Part II Relativity

Lectured by Anastasia Fialkov

Additional notes from Hobson's *General Relativity* and Carroll's *Spacetime and Geometry*

If you spot any mistakes please let me know at jhyl3@cam.ac.uk

Michaelmas 2023

Contents

1 Introduction

1.1 A note on notation

Notation becomes very relevant and important from Chapter 4 onwards. 4-vectors are denoted as bold, e.g. $\mathbf{v}$, and 3-vectors with arrows, e.g. $\vec{v}$, so that a 4-vector may be written as $\mathbf{v} = (c, \vec{v})$ for instance. Similarly ∇ is a covariant derivative (which acts on a 4-vector) but $\vec{\nabla}$ is a vector differential operator acting on 3-vectors. Partial derivatives are written as $\partial/\partial x^a \equiv \partial_a$ and $\mathrm{d}x^a/\mathrm{d}u \equiv \dot{x}^a$ unless otherwise specified. In later chapters this is usually w.r.t. the affine parameter of proper time τ.

1.2 The universality of free-fall: equivalence principles

This section concerns the *strong equivalence principle*, which is motivated by the weak equivalence principle described below.

Newtonian gravity is inconsistent with special relativity. It is described by a gravitational potential $\Phi(\mathbf{x}, t)$ determined via Poisson's equation; for a mass density ρ

$$\nabla^2 \Phi = 4\pi G \rho \tag{1.1}$$

which implies that Φ and thus $\nabla\Phi$ respond instantaneously to changes in ρ. But in SR no causal influence can travel faster than c.

Equality of passive and active gravitational mass

A test particle of *passive gravitational mass* m_G has a force exerted on it by Φ of

$$\mathbf{f} = -m_G \nabla \Phi \tag{1.2}$$

We consider an *active gravitational mass* m_A which contributes to ρ in Poisson's equation (Eq. 1.1). The relevant mass density for a point particle at position $\mathbf{y}(t)$ at time t is

$$\rho(\mathbf{x}, t) = m_A \delta^{(3)}(\mathbf{x} - \mathbf{y}(t))$$

The solution (3-dim, Dirichlet BCs) is

$$\Phi(\mathbf{x}, t) = -\frac{G m_A}{|\mathbf{x} - \mathbf{y}(t)|} \tag{1.3}$$

Then for two particles indicated by subscripts 1 and 2 respectively

$$\mathbf{f}_{2\text{ on }1} = -G m_{G,1} m_{A,2} \frac{\mathbf{x}_1 - \mathbf{x}_2}{|\mathbf{x}_1 - \mathbf{x}_2|^3} \quad \text{and} \quad \mathbf{f}_{1\text{ on }2} = -G m_{G,2} m_{A,1} \frac{\mathbf{x}_2 - \mathbf{x}_1}{|\mathbf{x}_1 - \mathbf{x}_2|^3}$$

Conservation of momentum requires $\mathbf{f}_{2\text{ on }1} = -\mathbf{f}_{1\text{ on }2}$ so it follows that $m_{G,1} m_{A,2} = m_{G,2} m_{A,1}$, or

$$\frac{m_{G,1}}{m_{A,1}} = \frac{m_{G,2}}{m_{A,2}} = \text{const.}$$

which must hold for arbitrary masses, so we conclude that the ratio of passive to active gravitational mass is the same for all particles. We can further take $m_G = m_A$ by absorbing the universal ratio into the gravitational constant G.

Equality of gravitational and inertial mass - the weak equivalence principle

By Newton's second law, the inertial mass m_I obeys

$$\mathbf{f} = m_I \frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2}$$

In combination with Eq. 1.2

$$\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} = -\frac{m_G}{m_I}\nabla\Phi$$

Experimental results have determined this ratio to be $m_G/m_I = 1 + \mathcal{O}(10^{-13})$ so we can always take $m_G = m_I$. The *weak equivalence principle* states that

> Freely-falling particles with negligible gravitational self-interaction follow the same path through space and time if they have the same initial position and velocity, independent of their composition.

The strong equivalence principle

> In an arbitrary gravitational field, *all* the laws of physics in a free-falling, non-rotating laboratory occupying a sufficiently small region of spacetime look locally like special relativity (with no gravity).

