + \sin^2 \theta \, d\varphi^2 \right) \]

(where \(\psi \in [0, +\infty)\)), called the anti-de Sitter universe. It represents a static hyperbolic universe (with radius 1 in our units).
To obtain the Penrose diagram for the anti-de Sitter universe we write its metric as
\[ ds^2 = \cosh^2 \psi \left[ -dt^2 + \frac{d\psi^2}{\cosh^2 \psi} \right] + \sinh^2 \psi \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right) \]
\[ = \cosh^2 \psi \left[ -dt^2 + dx^2 \right] + r^2 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right) \]
where
\[ x = \int_0^\psi \frac{d\psi}{\cosh \psi} \]
and
\[ r = \sinh \psi. \]
Since
\[ \int_0^{+\infty} \frac{d\psi}{\cosh \psi} = \pi, \]
we see that the quotient metric is conformal to the strip \( \mathbb{R} \times [0, \pi] \) of the Minkowski 2-dimensional spacetime, and so the Penrose diagram is as depicted in Figure 14. The FLRW model above corresponds to the triangular region in the diagram. Note also that in this case \( I^- \equiv J^+ \equiv J \) is a timelike boundary.

4.4. Universes with matter and \( \Lambda = 0 \). If \( \alpha > 0 \) and \( \Lambda = 0 \), the first Friedmann equation becomes
\[ \dot{a}^2 - \frac{2\alpha}{a} = -k. \]
In this case all three values \( k = 1 \), \( k = 0 \) and \( k = -1 \) are possible. Although it is possible to obtain explicit formulas for the solutions of these equations, it is simpler to analyze the graph of the effective potential \( V(a) \) (Figure 15). Possibly by reversing and translating \( t \), we can assume that all solutions are defined for \( t > 0 \), with \( \lim_{t \to 0} a(t) = 0 \), implying \( \lim_{t \to 0} \rho(t) = +\infty \). Therefore all three models have a true singularity at \( t = 0 \), known as the
4. FRIEDMANN-LEMAÎTRE-ROBERTSON-WALKER MODELS

\[ r = 0 \]

**Figure 14.** Penrose diagram for the anti-de Sitter universe.

**Big Bang**, where the scalar curvature \( R = 8\pi\rho \) also blows up; this is not true for the Milne universe or the open region in the anti-de Sitter universe, which can be extended across the Big Bang. The spherical universe \( (k = 1) \) reaches a maximum radius \( 2\alpha \) and re-collapses, forming a second singularity (the **Big Crunch**); the radius of the flat \( (k = 0) \) and hyperbolic \( (k = -1) \) universes increases monotonically.

To obtain the Penrose diagram for the spherical universe we write its metric as

\[
\begin{align*}
 ds^2 &= -dt^2 + a^2(t) \left[ d\psi^2 + \sin^2 \psi \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right) \right] \\
 &= a^2(t) \left[ -d\tau^2 + d\psi^2 + \sin^2 \psi \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right) \right] \\
 &= a^2(t) \left[ -d\tau^2 + d\psi^2 \right] + r^2 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right),
\end{align*}
\]

where \( \psi \in [0, \pi] \),

\[
\tau = \int_0^t \frac{dt}{a(t)}
\]

and

\[
r = a(t) \sin \psi.
\]

Since

\[
\int_0^{t_{\text{max}}} \frac{dt}{a(t)} = 2 \int_0^{a_{\text{max}}} \frac{da}{a\dot{a}} = 2 \int_0^{2\alpha} \frac{da}{a} \sqrt{\frac{2\alpha}{a} - 1} = 2\pi,
\]
we see that the quotient metric is conformal to the rectangle \((0, 2\pi) \times [0, \pi]\) of the Minkowski 2-dimensional spacetime, and so the Penrose diagram is as depicted in Figure 16. Note that there are two lines where \(r = 0\), corresponding to two antipodal points of the 3-sphere. A light ray emitted from one of these points at \(t = 0\) has just enough time to circle once around the universe and return at \(t = t_{\text{max}}\) (dashed line in the diagram). Note also that the Big Bang and the Big Crunch are spacelike boundaries.

To obtain the Penrose diagram for the flat universe we write its metric as

\[
\begin{align*}
\frac{ds^2}{a^2} &= -dt^2 + a^2(t) \left[ d\rho^2 + \rho^2 \left( d\psi^2 + \sin^2 \theta d\varphi^2 \right) \right] \\
&= a^2(t) \left[ -d\tau^2 + d\rho^2 + \rho^2 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right) \right] \\
&= a^2(t) \left[ -d\tau^2 + d\rho^2 \right] + r^2 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right),
\end{align*}
\]

where \(\rho \in [0, +\infty)\) and \(r = a(t)\rho\). Since

\[
\int_0^{+\infty} \frac{dt}{a(t)} = \int_0^{+\infty} \frac{da}{a} = \int_0^{+\infty} \frac{da}{a \sqrt{2a}} = +\infty,
\]

we see that the quotient metric is conformal to the region \((0, +\infty) \times [0, +\infty)\) of the Minkowski 2-dimensional spacetime, and so the Penrose diagram is as depicted in Figure 17. Note that the Big Bang is a spacelike boundary.

The Penrose diagram for the hyperbolic universe turns out to be the same as for the flat universe. To see this we write its metric as

\[
\begin{align*}
\frac{ds^2}{a^2} &= -dt^2 + a^2(t) \left[ d\psi^2 + \sinh^2 \psi \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right) \right] \\
&= a^2(t) \left[ -d\tau^2 + d\psi^2 + \sinh^2 \psi \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right) \right] \\
&= a^2(t) \left[ -d\tau^2 + d\psi^2 \right] + r^2 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right),
\end{align*}
\]
4. FRIEDMANN-LEMAITRE-ROBERTSON-WALKER MODELS

where $\psi \in [0, +\infty)$ and $r = a(t) \sinh \psi$, and note that

$$
\int_0^{+\infty} \frac{dt}{a(t)} = \int_0^{+\infty} \frac{da}{a\dot{a}} = \int_0^{+\infty} \frac{da}{a\sqrt{\frac{2a}{2a} + 1}} = +\infty.
$$
4.5. **Universes with matter and** $\Lambda > 0$. If $\alpha > 0$ and $\Lambda > 0$ we can choose units such that $\Lambda = 3$. The first Friedmann equation then becomes

$$\dot{a}^2 - \frac{2\alpha}{a} - a^2 = -k.$$  

In this case all three values $k = 1$, $k = 0$ and $k = -1$ are possible. As before, we analyze the graph of the effective potential $V(a)$ (Figure 18). The hyperbolic and flat universes behave qualitatively like when $\Lambda = 0$, although $\dot{a}(t)$ is now unbounded as $t \to +\infty$, instead of approaching some constant. The spherical universe has a richer spectrum of possible behaviors, depending on $\alpha$, represented in Figure 18 by drawing the line of constant energy $-k = -1$ at three different heights. The higher line (corresponding to $\alpha > \sqrt{3}$) yields a behaviour similar to that of the hyperbolic and flat universes. The intermediate line (corresponding to $\alpha = \frac{\sqrt{3}}{3}$) gives rise to an unstable equilibrium point $a = \sqrt[3]{3}$, where the attraction force of the matter is balanced by the repulsion force of the cosmological constant; it corresponds to the so-called **Einstein universe**, the first cosmological model ever proposed. The intermediate line also yields two solutions asymptotic to the Einstein universe, one containing a Big Bang and the other endless expansion. Finally, the lower line (corresponding to $\alpha < \frac{\sqrt{3}}{3}$) yields two different types of behaviour (depending on the initial conditions): either similar to the spherical model with $\Lambda = 0$, or to the de Sitter universe.

![Figure 18. Effective potential for FLRW models with $\Lambda > 0$.](image-url)

It is currently believed that the best model for our physical Universe is the flat universe with $\Lambda > 0$. If we write the first Friedmann equation as

$$H^2 \equiv \frac{\dot{a}^2}{a^2} = \frac{8\pi}{3} \rho + \frac{\Lambda}{3}.$$
then the terms on the right-hand side are in the proportion 2 : 5 at the present time.

4.6. Universes with matter and \( \Lambda < 0 \). If \( \alpha > 0 \) and \( \Lambda < 0 \) we can choose units such that \( \Lambda = -3 \). The first Friedmann equation then becomes

\[
\dot{a}^2 - \frac{2\alpha}{a} + a^2 = -k.
\]

In this case all three values \( k = 1 \), \( k = 0 \) and \( k = -1 \) are possible. As before, we analyze the graph of the effective potential \( V(a) \) (Figure 19).

The qualitative behaviour of the hyperbolic, flat and spherical universes is the same as the spherical universe with \( \Lambda = 0 \), namely starting at a Big Bang and ending at a Big Crunch.

\[
V(a)
\]

\[
\begin{align*}
&k = -1 \\
&k = 0 \\
&k = 1
\end{align*}
\]

\( a \)

**Figure 19.** Effective potential for FLRW models with \( \Lambda > 0 \).

5. Matching

Let \( (M_1, g_1) \) and \( (M_2, g_2) \) be solutions of the Einstein field equation containing open sets \( U_1 \) and \( U_2 \) whose boundaries \( \Sigma_1 \) and \( \Sigma_2 \) are timelike hypersurfaces, that is, hypersurfaces whose induced metric is Lorentzian (or, equivalently, whose normal vector is spacelike). If \( \Sigma_1 \) is diffeomorphic to \( \Sigma_2 \) then we can identify them to obtain a new manifold \( M \) gluing \( U_1 \) to \( U_2 \) along \( \Sigma_1 \cong \Sigma_2 \) (Figure 20).

Let \( n \) be the unit normal vector to \( \Sigma_1 \) pointing out of \( U_1 \), which we identify with the unit normal vector to \( \Sigma_2 \) pointing into \( U_2 \). If \( (x^1, x^2, x^3) \) are local coordinates on \( \Sigma \cong \Sigma_1 \cong \Sigma_2 \), we can construct a system of local coordinates \( (t, x^1, x^2, x^3) \) in a neighbourhood of \( \Sigma \) by moving a distance \( t \) along the geodesics with initial condition \( n \). Note that \( U_1, \Sigma \) and \( U_2 \)
correspond to $t < 0$, $t = 0$ and $t > 0$ in these coordinates. Since $\frac{\partial}{\partial t}$ is the unit tangent vector to the geodesics, we have

\[
\frac{\partial}{\partial t} \left( \frac{\partial}{\partial t} \cdot \frac{\partial}{\partial x^i} \right) = \left\langle \nabla_{\frac{\partial}{\partial t}} \frac{\partial}{\partial x^i}, \frac{\partial}{\partial t} \right\rangle + \left\langle \frac{\partial}{\partial t}, \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial t} \right\rangle
\]

\[
= \left\langle \frac{\partial}{\partial t}, \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial t} \right\rangle = \frac{\partial}{\partial x^i} \left( \frac{1}{2} \left\langle \frac{\partial}{\partial t}, \frac{\partial}{\partial t} \right\rangle \right) = 0
\]

($i = 1, 2, 3$), where we used

\[
\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial t} - \nabla_{\frac{\partial}{\partial t}} \frac{\partial}{\partial x^i} = \left[ \frac{\partial}{\partial t}, \frac{\partial}{\partial x^i} \right] = 0.
\]

Since for $t = 0$ we have

\[
\left\langle \frac{\partial}{\partial t}, \frac{\partial}{\partial x^i} \right\rangle = \left\langle n, \frac{\partial}{\partial x^i} \right\rangle = 0,
\]

we see that $\frac{\partial}{\partial t}$ remains orthogonal to the surfaces of constant $t$. This result will be used repeatedly.

**Lemma 5.1.** (Gauss Lemma) Let $(M, g)$ be a Riemannian or Lorentzian manifold, and $\Sigma \subset M$ a hypersurface whose normal vector field $n$ satisfies $g(n, n) \neq 0$. The hypersurfaces $\Sigma_t$ obtained from $\Sigma$ by moving a distance $t$ along the geodesics orthogonal to $\Sigma$ remain orthogonal to the geodesics.
In this coordinate system the metrics $g_1$ and $g_2$ are given on $t \leq 0$ and $t \geq 0$, respectively, by
\[ g_A = dt^2 + h_{ij}^a(t, x)dx^idx^j \]
($A = 1, 2$). Therefore we can define a continuous metric $g$ on $M$ if
\[ h_{ij}^1(0, x)dx^idx^j = h_{ij}^2(0, x)dx^idx^j, \]
that is, if
\[ g_1|_{\Sigma} = g_2|_{\Sigma}. \]
This also guarantees continuity of all tangential derivatives of the metric, but not of the normal derivatives. In order to have a $C^1$ metric we must have
\[ \frac{\partial h_{ij}^1}{\partial t}(0, x)dx^idx^j = \frac{\partial h_{ij}^2}{\partial t}(0, x)dx^idx^j, \]
that is,
\[ \mathcal{L}_n g_1|_{\Sigma} = \mathcal{L}_n g_2|_{\Sigma}. \]
Note that in this case the curvature tensor (hence the energy-momentum tensor) is at most discontinuous across $\Sigma$. More importantly, as we will see in Chapter ??, the components $T_{\mu\nu}$ of the energy-momentum tensor are continuous across $\Sigma$ (that is, the flow of energy and momentum across $\Sigma$ is equal on both sides), implying that the energy-momentum tensor satisfies the integrated version of the conservation equation $\nabla_{\mu} T_{\mu\nu} = 0$. Therefore we can consider $(M, g)$ a solution of the Einstein equations.

Recall that
\[ K_A = \frac{1}{2} \mathcal{L}_n g_A|_{\Sigma}, \]
is known as the **extrinsic curvature**, or **second fundamental form**, of $\Sigma_A$. We can summarize the discussion above in the following statement.

**Proposition 5.2.** Two solutions $(M_1, g_1)$ and $(M_2, g_2)$ of the Einstein field equation can be matched along diffeomorphic timelike boundaries $\Sigma_1$ and $\Sigma_2$ if and only if the induced metrics and second fundamental forms coincide:
\[ g_1 = g_2 \quad \text{and} \quad K_1 = K_2. \]

6. Oppenheimer-Snyder collapse

We can use the matching technique to construct a solution of the Einstein field equation which describes a spherical cloud of dust collapsing to a black hole. This is a physically plausible model for a black hole, as opposed to the eternal black hole.

Let us take $(M_1, g_1)$ to be a flat collapsing FLRW universe:
\[ g_1 = -d\tau^2 + a^2(\tau) \left[ d\sigma^2 + \sigma^2 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right) \right]. \]
We choose $\Sigma_1$ to be the hypersurface $\sigma = \sigma_0$, with normal vector
\[ n = \frac{1}{a} \frac{\partial}{\partial \sigma}. \]
The induced metric then is
\[ g_{1|T\Sigma_1} = -d\tau^2 + a^2(\tau)\sigma_0^2 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right), \]
and the second fundamental form
\[ K_1 = a(\tau)\sigma_0 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right). \]
Here we used **Cartan’s magic formula**: if \( \omega \) is a differential form and \( X \) is a vector field then
\[ \mathcal{L}_X \omega = X \lrcorner d\omega + d(X \lrcorner \omega). \]
Thus, for example,
\[ \mathcal{L}_n d\tau = n \lrcorner d^2\tau + d(n \lrcorner d\tau) = 0. \]
Note that the function \( a(\tau) \) is constrained by the first Friedmann equation:
\[ (10) \quad \dot{a}^2 = \frac{2\alpha}{a} \]
(we assume \( \Lambda = 0 \)).

We now take \( (M_2, g_2) \) to be the Schwarzschild solution,
\[ g_2 = -V dt^2 + V^{-1} dr^2 + r^2 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right), \]
where
\[ V = 1 - \frac{2M}{r}, \]
and choose \( \Sigma_2 \) to be a spherically symmetric timelike hypersurface given by the parameterization
\[ \begin{cases} 
  t = t(\tau) \\
  r = r(\tau) 
\end{cases}. \]
The exact form of the functions \( t(\tau) \) and \( r(\tau) \) will be fixed by the matching conditions; for the time being they are constrained only by the condition that \( \tau \) is the proper time along \( \Sigma_2 \):
\[ -V \dot{t}^2 + V^{-1} \dot{r}^2 = -1. \]
The induced metric is then
\[ g_{2|T\Sigma_2} = -d\tau^2 + r^2(\tau) \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right), \]
and so the two induced metrics coincide if and only if
\[ (11) \quad r(\tau) = a(\tau)\sigma_0. \]
Since \( \Sigma_1 \) is ruled by timelike geodesics, so is \( \Sigma_2 \), because the induced metrics and extrinsic curvatures are the same (therefore so are the Christoffel symbols). Therefore \( t(\tau) \) and \( r(\tau) \) must be a solution of the radial geodesic equations, which are equivalent to
\[ (12) \quad \begin{cases} 
  V \dot{t} = E \\
  -V \dot{t}^2 + V^{-1} \dot{r}^2 = -1 
\end{cases} \quad \Leftrightarrow \quad \begin{cases} 
  \dot{t} = E \\
  \dot{r}^2 = E^2 - 1 + \frac{2M}{r} 
\end{cases} \]
(where \( E > 0 \) is a constant). Equations (10), (11) and (12) are compatible if and only if
\[
\begin{cases}
E = 1 \\
M = \alpha \sigma_0^3
\end{cases}
\]
that is, if and only if \( \Sigma_2 \) represents a spherical shell dropped from infinity with zero velocity and the mass parameter of the Schwarzschild spacetime is related to the density \( \rho(\tau) \) of the collapsing dust by
\[
M = \frac{4\pi}{3} r^3(\tau) \rho(\tau).
\]
To compute the second fundamental form of \( \Sigma_2 \) we consider a family of free-falling spherical shells which includes \( \Sigma_2 \). If \( s \) is the parameter indexing the shells (with \( \Sigma_2 \) corresponding to, say, \( s = 0 \)) then by the Gauss Lemma (used twice) we can write the Schwarzschild metric in the form
\[
g_2 = -d\tau^2 + A^2(\tau, s) ds^2 + r^2(\tau, s) \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right).
\]
The unit normal vector field \( n \) to the hypersurfaces of constant \( s \) is then
\[
n = \frac{1}{A} \frac{\partial}{\partial s},
\]
and so we have
\[
K_2 = \frac{1}{2} \partial_n g_{2|\tau=\Sigma_2} = r(n \cdot r) \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right).
\]
On the other hand, in Schwarzschild coordinates
\[
n = V^{-1} r \frac{\partial}{\partial t} + V i \frac{\partial}{\partial r},
\]
since \( n \) must be unit and orthogonal to
\[
i \frac{\partial}{\partial t} + r \frac{\partial}{\partial r}.
\]
Therefore we have
\[
K_2 = rV i \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right) = rE \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right),
\]
or, using \( E = 1 \),
\[
K_2 = r(\tau) \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right).
\]
In other words, \( K_1 = K_2 \) follows from the previous conditions, and so we indeed have a solution of the Einstein equations.