The requirement of *a sufficiently small region of spacetime* is necessary because tidal effects arising from a non-uniform gravitational field (e.g. that of Earth) are undetectable in the free-falling frame. Physics relative to the free-falling frame looks just like SR physics in an inertial frame of reference in the absence of gravity. Thus a *local* equivalence is established. Inertial reference frames should be defined locally by free-falling observers, rather than comoving, unaccelerated observers with synchronised clocks.

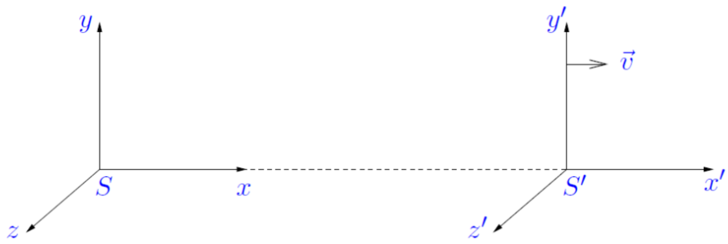
The trajectories of free-falling particles should be determined by the local structure of spacetime; local inertial reference frames correspond to local systems of coordinates over spacetime so that the geometry over a small region looks like that of the spacetime of special relativity.

2 Special relativity

2.1 Fundamentals

Now we take a step back and revisit the fundamentals of SR without gravity.

- *Events* in spacetime are defined uniquely by three spatial coordinates and one time coordinate, so that it has coordinates (ct, x, y, z), in a frame S.
- *Inertial frames* are defined by the requirement that a free particle is at rest or moves in a straight line with fixed speed, i.e. $\ddot{\mathbf{x}} = \mathbf{0}$. Two inertial frames S and S' differ only by 1. a translation; 2. a rotation; 3. relative motion at constant velocity.
- *Standard configuration* implies aligned axes, the same spacetime origin (coincident origins) and relative velocity v along the x, x' axes:



- *Minkowski* spacetime: a 4-dim continuum with invariant geometry characterised by Δs^2.

Postulates of SR:

1. The laws of physics are invariant in all inertial frames of reference (the principle of relativity)
2. The speed of light c in vacuum is the same for all observers regardless their motion

Recap: Newtonian theory; Galilean transformation

The notion of *absolute time* induces $t' = t$;

$$t' = t, \quad x' = x - vt, \quad y' = y, \quad z' = z$$

For two events A and B the Galilean transformation implies that $\Delta t \equiv t_B - t_A$ and $\Delta r^2 \equiv \Delta x^2 + \Delta y^2 + \Delta z^2$ (for simultaneous events) are invariant.

SR; Lorentz transformation

The Lorentz transformation under standard configuration is

$$ct' = \gamma(ct - \beta x), \quad x' = \gamma(x - \beta ct), \quad y' = y, \quad z' = z \tag{2.1}$$

where $\beta \equiv v/c$ and $\gamma \equiv (1 - \beta^2)^{-1/2}$

This is a *Lorentz boost* in the x-direction; note that this reduces to the Galilean transform when $\beta \ll 1$. In both cases (Galilean and Lorentz) the inverse transformation is obtained by setting $v \to -v$.

For two events A and B the squared interval Δs^2 is invariant under any Lorentz transformation:

$$\Delta s^2 = c^2 \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 \tag{2.2}$$

It follows from Eq. 2.1 that *simultaneity* is not Lorentz invariant: events that are simultaneous in one inertial frame (e.g. $\Delta t = 0$ in S) will generally not be in another.

2.2 Lorentz transformations as 4-dim 'rotations'

In Minkowski spacetime, (Cartesian) inertial frames S and S' 'relabel' events by performing the coordinate transformation $(ct, x, y, z) \to (ct', x', y', z')$. We define a *rapidity* parameter $\psi \in (-\infty, \infty)$ by $\beta = \tanh \psi$ s.t.

$$\beta = \tanh \psi \implies \gamma = \cosh \psi \quad \text{and} \quad \gamma\beta = \sinh \psi \tag{2.3}$$

For S and S' in standard configuration,

$$\begin{aligned} ct' &= ct \cosh \psi - x \sinh \psi \\ x' &= -ct \sinh \psi + x \cosh \psi \\ y' &= y \\ z' &= z \end{aligned} \tag{2.4}$$

Then a Lorentz boost resembles a *hyperbolic* rotation of Minkowski spacetime.