To construct the Penrose diagram for this solution we represent \( \Sigma_1 \) and \( \Sigma_2 \) in the Penrose diagrams of the collapsing flat FLRW universe (obtained by reversing the time direction in the expanding flat FLRW universe) and the Schwarzschild solution (Figure 21). Identifying these hypersurfaces results in the Penrose diagram depicted in Figure 22.
2. EXACT SOLUTIONS

Figure 21. Matching hypersurfaces in the collapsing flat FLRW universe and the Schwarzschild solution.

Figure 22. Penrose diagram for the Oppenheimer-Snyder collapse.

7. Exercises

(1) In this exercise we will solve the vacuum Einstein equation (without cosmological constant) for the spherically symmetric Lorentzian metric given by

\[ ds^2 = -(A(t,r))^2 dt^2 + (B(t,r))^2 dr^2 + r^2 (d\theta^2 + r^2 \sin^2 \theta d\varphi^2), \]

where \( A \) and \( B \) are positive smooth functions.
(a) Use Cartan’s first structure equations,

\[
\begin{align*}
\omega_{\mu\nu} &= -\omega_{\nu\mu} \\
\omega_{\mu} - \omega_{\nu} \wedge \omega_{\nu} &= 0
\end{align*}
\]

to show that the nonvanishing connection forms for the orthonormal frame dual to

\[
\begin{align*}
\omega^0 &= Adt, \quad \omega^r &= Bdr, \quad \omega^\theta &= r d\theta, \quad \omega^\phi &= r \sin \theta d\phi
\end{align*}
\]

are (using the notation \(\dot{\cdot} = \frac{\partial}{\partial t}\) and \(\cdot' = \frac{\partial}{\partial r}\))

\[
\begin{align*}
\omega^0_r &= \omega^r_0 = \frac{A'}{B} dt + \frac{\dot{B}}{A} dr; \\
\omega^\theta_r &= -\omega^r_\theta = \frac{1}{B} d\theta; \\
\omega^\phi_r &= -\omega^r_\phi = \frac{\sin \theta}{B} d\phi; \\
\omega^\phi_\theta &= -\omega^\theta_\phi = \cos \theta d\phi.
\end{align*}
\]

(b) Use Cartan’s second structure equations,

\[
\Omega^\mu_{\nu} = d\omega^\mu_{\nu} + \omega^\mu_{\alpha} \wedge \omega^\alpha_{\nu},
\]

to show that the curvature forms on this frame are

\[
\begin{align*}
\Omega^0_r &= \Omega^r_0 = \left( \frac{A'^3B - A'B'^3}{AB^3} + \frac{\dot{A}B - A\dot{B}}{A^3B} \right) \omega^r \wedge \omega^0; \\
\Omega^0_\theta &= \Omega^\theta_0 = \frac{A'}{rAB^2} \omega^\theta \wedge \omega^0 + \frac{\dot{B}}{rAB^2} \omega^\theta \wedge \omega^r; \\
\Omega^0_\phi &= \Omega^\phi_0 = \frac{A'}{rAB^2} \omega^\phi \wedge \omega^0 + \frac{\dot{B}}{rAB^2} \omega^\phi \wedge \omega^r; \\
\Omega^r_\theta &= -\Omega^\theta_\phi = \frac{B'}{rB^3} \omega^\phi \wedge \omega^r + \frac{\dot{B}}{rAB^2} \omega^\theta \wedge \omega^0; \\
\Omega^r_\phi &= -\Omega^\phi_\theta = \frac{B'}{rB^3} \omega^\theta \wedge \omega^r + \frac{\dot{B}}{rAB^2} \omega^\phi \wedge \omega^0; \\
\Omega^\phi_\theta &= -\Omega^\theta_\phi = \frac{B^2 - 1}{r^2B^2} \omega^\phi \wedge \omega^\theta.
\end{align*}
\]

(c) Using

\[
\Omega^\mu_{\nu} = \sum_{\alpha<\beta} R_{\alpha\beta}^\mu \omega^\alpha \wedge \omega^\beta,
\]
determine the components $R_{\alpha \beta \nu}^\mu$ of the curvature tensor in this orthonormal frame, and show that the nonvanishing components of the Ricci tensor in this frame are

$$R_{00} = \frac{A''B - A'B'}{AB^3} + \frac{\dot{A}\dot{B} - A\ddot{B}}{A^3B} + \frac{2A'}{rAB^2};$$

$$R_{0r} = R_{r0} = \frac{2\dot{B}}{rAB^2};$$

$$R_{rr} = \frac{A'B' - A''B}{AB^3} + \frac{\ddot{A}B - A\dot{B}}{A^3B} + \frac{2B'}{rB^3};$$

$$R_{\theta\theta} = R_{\varphi\varphi} = -\frac{A'}{rAB^2} + \frac{B'}{rB} + \frac{B^2 - 1}{r^2B^2}.$$

Conclude that the nonvanishing components of the Einstein tensor in this frame are

$$G_{00} = \frac{2B'}{rB^3} + \frac{B^2 - 1}{r^2B^2};$$

$$G_{0r} = G_{r0} = \frac{2\dot{B}}{rAB^2};$$

$$G_{rr} = \frac{2A'}{rAB^2} - \frac{B^2 - 1}{r^2B^2};$$

$$G_{\theta\theta} = G_{\varphi\varphi} = \frac{A''B - A'B'}{AB^3} + \frac{\dot{A}\dot{B} - A\ddot{B}}{A^3B} + \frac{2A'}{rAB^2} - \frac{B'}{rB^3}.$$

(d) Show that if we write

$$B(t, r) = \left(1 - \frac{2m(t, r)}{r}\right)^{-\frac{1}{2}}$$

for some smooth function $m$ then

$$G_{00} = \frac{2m'}{r^2}.$$ 

Conclude that the Einstein equations $G_{00} = G_{0r} = 0$ are equivalent to

$$B = \left(1 - \frac{2M}{r}\right)^{-\frac{1}{2}},$$

where $M \in \mathbb{R}$ is an integration constant.

(e) Show that the Einstein equation $G_{00} + G_{rr} = 0$ is equivalent to $A = \frac{a(t)}{B}$ for some positive smooth function $a$.

(f) Check that if $A$ and $B$ are as above then the remaining Einstein equations $G_{\theta\theta} = G_{\varphi\varphi} = 0$ are automatically satisfied.

(g) Argue that it is always possible to rescale the time coordinate $t$ so that the metric is written

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \left(1 - \frac{2M}{r}\right)^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2)$$
(the statement that any spherically symmetric solution of the vacuum Einstein equation without cosmological constant is of this form is known as Birkhoff’s theorem).

(2) Show that the Riemannian manifold obtained by gluing the hypersurfaces \( t = 0 \) of the two exterior regions in the maximally extended Schwarzschild solution along the horizon \( r = 2M \) is isometric to the Flamm paraboloid, that is, the hypersurface in \( \mathbb{R}^4 \) with equation

\[
\sqrt{x^2 + y^2 + z^2} = 2M + \frac{w^2}{8M}
\]

(Figure 23).

\[\text{Figure 23. Two-dimensional analogue of the Flamm paraboloid.}\]

(3) Recall that the nonvanishing components of the Einstein tensor of the spherically symmetric Lorentzian metric

\[
ds^2 = -(A(t, r))^2 dt^2 + (B(t, r))^2 dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\varphi^2)
\]

in the orthonormal frame dual to

\[
\omega^0 = Adt, \quad \omega^r = Bdr, \quad \omega^\theta = rd\theta, \quad \omega^\varphi = r \sin \theta d\varphi,
\]

are given by (using the notation \( \dot{=} = \frac{\partial}{\partial t} \) and \( \prime = \frac{\partial}{\partial r} \))

\[
G_{00} = \frac{2B'}{rB^3} + \frac{B^2 - 1}{r^2 B^2} = \frac{2m'}{r^2};
\]

\[
G_{0r} = G_{r0} = \frac{2B}{rAB^2};
\]

\[
G_{rr} = \frac{2A'}{rAB^2} - \frac{B^2 - 1}{r^2 B^2};
\]

\[
G_{\theta\theta} = G_{\varphi\varphi} = \frac{A''B - A'B'}{AB^3} + \frac{\dot{A}B - \dot{A}\dot{B}}{A^3B} + \frac{A'}{rAB^2} - \frac{B'}{rB^3};
\]
where
\[ B(t,r) = \left( 1 - \frac{2m(t,r)}{r} \right)^{-\frac{1}{2}}. \]

(a) Assuming
- \( G_{0r} = 0 \) (so that \( B \), and hence \( m \), do not depend on \( t \));
- \( G_{00} + G_{rr} = 0 \) (so that \( A = \frac{\alpha(t)}{r} \) for some positive smooth function \( \alpha(t) \));
- \( \alpha(t) = 1 \) (which can always be achieved by rescaling \( t \)),
show that
\[ G_{\theta\theta} = G_{\varphi\varphi} = \frac{1}{2} \left( A^2 \right)'' + \frac{1}{r} \left( A^2 \right)'. \]

(b) Prove that the general spherically symmetric solution of the vacuum Einstein field equation with a cosmological constant \( \Lambda \) is the Kottler metric
\[ g = -\left( 1 - \frac{2M}{r} - \frac{\Lambda}{3} r^2 \right) dt^2 + \left( 1 - \frac{2M}{r} - \frac{\Lambda}{3} r^2 \right)^{-1} dr^2 + r^2 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right). \]

(c) Obtain the Penrose diagram for the maximal extension of the Kottler solution with \( \Lambda > 0 \) and \( 0 < M < \frac{1}{3\sqrt{\Lambda}} \).

(d) Consider now the spherically symmetric electromagnetic field
\[ F = E(t,r) \omega^r \wedge \omega^0. \]
Show that this field satisfies the vacuum Maxwell equations
\[ dF = d \star F = 0 \]
(where \( \star \) is the Hodge star) if and only if
\[ E(t,r) = \frac{e}{r^2} \]
for some constant \( e \in \mathbb{R} \) (in geometrized units, where \( \varepsilon_0 = 1 \), the electric charge is \( 4\pi e \)).

(e) As we shall see in Chapter ??, this electromagnetic field corresponds to the energy-momentum tensor
\[ T = \frac{E^2}{2} \left( \omega^0 \otimes \omega^0 - \omega^r \otimes \omega^r + \omega^\theta \otimes \omega^\theta + \omega^\varphi \otimes \omega^\varphi \right). \]
Prove that the general spherically symmetric solution of the Einstein field equation with an electromagnetic field of this kind is the Reissner-Nordström metric
\[ g = -\left( 1 - \frac{2M}{r} + \frac{e^2}{r^2} \right) dt^2 + \left( 1 - \frac{2M}{r} + \frac{e^2}{r^2} \right)^{-1} dr^2 + r^2 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right). \]

(f) Obtain the Penrose diagram for the maximal extension of the Reissner-Nordström solution with \( M > 0 \) and \( 0 < e^2 < M^2 \).
(4) Consider the spherically symmetric Lorentzian metric
\[ ds^2 = -dt^2 + a^2(t) \left( \frac{1}{1 - kr^2} dr^2 + r^2 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right) \right), \]
where \( a \) is a positive smooth function.

(a) Use Cartan’s first structure equations,
\[
\begin{align*}
\omega_{\mu\nu} &= -\omega_{\nu\mu} \\
d\omega^\mu + \omega^\mu_{\alpha} \wedge \omega^\alpha &= 0
\end{align*}
\]
to show that the nonvanishing connection forms for the orthonormal frame dual to
\[
\begin{align*}
\omega^0 &= dt, \\
\omega^r &= a(t) \left( 1 - kr^2 \right)^{-\frac{1}{2}} dr, \\
\omega^\theta &= a(t) rdr, \\
\omega^\varphi &= a(t) r \sin \theta d\varphi
\end{align*}
\]
are
\[
\begin{align*}
\omega^0_r &= \omega^r_0 = \dot{a} \left( 1 - kr^2 \right)^{-\frac{1}{2}} dr; \\
\omega^0_\theta &= \omega^\theta_0 = \dot{a} r d\theta; \\
\omega^0_\varphi &= \omega^\varphi_0 = \dot{a} r \sin \theta d\varphi; \\
\omega^r_\theta &= -\omega^\theta_r = \left( 1 - kr^2 \right)^{\frac{3}{2}} d\theta; \\
\omega^r_\varphi &= -\omega^\varphi_r = \left( 1 - kr^2 \right)^{\frac{1}{2}} \sin \theta d\varphi; \\
\omega^\theta_\varphi &= -\omega^\varphi_\theta = \cos \theta d\varphi.
\end{align*}
\]

(b) Use Cartan’s second structure equations,
\[
\Omega^\mu_{\nu} = d\omega^\mu_{\nu} + \omega^\mu_{\alpha} \wedge \omega^\alpha_{\nu},
\]
to show that the curvature forms on this frame are
\[
\begin{align*}
\Omega^0_r &= \Omega^r_0 = \frac{\dot{a}}{a} \omega^0 \wedge \omega^r; \\
\Omega^0_\theta &= \Omega^\theta_0 = \frac{\dot{a}}{a} \omega^0 \wedge \omega^\theta; \\
\Omega^0_\varphi &= \Omega^\varphi_0 = \frac{\dot{a}}{a} \omega^0 \wedge \omega^\varphi; \\
\Omega^r_\theta &= -\Omega^\theta_r = \left( \frac{k}{a^2} + \frac{\dot{a}^2}{a^2} \right) \omega^\theta \wedge \omega^r; \\
\Omega^r_\varphi &= -\Omega^\varphi_r = \left( \frac{k}{a^2} + \frac{\dot{a}^2}{a^2} \right) \omega^\varphi \wedge \omega^r; \\
\Omega^\theta_\varphi &= -\Omega^\varphi_\theta = \left( \frac{k}{a^2} + \frac{\dot{a}^2}{a^2} \right) \omega^\varphi \wedge \omega^\theta.
\end{align*}
\]

(c) Using
\[
\Omega^\mu_{\nu} = \sum_{\alpha<\beta} R^\mu_{\alpha\beta} \omega^\alpha \wedge \omega^\beta,
\]
determine the components $R_{\alpha\beta}^{\mu\nu}$ of the curvature tensor on this orthonormal frame, and show that the nonvanishing components of the Ricci tensor on this frame are

$$R_{00} = -\frac{3\dot{a}}{a};$$

$$R_{rr} = R_{\theta\theta} = R_{\phi\phi} = \frac{\ddot{a}}{a} + \frac{2\dot{a}^2}{a^2} + \frac{2k}{a^2}.$$  

Conclude that the nonvanishing components of the Einstein tensor on this frame are

$$G_{00} = \frac{3\dot{a}^2}{a^2} + \frac{3k}{a^2};$$

$$G_{rr} = G_{\theta\theta} = G_{\phi\phi} = -\frac{2\ddot{a}}{a} - \frac{2\dot{a}^2}{a^2} - \frac{k}{a^2}.$$  

(d) Show that the Einstein equation with a cosmological constant $\Lambda$ for a comoving pressureless perfect fluid of nonnegative density $\rho$, $G + \Lambda g = 8\pi\rho dt^2$, is equivalent to the system

$$\begin{cases}
\frac{\dot{a}^2}{a^2} + \frac{k}{a^2} = \frac{8\pi\rho}{3} + \frac{\Lambda}{3} \\
\frac{2\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{k}{a^2} = \Lambda
\end{cases}.$$  

Show that this system can be integrated to

$$\begin{cases}
\frac{4\pi\rho}{3}a^3 = \alpha \\
\frac{1}{2}\dot{a}^2 - \frac{\alpha}{a} - \frac{\Lambda}{6}a^2 = -\frac{k}{2}
\end{cases},$$  

where $\alpha$ is a nonnegative integration constant.

(e) Draw the Penrose diagram of the solutions with $\alpha > 0$, $\Lambda > 0$ and $k = 0$ (currently believed to model our physical Universe).

(5) Recall that the nonvanishing components of the Einstein tensor of the static, spherically symmetric Lorentzian metric

$$ds^2 = -(A(r))^2 dt^2 + (B(r))^2 dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\phi^2)$$  

in the orthonormal frame dual to

$$\omega^0 = Adt, \quad \omega^r = Bdr, \quad \omega^\theta = r d\theta, \quad \omega^\phi = r \sin \theta d\phi,$$
are given by
\[ G_{00} = \frac{2B'}{rB^3} + \frac{B^2 - 1}{r^2B^2} = \frac{2m'}{r^2}; \]
\[ G_{rr} = \frac{2A'}{rAB^2} - \frac{B^2 - 1}{r^2B^2}; \]
\[ G_{\theta\theta} = G_{\phi\phi} = \frac{A'B - A'B'}{AB^3} + \frac{A'}{rAB^2} - \frac{B'}{rB^3}, \]
where
\[ B(r) = \left( 1 - \frac{2m(r)}{r} \right)^{-\frac{1}{2}}. \]
In this exercise we will solve the Einstein equation (without cosmological constant) for a static perfect fluid of constant rest density $\rho$ and rest pressure $p$, and match it to a Schwarzschild exterior.

(a) Show that
\[ B(r) = (1 - kr^2)^{-\frac{1}{2}}, \]
where $k = \frac{8\pi}{3}\rho$. Conclude that the spatial metric is that of a sphere $S^3$ with radius $\frac{1}{\sqrt{k}}$.

(b) Solve the ordinary differential equation $G_{rr} = G_{\theta\theta}$ to obtain
\[ A(r) = C \left( 1 - kr^2 \right)^{\frac{1}{2}} + D, \]
where $C, D \in \mathbb{R}$ are integration constants.

(c) Show that the matching conditions to a Schwarzschild exterior of mass $M > 0$ across a surface $r = R$ are
\[
\begin{align*}
A(R) &= \left( 1 - \frac{2M}{R} \right)^{\frac{1}{2}} \\
A'(R) &= \frac{M}{R^2} \left( 1 - \frac{2M}{R} \right)^{-\frac{1}{2}} \\
B(R) &= \left( 1 - \frac{2M}{R} \right)^{-\frac{1}{2}}
\end{align*}
\]
(d) Conclude that
\[ A(r) = \frac{3}{2} \left( 1 - \frac{2M}{R} \right)^{\frac{1}{2}} - \frac{1}{2} \left( 1 - \frac{2Mr^2}{R^3} \right)^{\frac{1}{2}}. \]
(e) Show that
\[ p(r) = \frac{k \left( 1 - kr^2 \right)^{\frac{1}{2}}}{\frac{3}{2} \left( 1 - kR^2 \right)^{\frac{1}{2}} - \frac{1}{2} \left( 1 - kr^2 \right)^{\frac{1}{2}}} - k. \]
What is the value of $p(R)$?
2. EXACT SOLUTIONS

(f) Show that $M$ and $R$ must satisfy **Buchdahl’s limit:**

$$\frac{2M}{R} < \frac{8}{9}$$

What happens to $p(0)$ as $\frac{2M}{R} \to \frac{8}{9}$?
CHAPTER 3

Causality

1. Past and future

A spacetime \((M, g)\) is said to be **time-orientable** if there exists a timelike vector field, that is, a vector field \(X\) satisfying \(g(X, X) < 0\). In this case, we can define a time orientation on each tangent space \(T_pM\) by declaring causal vectors \(v \in T_pM\) to be **future-pointing** if \(g(v, X_p) < 0\). It can be shown that any non-time-orientable spacetime admits a **time-orientable double covering** (just like any non-orientable manifold admits an orientable double covering).