General case: S' moves with constant velocity v w.r.t. S in an *arbitrary* direction; the axes of S' are rotated w.r.t. S; at $t = t' = 0$ the origins of S and S' are not coincident (separated by a vector displacement, say $\mathbf{a}$, as measured in S).

- To deal with non-coincident origins, apply appropriate temporal and spatial displacements; these are *inhomogeneous* Lorentz transfomations or Poincaré transformations;
- The 'remaining' Lorentz boost then consists of
  1. Apply a purely spatial rotation in S to align the new x-axis with the relative velocity of the two frames;
  2. Apply a standard Lorentz transformations (Eq. 2.1);
  3. Apply a spatial rotation in the transformed coordinates to align the axes with those of S'.

2.3 Light cones and spacetime diagrams

The interval (Eq. 2.2) is invariant under a Lorentz boost (Eq. 2.1):

$$\begin{aligned} \Delta s^2 &= c^2 \Delta t'^2 - \Delta x'^2 - \Delta y'^2 - \Delta z'^2 \\ &= [(c\Delta t) \cosh \psi - (\Delta x) \sinh \psi]^2 - [-(c\Delta t) \sinh \psi + (\Delta x) \cosh \psi]^2 - \Delta y^2 - \Delta z^2 \\ &= c^2 \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 \end{aligned}$$

This is true also for any Poincaré transformation. It follows that Δs^2 is an underlying geometrical property of spacetime itself, and that the sign of Δs^2 is defined invariantly, with the *lightcone* separating timelike and spacelike events:

$\Delta s^2 > 0$ - the interval is timelike; $\exists$ a S in which $x_A = x_B$ but $t_A \neq t_B$; this is *in* the lightcone;
$\Delta s^2 = 0$ - the interval is null or lightlike; this is *on* the lightcone;
$\Delta s^2 < 0$ - the interval is spacelike; $\exists$ a S in which $t_A = t_B$ but $x_A \neq x_B$; this is *outside of* the lightcone.

Events outside of the lightcone of A cannot influence/be influenced by A.

Simultaneity is not Lorentz invariant, but the *temporal ordering of events* is, provided that they are *timelike or null-separated* (i.e. causally connected, $\Delta s^2 \geq 0$). This falls out of the Lorentz transform

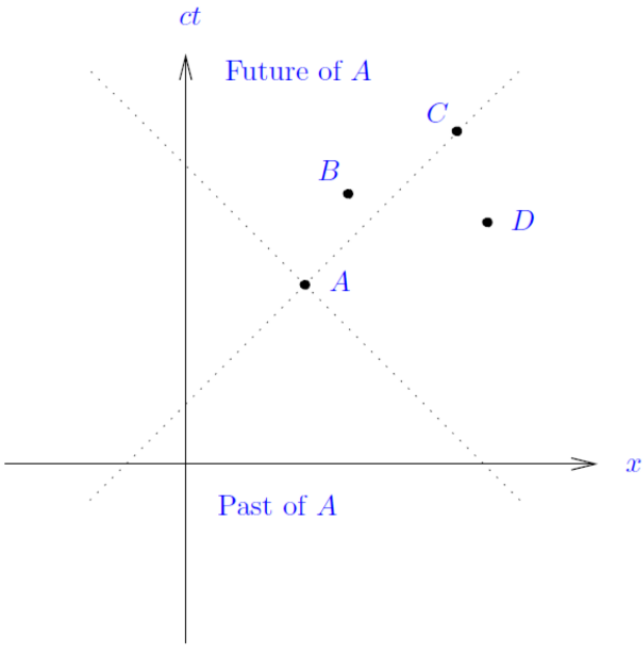


Figure 1: B and A are separated by a timelike interval and B lies in the *forward* lightcone of A. These events could be causally connected. C and A are separated by a null interval and could be connected by a *light signal*. D and A are separated by a spacelike interval and cannot be causally connected.

2.4 Length contraction and time dilation

We consider two inertial frames S and S' in standard configuration. A rod of *proper length* $\ell_0 = x'_B - x'_A$ is at rest in S'. In S the rod moves at speed v along the x-axis with length $\ell = x_B - x_A$, with x_A and x_B being measured at the same time t. With Eq. 2.1 we have

$$x'_A = \gamma(x_A(t) - vt) \quad \text{and} \quad x'_B = \gamma(x_B(t) - vt)$$

and so

$$\ell = \frac{\ell_0}{\gamma} = \ell_0(1 - v^2/c^2)^{1/2} \tag{2.5}$$

The rod suffers no contraction in directions perpendicular to its velocity.

- It follows that the volume V of a moving object is $V = V_0(1 - v^2/c^2)^{1/2}$;
- Number densities transform from the rest frame as $n = \gamma n_0$ as the total number of objects in a system is Lorentz invariant.

Consider a clock at rest in S' with period T_0 so that $x'_A = x'_B$. Let events A and B correspond to contiguous ticks so that $t'_B = t'_A + T_0$. Then in S

$$t_A = \gamma(t'_A + vx'_A/c^2) \quad \text{and} \quad t_B = \gamma(t'_A + T_0 + vx'_A/c^2)$$

$$T = t_B - t_A = \gamma T_0 \implies T = T_0(1 - v^2/c^2)^{-1/2} \tag{2.6}$$

2.5 Paths through spacetime

For a general path through spacetime, express the intrinsic geometry of Minkowski spacetime in infinitesimal form using the invariant *Minkowski line element* for infinitesimally-separated events

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2 \tag{2.7}$$

Then for a path connecting events A and B the invariant interval along the path is

$$\Delta s = \int_A^B \mathrm{d}s \tag{2.8}$$

which is path-dependent but invariant under a Lorentz transform.

A *worldline* is the path of a particle in spacetime. For a massive particle (which moves at speed less than c) passing through an event A its worldline must lie within the lightcone; for a massless particle (which moves at speed c) the worldline travels along the lightcone.

The path of a massive particle is parametrised using *proper time* τ, the time measured in the rest frame of the particle. So $\mathrm{d}\tau$ is an increment in time in the *instantaneous rest frame* (IRF) of the particle and where $\mathrm{d}x_i' = 0$. It follows that

$$c^2\mathrm{d}\tau^2 = \mathrm{d}s^2 = c^2\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2$$

Dividing the spatial terms by $\mathrm{d}t^2$ gives just the velocity squared v^2, so

$$\boxed{\frac{\mathrm{d}\tau}{\mathrm{d}t} = (1 - v^2/c^2)^{1/2} = \frac{1}{\gamma_v}} \quad \text{and} \quad \Delta\tau = \int_A^B \mathrm{d}\tau \tag{2.9}$$

2.6 Doppler effect

Let successive wavecrests (of light) be emitted by $\mathcal{E}$, moving at speed v along the x-axis of S, at events A and B, separated by proper time $\Delta\tau_{AB}$ (proper period of source), where from Eq. 2.6

$$\Delta\tau_{AB} = \left(1 - \frac{v^2}{c^2}\right)^{1/2} \Delta t_e$$

The wavecrests are received by $\mathcal{O}$ at events C and D, separated by time Δt_0 in S. $\mathcal{O}$ is at rest at x_0 in S so the proper time is $\Delta\tau_{CD} = \Delta t_0$. In Δt_e, $\mathcal{E}$ moves a distance $\Delta x_e = v\Delta t_e$; the second wavecrest travels an extra distance Δx_e to be received by $\mathcal{O}$ at x_0. Then

$$\Delta\tau_{CD} = \Delta t_0 = \left(1 + \frac{v}{c}\right) \Delta t_e$$

and the ratio of proper times is

$$\frac{\Delta\tau_{AB}}{\Delta\tau_{CD}} = \frac{(1-\beta)^{1/2}}{(1+\beta)^{1/2}} \tag{2.10}$$

which is the ratio of the received frequency (measured by $\mathcal{O}$) to the proper frequency (in the rest frame of $\mathcal{E}$).