Assume that \((M, g)\) is **time-oriented** (i.e. time-orientable with a definite choice of time orientation). A timelike or causal curve \(c : I \subset \mathbb{R} \rightarrow M\) is said to be **future-directed** if \(\dot{c}\) is future-pointing. The **chronological future** of \(p \in M\) is the set \(I^+(p)\) of all points to which \(p\) can be connected by a future-directed timelike curve. The **causal future** of \(p \in M\) is the set \(J^+(p)\) of all points to which \(p\) can be connected by a future-directed causal curve. Notice that \(I^+(p)\) is simply the set of all events which are accessible to a particle with nonzero mass at \(p\), whereas \(J^+(p)\) is the set of events which can be causally influenced by \(p\) (as this causal influence cannot propagate faster than the speed of light). Analogously, the **chronological past** of \(p \in M\) is the set \(I^-(p)\) of all points which can be connected to \(p\) by a future-directed timelike curve, and the **causal past** of \(p \in M\) is the set \(J^-(p)\) of all points which can be connected to \(p\) by a future-directed causal curve.

In general, the chronological and causal pasts and futures can be quite complicated sets, because of global features of the spacetime. Locally, however, causal properties are similar to those of Minkowski spacetime. More precisely, we have the following statement:

**Proposition 1.1.** Let \((M, g)\) be a time-oriented spacetime. Then each point \(p_0 \in M\) has an open neighborhood \(V \subset M\) such that the spacetime \((V, g)\) obtained by restricting \(g\) to \(V\) satisfies:

1. \(V\) is **geodesically convex**, that is, \(V\) is a normal neighborhood of each of its points such that given \(p, q \in V\) there exists a unique geodesic (up to reparameterization) connecting \(p\) to \(q\);
2. \(q \in I^+(p)\) if and only if there exists a future-directed timelike geodesic connecting \(p\) to \(q\);
3. \(J^+(p) = \overline{I^+(p)}\);
3. CAUSALITY

(4) \( q \in I^+(p) \setminus I^+(p) \) if and only if there exists a future-directed null geodesic connecting \( p \) to \( q \).

**Proof.** Recall that the exponential map \( \exp_p : U \subset T_p M \to M \) is the map given by

\[
\exp_p(v) = c_v(1),
\]

where \( c_v \) is the geodesic with initial conditions \( c_v(0) = p, \dot{c}_v(0) = v \); equivalently, \( \exp_p(tv) = c_v(t) \). Recall also that \( V \) is a normal neighborhood of \( p \) if \( \exp_p : U \to V \) is a diffeomorphism. The existence of geodesically convex neighborhoods is true for any affine connection and is proved for instance in [KN96].

To prove assertion (2), we start by noticing that if there exists a future-directed timelike geodesic connecting \( p \) to \( q \) then it is obvious that \( q \in I^+(p) \). Suppose now that \( q \in I^+(p) \); then there exists a future-directed timelike curve \( c : [0,1] \to V \) such that \( c(0) = p \) and \( c(1) = q \). Choose normal coordinates \((x^0, x^1, x^2, x^3)\), given by the parameterization

\[
\varphi(x^0, x^1, x^2, x^3) = \exp_p(x^0E_0 + x^1E_1 + x^2E_2 + x^3E_3),
\]

where \( \{E_0, E_1, E_2, E_3\} \) is an orthonormal basis of \( T_p M \) with \( E_0 \) timelike and future-pointing. These are global coordinates in \( V \), since \( \exp_p : U \to V \) is a diffeomorphism. Defining

\[
W_p(q) = -(x^0(q))^2 + (x^1(q))^2 + (x^2(q))^2 + (x^3(q))^2
= \sum_{\mu, \nu=0}^{3} \eta_{\mu \nu} x^\mu(q) x^\nu(q),
\]

with \( \eta_{\mu \nu} = \text{diag}(-1, 1, 1, 1) \), we have to show that \( W_p(q) < 0 \). Let \( W_p(t) = W_p(c(t)) \). Since \( x^\mu(p) = 0 \), we have \( W_p(0) = 0 \). Setting \( x^\mu(t) = x^\mu(c(t)) \), we obtain

\[
\dot{W}_p(t) = 2 \sum_{\mu, \nu=0}^{3} \eta_{\mu \nu} x^\mu(t) \dot{x}^\nu(t);
\]

\[
\ddot{W}_p(t) = 2 \sum_{\mu, \nu=0}^{3} \eta_{\mu \nu} x^\mu(t) \ddot{x}^\nu(t) + 2 \sum_{\mu, \nu=0}^{3} \eta_{\mu \nu} \dot{x}^\mu(t) \dot{x}^\nu(t),
\]

and consequently

\[
W_p(0) = 0;
\]

\[
\dot{W}_p(0) = 2 \langle \dot{c}(0), \dot{c}(0) \rangle < 0.
\]

Therefore there exists \( \varepsilon > 0 \) such that \( W_p(t) < 0 \) for \( t \in (0, \varepsilon) \).

A slightly modified version of the Gauss Lemma shows that the level surfaces of \( W_p \) are orthogonal to the geodesics through \( p \). Therefore, if \( c_v(t) = \exp_p(tv) \) is the geodesic with initial condition \( v \in T_p M \), we have

\[
(\text{grad } W_p)_{c_v(1)} = a(v) \dot{c}_v(1).
\]
Now
\[ \langle (\nabla W_{p})_{c_{v}}, \dot{c}_{v}(t) \rangle = \frac{d}{dt} W_{p}(c_{v}(t)) = \frac{d}{dt} W_{p}(c_{v}(1)) = \frac{d}{dt} (t^{2} W_{p}(c_{v}(1))) = 2t W_{p}(c_{v}(1)), \]
and hence
\[ \langle (\nabla W_{p})_{c_{v}(1)}, \dot{c}_{v}(1) \rangle = 2 W_{p}(c_{v}(1)). \]
On the other hand,
\[ \langle (\nabla W_{p})_{c_{v}(1)}, \dot{c}_{v}(1) \rangle = \langle a(v) \dot{c}_{v}(1), \dot{c}_{v}(1) \rangle = a(v) \langle v, v \rangle = a(v) W_{p}(c_{v}(1)). \]
We conclude that \( a(v) = 2 \), and therefore
\[ (\nabla W_{p})_{c_{v}(1)} = 2 \dot{c}_{v}(1). \]
Consequently, \( \nabla W_{p} \) is tangent to geodesics through \( p \), being future-pointing on future-directed geodesics.

Suppose that \( W_{p}(t) < 0 \). Then \( (\nabla W_{p})_{c_{v}(t)} \) is timelike future-pointing, and so
\[ W(t) = \langle (\nabla W_{p})_{c_{v}(t)}, \dot{c}(t) \rangle < 0, \]
as \( \dot{c}(t) \) is also timelike future-pointing. We conclude that we must have \( W_{p}(t) < 0 \) for all \( t \in [0,1] \). In particular, \( W_{p}(q) = W_{p}(1) < 0 \), and hence there exists a future-directed timelike geodesic connecting \( p \) to \( q \).

To prove assertion (3), let us see first that \( \overline{I^{+}(p)} \subset J^{+}(p) \). If \( q \in \overline{I^{+}(p)} \), then \( q \) is the limit of a sequence of points \( q_{n} \in I^{+}(p) \). By (2), \( q_{n} = \exp_{p}(v_{n}) \) with \( v_{n} \in T_{p}M \) timelike future-pointing. Since \( \exp_{p} \) is a diffeomorphism, \( v_{n} \) converges to a causal future-pointing vector \( v \in T_{p}M \), and so \( q = \exp_{p}(v) \) can be reached from \( p \) by a future-directed causal geodesic. The converse inclusion \( J^{+}(p) \subset \overline{I^{+}(p)} \) holds in general (cf. Proposition 1.2).

Finally, (4) is obvious from (3) and the fact that \( \exp_{p} \) is a diffeomorphism onto \( V \). □

This local behavior can be used to prove the following global result.

**Proposition 1.2.** Let \((M,g)\) be a time oriented spacetime and \( p \in M \). Then:
1. \( I^{+}(p) \) is open;
2. \( J^{+}(p) \subset \overline{I^{+}(p)} \);
3. \( I^{+}(p) = \text{int} \overline{I^{+}(p)} \);
4. if \( r \in J^{+}(p) \) and \( q \in I^{+}(r) \) then \( q \in I^{+}(p) \);
5. if \( r \in I^{+}(p) \) and \( q \in J^{+}(r) \) then \( q \in I^{+}(p) \).

**Proof.** Exercise. □

The twin paradox also holds locally for general spacetimes. More precisely, we have the following statement:
3. CAUSALITY

Proposition 1.3. Let \((M, g)\) be a time-oriented spacetime, \(p_0 \in M\) and \(V \subset M\) a geodesically convex open neighborhood of \(p_0\). The spacetime \((V, g)\) obtained by restricting \(g\) to \(V\) satisfies the following property: if \(p, q \in V\) with \(q \in I^+(p)\), \(c\) is the timelike geodesic connecting \(p\) to \(q\) and \(\gamma\) is any timelike curve connecting \(p\) to \(q\), then \(\tau(\gamma) \leq \tau(c)\), with equality if and only if \(\gamma\) is a reparameterization of \(c\).

Proof. Any timelike curve \(\gamma : [0, 1] \to V\) satisfying \(\gamma(0) = p\), \(\gamma(1) = q\) can be written as 
\[\gamma(t) = \exp_p(r(t)n(t)),\]
for \(t \in [0, 1]\), where \(r(t) \geq 0\) and \((n(t), n(t)) = -1\). We have 
\[\dot{\gamma}(t) = (\exp_p)_* (\dot{r}(t)n(t) + r(t)\dot{n}(t)).\]
Since \((n(t), n(t)) = -1\), we have \(\langle \dot{n}(t), n(t) \rangle = 0\), and consequently \(\dot{n}(t)\) is tangent to the level surfaces of the function \(v \mapsto \langle v, v \rangle\). We conclude that 
\[\dot{\gamma}(t) = \dot{r}(t)X_{\gamma(t)} + Y(t),\]
where \(X\) is the unit tangent vector field to timelike geodesics through \(p\) and \(Y(t) = r(t)(\exp_p)_* \dot{n}(t)\) is tangent to the level surfaces of \(W_p\) (hence orthogonal to \(X_{\gamma(t)}\)). Consequently,
\[
\tau(\gamma) = \int_0^1 \left| \langle \dot{r}(t)X_{\gamma(t)} + Y(t), \dot{r}(t)X_{\gamma(t)} + Y(t) \rangle \right|^\frac{1}{2} dt \\
= \int_0^1 (\dot{r}(t)^2 - |Y(t)|^2)^\frac{1}{2} dt \\
\leq \int_0^1 \dot{r}(t) dt = r(1) = \tau(c),
\]
where we have used the facts that \(\gamma\) is timelike, \(\dot{r}(t) > 0\) for all \(t \in [0, 1]\) (as \(\gamma\) is future-pointing) and \(\tau(c) = r(1)\) (as \(q = \exp_p(r(1)n(1))\)). It should be clear that \(\tau(\gamma) = \tau(c)\) if and only if \(|Y(t)| \equiv 0 \iff Y(t) \equiv 0\) (\(Y(t)\) is spacelike or zero) for all \(t \in [0, 1]\), implying that \(n\) is constant. In this case, \(\gamma(t) = \exp_p(r(t)n)\) is, up to reparameterization, the geodesic through \(p\) with initial condition \(n \in T_pM\). \(\square\)

There is also a local property characterizing null geodesics.

Proposition 1.4. Let \((M, g)\) be a time-oriented spacetime, \(p_0 \in M\) and \(V \subset M\) a geodesically convex open neighborhood of \(p_0\). The spacetime \((V, g)\) obtained by restricting \(g\) to \(V\) satisfies the following property: if \(p, q \in V\) there exists a future-directed null geodesic \(c\) connecting \(p\) to \(q\) and \(\gamma\) is a causal curve connecting \(p\) to \(q\) then \(\gamma\) is a reparameterization of \(c\).

Proof. Since \(p\) and \(q\) are connected by a null geodesic, we conclude from Proposition 1.1 that \(q \in J^+(p) \setminus I^+(p)\). Let \(\gamma : [0, 1] \to V\) be a causal curve connecting \(p\) to \(q\). Then we must have \(\gamma(t) \in J^+(p) \setminus I^+(p)\).
2. CAUSALITY CONDITIONS

for all \( t \in [0, 1] \), since \( \gamma(t_0) \in I^+(p) \) implies \( \gamma(t) \in I^+(p) \) for all \( t > t_0 \) (see Proposition 1.2). Consequently, we have

\[
W_p(\gamma(t)) = 0 \Rightarrow \langle (\text{grad} \ W_p)_{\gamma(t)}, \dot{\gamma}(t) \rangle = 0,
\]

where \( W_p \) was defined in the proof of Proposition 1.1. The formula

\[
(\text{grad} \ W_p)_{c_\nu}(1) = 2 \dot{c}_\nu(1),
\]

which was proved for timelike geodesics \( c_\nu \) with initial condition \( \nu \in T_pM \), must also hold for null geodesics (by continuity). Hence \( \text{grad} \ W_p \) is tangent to the null geodesics ruling \( J^+(p) \setminus I^+(p) \) and future-pointing. Since \( \dot{\gamma}(t) \) is also future-pointing, we conclude that \( \dot{\gamma} \) is proportional to \( \text{grad} W_p \), and therefore \( \gamma \) is a reparameterization of a null geodesic, which must be \( c \).

Corollary 1.5. Let \((M, g)\) be a time-oriented spacetime and \( p \in M \). If \( q \in J^+(p) \setminus I^+(p) \) then any future-directed causal curve connecting \( p \) to \( q \) must be a reparameterized null geodesic.

2. Causality conditions

For physical applications, it is important to require that the spacetime satisfies reasonable causality conditions. The simplest of these conditions excludes time travel, i.e. the possibility of a particle returning to an event in its past history.

Definition 2.1. A spacetime \((M, g)\) is said to satisfy the chronology condition if it does not contain closed timelike curves.

This condition is violated by compact spacetimes:

Proposition 2.2. Any compact spacetime \((M, g)\) contains closed timelike curves.

Proof. Taking if necessary the time-orientable double covering, we can assume that \((M, g)\) is time-oriented. Since \( I^+(p) \) is an open set for any \( p \in M \), it is clear that \( \{I^+(p)\}_{p \in M} \) is an open cover of \( M \). If \( M \) is compact, we can obtain a finite subcover \( \{I^+(p_1), \ldots, I^+(p_N)\} \). Now if \( p_i \in I^+(p_i) \) for \( i \neq 1 \) then \( I^+(p_1) \subset I^+(p_i) \), and we can exclude \( I^+(p_1) \) from the subcover. Therefore, we can assume without loss of generality that \( p_1 \in I^+(p_1) \), and hence there exists a closed timelike curve starting and ending at \( p_1 \).

A stronger restriction on the causal behavior of the spacetime is the following:

Definition 2.3. A spacetime \((M, g)\) is said to be stably causal if there exists a global time function, i.e. a smooth function \( t : M \rightarrow \mathbb{R} \) such that \( \text{grad}(t) \) is timelike.

In particular, a stably causal spacetime is time-orientable. We choose the time orientation defined by \( -\text{grad}(t) \), so that \( t \) increases along future-directed timelike curves. Notice that this implies that no closed timelike
curves can exist, i.e. any stably causal spacetime satisfies the chronology condition. In fact, any small perturbation of a stably causal spacetime still satisfies the chronology condition (Exercise 4).

Let \((M, g)\) be a time-oriented spacetime. A smooth future-directed causal curve \(c : (a, b) \to M\) (with possibly \(a = -\infty\) or \(b = +\infty\)) is said to be future-inextendible if \(\lim_{t \to b} c(t)\) does not exist. The definition of a past-inextendible causal curve is analogous. The future domain of dependence of \(S \subset M\) is the set \(D^+(S)\) of all events \(p \in M\) such that any past-inextendible causal curve starting at \(p\) intersects \(S\). Therefore any causal influence on an event \(p \in D^+(S)\) had to register somewhere in \(S\), and one can expect that what happens at \(p\) can be predicted from data on \(S\). Similarly, the past domain of dependence of \(S\) is the set \(D^-(S)\) of all events \(p \in M\) such that any future-inextendible causal curve starting at \(p\) intersects \(S\). Therefore any causal influence of an event \(p \in D^-(S)\) will register somewhere in \(S\), and one can expect that what happened at \(p\) can be retrodicted from data on \(S\). The domain of dependence of \(S\) is simply the set \(D(S) = D^+(S) \cup D^-(S)\).

Let \((M, g)\) be a stably causal spacetime with time function \(t : M \to \mathbb{R}\). The level sets \(S_a = t^{-1}(a)\) are said to be Cauchy hypersurfaces if \(D(S_a) = M\). Spacetimes for which this happens have particularly good causal properties.

**Definition 2.4.** A stably causal spacetime possessing a time function whose level sets are Cauchy hypersurfaces is said to be globally hyperbolic.

Notice that the future and past domains of dependence of the Cauchy hypersurfaces \(S_a\) are \(D^+(S_a) = t^{-1}([a, +\infty))\) and \(D^-(S_a) = t^{-1}((-\infty, a])\).

3. Exercises

1. Let \((M, g)\) be the quotient of the 2-dimensional Minkowski spacetime by the discrete group of isometries generated by the map \(f(t, x) = (-t, x + 1)\). Show that \((M, g)\) is not time orientable.

2. Let \((M, g)\) be a time oriented spacetime and \(p \in M\). Show that:
   (a) \(I^+(p)\) is open;
   (b) \(J^+(p)\) is not necessarily closed;
   (c) \(I^+(p) \subset \overline{J^+(p)}\);  
   (d) \(I^+(p) = \text{int} J^+(p)\);
   (e) if \(r \in J^+(p)\) and \(q \in I^+(r)\) then \(q \in I^+(p)\);
   (f) if \(r \in I^+(p)\) and \(q \in J^+(r)\) then \(q \in I^+(p)\);
   (g) it may happen that \(I^+(p) = M\).

3. Consider the 3-dimensional Minkowski spacetime \((\mathbb{R}^3, g)\), where

\[ g = -dt \otimes dt + dx \otimes dx + dy \otimes dy. \]
Let $c : \mathbb{R} \to \mathbb{R}^3$ be the curve $c(t) = (t, \cos t, \sin t)$. Show that although $\dot{c}(t)$ is null for all $t \in \mathbb{R}$ we have $c(t) \in I^+(c(0))$ for all $t > 0$. What kind of motion does this curve represent?

(4) Let $(M, g)$ be a stably causal spacetime and $h$ an arbitrary symmetric $(2, 0)$-tensor field with compact support. Show that for sufficiently small $|\varepsilon|$ the tensor field $g_\varepsilon = g + \varepsilon h$ is still a Lorentzian metric on $M$, and $(M, g_\varepsilon)$ satisfies the chronology condition.

(5) Let $(M, g)$ be the quotient of the 2-dimensional Minkowski spacetime by the discrete group of isometries generated by the map $f(t, x) = (t + 1, x + 1)$. Show that $(M, g)$ satisfies the chronology condition, but there exist arbitrarily small perturbations of $(M, g)$ (in the sense of Exercise 4) which do not.

(6) Let $(M, g)$ be a time oriented spacetime and $S \subset M$. Show that:
   (a) $S \subset D^+(S)$;
   (b) $D^+(S)$ is not necessarily open;
   (c) $D^+(S)$ is not necessarily closed.

(7) Let $(M, g)$ be the 2-dimensional spacetime obtained by removing the positive $x$-semi-axis of Minkowski 2-dimensional spacetime (cf. Figure 1). Show that:
   (a) $(M, g)$ is stably causal but not globally hyperbolic;
   (b) there exist points $p, q \in M$ such that $J^+(p) \cap J^-(q)$ is not compact;
   (c) there exist points $p, q \in M$ with $q \in I^+(p)$ such that the supremum of the lengths of timelike curves connecting $p$ to $q$ is not attained by any timelike curve.

![Figure 1. Stably causal but not globally hyperbolic spacetime.](image)

(8) Let $(\Sigma, h)$ be a 3-dimensional Riemannian manifold. Show that the spacetime $(M, g) = (\mathbb{R} \times \Sigma, -dt \otimes dt + h)$ is globally hyperbolic if and only if $(\Sigma, h)$ is complete.

(9) Show that the following spacetimes are globally hyperbolic:
(a) the Minkowski spacetime;
(b) the FLRW spacetimes;
(c) the region \( \{r > 2m\} \) of the Schwarzschild spacetime;
(d) the region \( \{r < 2m\} \) of the Schwarzschild spacetime;
(e) the maximal analytical extension of the Schwarzschild spacetime.

(10) Let \((M, g)\) be a global hyperbolic spacetime with Cauchy hypersurface \(S\). Show that \(M\) is diffeomorphic to \(\mathbb{R} \times S\).
Appendix: Mathematical concepts for physicists

In this appendix we list some mathematical concepts which will be used in the main text, for the benefit of readers whose background is in Physics.

**Topology**

**Definition 3.1.** A **topological space** is a set $M$ with a topology, that is, a list of the **open subsets** of $M$, satisfying:

1. Both $\emptyset$ and $M$ are open;
2. Any union of open sets is open;
3. Any finite intersection of open sets is open.

All the usual topological notions can now be defined. For instance, a **closed set** is a set whose complement is open. The **interior** $\text{int} A$ of a subset $A \subset M$ is the largest open set contained in $A$, its **closure** $\overline{A}$ is the smallest closed set containing $A$, and its **boundary** is $\partial A = \overline{A} \setminus \text{int} A$.

The main object of topology is the study of limits and continuity.

**Definition 3.2.** A sequence $\{p_n\}$ is said to **converge** to $p \in M$ if for any open set $U \ni p$ there exists $k \in \mathbb{N}$ such that $p_n \in U$ for all $n \geq k$.

**Definition 3.3.** A map $f : M \to N$ between two topological spaces is said to be **continuous** if for each open set $U \subset N$ the preimage $f^{-1}(U)$ is an open subset of $M$. A bijection $f$ is called a **homeomorphism** if both $f$ and its inverse $f^{-1}$ are continuous.

Two fundamental concepts in topology are compactness and connectedness.

**Definition 3.4.** A subset $A \subset M$ is said to be **compact** if every cover of $A$ by open sets admits a finite subcover. It is said to be **connected** it is impossible to write $A = (A \cap U) \cup (A \cap V)$ with $U, V$ open sets and $A \cap U, A \cap V \neq \emptyset$.

The following result generalizes the theorems of Weierstrass and Bolzano.

**Theorem 3.5.** Continuous maps carry compact sets to compact sets, and connected sets to connected sets.
Metric spaces

Definition 3.6. A metric space is a set $M$ and a distance function $d : M \times M \to [0, +\infty)$ satisfying:

1. **Positivity:** $d(p, q) \geq 0$ and $d(p, q) = 0$ if and only if $p = q$;
2. **Symmetry:** $d(p, q) = d(q, p)$;
3. **Triangle inequality:** $d(p, r) \leq d(p, q) + d(q, r)$,

for all $p, q, r \in M$.

The open ball with center $p$ and radius $\varepsilon$ is the set

$$B_\varepsilon(p) = \{ q \in M \mid d(p, q) < \varepsilon \}.$$

Any metric space has a natural topology, whose open sets are unions of open balls. In this topology $p_n \to p$ if and only if $d(p_n, p) \to 0$, $F \subset M$ is closed if and only if every convergent sequence in $F$ has limit in $F$, and $K \subset M$ is compact if and only if every sequence in $K$ has a sublimit in $K$.

A fundamental notion for metric spaces is completeness.

Definition 3.7. A sequence $\{p_n\}$ in $M$ is said to be a Cauchy sequence if for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $d(p_n, p_m) < \varepsilon$ for all $m, n > N$. A metric space is said to be complete if all its Cauchy sequences converge.

In particular any compact metric space is complete.

Hopf-Rinow theorem

Definition 3.8. A Riemannian manifold is said to be geodesically complete if any geodesic is defined for every value of its parameter.

Definition 3.9. Let $(M, g)$ be a connected Riemannian manifold and $p, q \in M$. The distance between $p$ and $q$ is defined as

$$d(p, q) = \inf \{ l(\gamma) \mid \gamma \text{ is a smooth curve connecting } p \text{ to } q \}.$$

It is easily seen that $(M, d)$ is a metric space. Remarkably, the completeness of this metric space is equivalent to geodesic completeness.

Theorem 3.10. (Hopf-Rinow) A connected Riemannian manifold $(M, g)$ is geodesically complete if and only if $(M, d)$ is a complete metric space.

Differential forms

Definition 3.11. A differential-form $\omega$ of degree $k$ is simply a completely anti-symmetric $k$-tensor: $\omega_{\alpha_1 \ldots \alpha_k} = \omega^{[\alpha_1 \ldots \alpha_k]}$.

For instance, covector fields are differential forms of degree 1. Differential forms are useful because of their rich algebraic and differential structure.
**Definition 3.12.** If $\omega$ is a $k$-form and $\eta$ is an $l$-form then their **exterior product** is the $(k + l)$-form

$$(\omega \wedge \eta)_{\alpha_1 \cdots \alpha_k \beta_1 \cdots \beta_l} = \frac{(k + l)!}{k! l!} \omega_{[\alpha_1 \cdots \alpha_k} \eta_{\beta_1 \cdots \beta_l]},$$

and the **exterior derivative** of $\omega$ is the $(k + 1)$-form

$$(d\omega)_{\alpha_1 \cdots \alpha_k} = (k + 1)\nabla_{[\alpha} \omega_{\alpha_1 \cdots \alpha_k]},$$

where $\nabla$ is any symmetric connection.

It is easy to see that any $k$-form $\omega$ is given in local coordinates by

$$\omega = \sum_{\alpha_1 < \cdots < \alpha_k} \omega_{\alpha_1 \cdots \alpha_k} (x) dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_k},$$

and therefore has $\binom{n}{k}$ independent components on an $n$-dimensional manifold.

**Proposition 3.13.** If $\omega$, $\eta$ and $\theta$ are differential forms then:

1. $\omega \wedge (\eta \wedge \theta) = (\omega \wedge \eta) \wedge \theta$;
2. $\omega \wedge \eta = (-1)^{\deg \omega (\deg \eta)} \eta \wedge \omega$;
3. $\omega \wedge (\eta + \theta) = \omega \wedge \eta + \omega \wedge \theta$;
4. $d(\omega + \eta) = d\omega + d\eta$;
5. $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^{\deg \omega} \omega \wedge d\eta$;
6. $d^2 \omega = 0$.

It is clear from these properties that if the $k$-form $\omega$ is given in local coordinates by (13) above then

$$d\omega = \sum_{\alpha_1 < \cdots < \alpha_k} \sum_{\alpha} \partial_{\alpha} \omega_{\alpha_1 \cdots \alpha_k} dx^\alpha \wedge dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_k}.$$

The last property in Proposition 3.13 has a converse, known as the **Poincaré Lemma**.

**Lemma 3.14.** (Poincaré) If $d\omega = 0$ then locally $\omega = d\eta$.

A related result is the **Frobenius Theorem**. Here we present a particular case of this result.

**Theorem 3.15.** (Frobenius) The nonvanishing 1-form $\omega$ is locally orthogonal to a family of hypersurfaces if and only if $\omega \wedge d\omega = 0$.

To prove the easy direction in this equivalence it suffices to note that $\omega$ is locally orthogonal to a family of hypersurfaces $\{ f = \text{constant} \}$ if and only if $\omega = gdf$ for some nonvanishing function $g$. Note that in particular this is always true for 1-forms in 2-dimensional manifolds.

We will now assume that our $n$-dimensional manifold is **oriented**, that is, that an orientation can be, and has been, consistently chosen on every tangent space. Any $n$-form $\omega$ is written in local coordinates as

$$\omega = a(x) dx^1 \wedge \cdots \wedge dx^n.$$
If the coordinate system is positive, that is, if the coordinate basis \{\partial_1, \ldots, \partial_n\} has positive orientation at all points, we define
\[
\int_U \omega = \int_{x(U)} a(x) dx^1 \ldots dx^n,
\]
where \(U\) is the coordinate neighborhood. This formula does not depend on the choice of local coordinates because \(dx^1 \wedge \cdots \wedge dx^n\) transforms by the determinant of the change of variables.

**Theorem 3.16. (Stokes)** If \(M\) is an oriented \(n\)-dimensional manifold with boundary \(\partial M\) and \(\omega\) is an \((n-1)\)-form then
\[
\int_M d\omega = \int_{\partial M} \omega.
\]

In this theorem the orientations of \(M\) and \(\partial M\) are related as follows: if \(\partial M\) is a level set of \(x^1\) in a positive coordinate system and \(\partial_1\) points outwards then the coordinate system \((x^2, \ldots, x^n)\) on \(\partial M\) is positive.

If \(M\) has a metric \(g\) then its **volume element** is the \(n\)-form \(\epsilon\) which is 1 when contracted with a positive orthonormal frame. In positive local coordinates we have
\[
\epsilon = \sqrt{|\det(g_{\mu\nu})|} dx^1 \wedge \cdots \wedge dx^n.
\]

It is easily seen that \(\nabla \epsilon = 0\), where \(\nabla\) is the Levi-Civita connection. If \(\omega\) is a \(k\)-form then its **Hodge dual** is the \((n-k)\)-form \(*\omega\) given by
\[
(*\omega)_{\beta_1 \cdots \beta_{n-k}} = \frac{1}{k!} \omega^{\alpha_1 \cdots \alpha_k} \epsilon_{\alpha_1 \cdots \alpha_k \beta_1 \cdots \beta_{n-k}}.
\]

The operator \(*\), called the Hodge star, can alternatively be defined as follows: if \(\{\omega^1, \ldots, \omega^n\}\) is any positively oriented orthonormal coframe then
\[
* (\omega^1 \wedge \cdots \wedge \omega^k) = g(\omega^1, \omega^1) \cdots g(\omega^k, \omega^k) \omega^{k+1} \wedge \cdots \wedge \omega^n.
\]

**Lie derivative**

A vector field \(X\) can be identified with the differential operator that corresponds to taking derivatives along \(X\). In local coordinates, this operator is given by
\[
X \cdot f = X^\mu \partial_\mu f.
\]
It turns out that the commutator of two vector fields \(X\) and \(Y\), regarded as differential operators, is also a vector field:
\[
[X, Y] \cdot f = X \cdot (Y^\mu \partial_\mu f) - Y \cdot (X^\mu \partial_\mu f)
= (X \cdot Y^\mu) \partial_\mu f + Y^\mu X^\nu \partial_\nu \partial_\mu f - (Y \cdot X^\mu) \partial_\mu f - X^\mu Y^\nu \partial_\nu \partial_\mu f
= (X \cdot Y^\mu - Y \cdot X^\mu) \partial_\mu f.
\]

**Definition 3.17.** The **Lie bracket** of two vector fields \(X\) and \(Y\) is the vector field
\[
[X, Y] = (X \cdot Y^\mu - Y \cdot X^\mu) \partial_\mu.
\]
This operation is intimately related with the exterior derivative.

**Proposition 3.18.** If \( \omega \) is a 1-form then
\[
d\omega(X, Y) = X \cdot \omega(Y) - Y \cdot \omega(X) - \omega([X, Y])
\]
for all vector fields \( X \) and \( Y \).

**Proof.** In local coordinates we have
\[
d\omega(X, Y) = (\partial_\mu \omega_\nu - \partial_\nu \omega_\mu) X^\mu Y^\nu = Y^\nu X \cdot \omega_\nu - X^\nu Y \cdot \omega_\nu
\]
\[
= X \cdot (\omega_\nu Y^\nu) - \omega_\nu X \cdot Y^\nu - Y \cdot (\omega_\mu X^\mu) + \omega_\mu Y \cdot X^\mu
\]
\[
= X \cdot \omega(Y) - Y \cdot \omega(X) - \omega_\nu (X \cdot Y^\nu - Y \cdot X^\nu).
\]
\[\square\]

If the vector field \( X \) is nonzero at some point \( p \) then there exists a coordinate system defined in a neighborhood of \( p \) such that \( X = \partial_1 \). In fact, we just have to fix local coordinates \((x^2, \ldots, x^n)\) on a hypersurface \( \Sigma \) transverse to \( X \) at \( p \) and let \( x^1 \) be the parameter for the flow of \( X \) starting at \( \Sigma \). If \( T \) is any tensor, we define its **Lie derivative** along \( X \) as the tensor with components
\[
(\mathcal{L}_X T)^{\alpha_1 \cdots \alpha_k}_{\beta_1 \cdots \beta_l} = \partial_1 T^{\alpha_1 \cdots \alpha_k}_{\beta_1 \cdots \beta_l}.
\]
This can be extended to points where \( X \) vanishes by continuity. Although this definition seems to depend on the coordinate system, it is actually invariant. To check this, we just have to find an invariant expression for the Lie derivative of functions, vector fields and 1-forms and then notice that the Leibnitz rule applies.

**Proposition 3.19.** If \( X \) is a vector field then:

1. \( \mathcal{L}_X f = X \cdot f \) for functions \( f \);
2. \( \mathcal{L}_X Y = [X, Y] \) for vector fields \( Y \);
3. \( \mathcal{L}_X \omega = X \lrcorner d\omega + d(X \lrcorner \omega) \) for 1-forms \( \omega \)

(where \( \lrcorner \) means contraction in the first index).

**Proof.** The formula for functions is immediate. In the coordinate system where \( X = \partial_1 \),
\[
\mathcal{L}_X Y = \partial_1 Y^\mu \partial_\mu = (X \cdot Y^\mu - Y \cdot X^\mu) \partial_\mu = [X, Y].
\]
Finally, we have
\[
(X \lrcorner d\omega + d(X \lrcorner \omega))(Y) = d\omega(X, Y) + Y \cdot \omega(X) = X \cdot \omega(Y) - \omega([X, Y])
\]
\[
= \mathcal{L}_X (\omega(Y)) - \omega(\mathcal{L}_X Y) = (\mathcal{L}_X \omega)(Y),
\]
where we used the Leibnitz rule. This formula is sometimes called **Cartan's magic formula.**  \[\square\]
Cartan structure equations

Let \( \{E_\mu\} \) be an orthonormal frame, and \( \{\omega^\mu\} \) the corresponding orthonormal coframe, so that
\[
\omega^\mu(E_\nu) = \delta^\mu_\nu.
\]
Note that the metric can be written as
\[
d\!s^2 = \eta_{\mu\nu} \omega^\mu \otimes \omega^\nu,
\]
where \( (\eta_{\mu\nu}) = \text{diag}(-1,1,1,1) \) is the flat space metric (which we will use to raise and lower indices).

Definition 3.20. The connection forms associated to the orthonormal frame \( \{E_\mu\} \) are the 1-forms \( \omega^\mu_\nu \) such that
\[
\nabla_X E_\nu = \omega^\mu_\nu(X)
\]
for all vector fields \( X \). The curvature forms associated this frame are the 2-forms \( \Omega^\mu_\nu \) such that
\[
R(X,Y)E_\nu = \Omega^\mu_\nu(X,Y)E_\mu
\]
for all vector fields \( X, Y \).

Note that the components of the Riemann tensor in the orthonormal frame can be retrieved from the curvature forms by noticing that
\[
\Omega^\mu_\nu = R_{\alpha\beta}^\mu_\nu \omega^\alpha \otimes \omega^\beta = \sum_{\alpha<\beta} R_{\alpha\beta}^\mu_\nu \omega^\alpha \wedge \omega^\beta.
\]
These forms can be computed by using the so-called Cartan structure equations. This is by far the most efficient way to compute the curvature.

Theorem 3.21. The connection forms are the unique solution of Cartan’s first structure equations
\[
\begin{cases}
\omega_{\mu\nu} = -\omega_{\nu\mu} \\
d\omega^\mu + \omega^\mu_\nu \wedge \omega^\nu = 0
\end{cases},
\]
and the curvature forms are given by Cartan’s second structure equations
\[
\Omega^\mu_\nu = d\omega^\mu_\nu + \omega^\mu_\alpha \wedge \omega^\alpha_\nu.
\]

Proof. The first condition is equivalent to
\[
X \cdot \langle E_\mu, E_\nu \rangle = 0
\]
for all vector fields \( X \), which in turn is equivalent to the compatibility of the metric. Using
\[
d\omega^\mu(X,Y) = X \cdot \omega^\mu(Y) - Y \cdot \omega^\mu(X) - \omega^\mu([X,Y])
\]
for all vector fields \( X \) and \( Y \), it is easy to see that the second condition is equivalent to
\[
[E_\alpha, E_\beta] = \nabla_{E_\alpha} E_\beta - \nabla_{E_\beta} E_\alpha,
\]
which in turn is equivalent to the symmetry of the connection. Since the Levi-Civita connection is the only connection which is symmetric and compatible with the metric, we conclude that Cartan’s first structure equations have a unique solution.

Finally, the third condition can be derived by writing

\[ \Omega^\mu_{\nu}(X,Y)E_\mu = R(X,Y)E_\nu = \nabla_X \nabla_Y E_\nu - \nabla_Y \nabla_X E_\nu - \nabla_{[X,Y]}E_\nu \]

in terms of the connection forms. \qed
Bibliography