2.7 Addition of velocities under the Lorentz transform

A particle on a worldline $x_i(t)$ for $x_i = (x, y, z)$ has velocity in S with components $u_i = \mathrm{d}x_i/\mathrm{d}t$. S' is related to S by a standard Lorentz boost. In S'

$$\begin{aligned} \mathrm{d}t' &= \gamma_v(\mathrm{d}t - v\,\mathrm{d}x/c^2) \\ \mathrm{d}x' &= \gamma_v(\mathrm{d}x - v\,\mathrm{d}t) \\ \mathrm{d}y' &= \mathrm{d}y \\ \mathrm{d}z' &= \mathrm{d}z \end{aligned} \tag{2.11}$$

Therefore

$$\begin{aligned} u'_x &= \frac{\mathrm{d}x'}{\mathrm{d}t'} = \frac{u_x - v}{1 - u_x v/c^2} \\ u'_y &= \frac{\mathrm{d}y'}{\mathrm{d}t'} = \frac{u_y}{\gamma_v(1 - u_x v/c^2)} \\ u'_z &= \frac{\mathrm{d}z'}{\mathrm{d}t'} = \frac{u_z}{\gamma_v(1 - u_x v/c^2)} \end{aligned} \tag{2.12}$$

The inverse transformations are obtained by replacing v with $-v$ and switching u'_i and u_i.

Consecutive colinear standard boosts

Let S' be related to S by a standard boost along the x-direction with speed v. Let S'' be related to S' by a standard boost along the x'-direction with speed u'. Then x and x' are in the same direction. The composition of two *colinear* boosts is a boost along this same direction; *the rapidities add*:

Consider the rapidities ψ_v and $\psi_{u'}$ both satisfying $\tanh\psi_w = w/c$, then we have, for example

$$\begin{aligned} x'' &= \cosh\psi_{u'} x' - \sinh\psi_{u'} ct' \\ &= \cosh\psi_{u'}(\cosh\psi_v x - \sinh\psi_v ct) - \sinh\psi_{u'}(\cosh\psi_v ct - \sinh\psi_v x) \\ &= \cosh(\psi_v + \psi_{u'})x - \sinh(\psi_v + \psi_{u'})ct \end{aligned} \tag{2.13}$$

and similarly for t''

$$ct'' = \cosh(\psi_v + \psi_{u'})ct - \sinh(\psi_v + \psi_{u'})x \tag{2.14}$$

So S'' is related to S by a Lorentz boost in the x-direction with speed $u = c\tanh(\psi_v + \psi_{u'})$.

2.8 Acceleration

Acceleration is not invariant but it is an *absolute* quantity, i.e. all observers agree whether a particle is accelerating or not.

Working out how the components of acceleration transform is simple but very tedious. We work out (1) the differentials of the velocity transformations in Eq. 2.12 then (2) use the chain rule:

$$\begin{aligned} a'_x &= \frac{\mathrm{d}u'_x}{\mathrm{d}t'} = \frac{a_x}{\gamma_v^3(1 - u_x v/c^2)^3} \\ a'_y &= \frac{\mathrm{d}u'_y}{\mathrm{d}t'} = \frac{a_y}{\gamma_v^3 21 - u_x v/c^2)^2} + \frac{u_y v a_x}{c^2\gamma_v^2(1 - u_x v/c^2)^3} \\ a'_z &= \frac{\mathrm{d}u'_z}{\mathrm{d}t'} = \frac{a_z}{\gamma_v^2(1 - u_x v/c^2)^2} + \frac{u_z v a_x}{c^2\gamma_v^2(1 - u_x v/c^2)^3} \end{aligned} \tag{2.15}$$

Rectilinear acceleration

Consider a particle moving at *variable* speed $u(\tau)$ along the x-axis in inertial frame S. The proper acceleration as measured by the particle in its instantaneous rest frame (IRF) at τ is $f(\tau)$. In this IRF at time τ S'

$$u'(\tau) = 0 \quad \text{and} \quad \frac{\mathrm{d}u'}{\mathrm{d}t'} = f(\tau) \implies \frac{\mathrm{d}u}{\mathrm{d}t} = \left(1 - \frac{u^2}{c^2}\right)^{3/2} f(\tau) \quad \text{in } S \tag{2.16}$$

Recall from Eq. 2.9 that $\mathrm{d}\tau = (1 - u^2/c^2)^{1/2}\,\mathrm{d}t$ to express w.r.t. proper time

$$\frac{\mathrm{d}u}{\mathrm{d}\tau} = \left(1 - \frac{u^2}{c^2}\right) f(\tau) = \frac{f(\tau)}{\gamma_u^2} \tag{2.17}$$

In terms of rapidity $\psi(\tau)$ (see Eq. 2.3) and taking $u(t=0) = 0$

$$\frac{\mathrm{d}u}{\mathrm{d}\tau} = c\,\mathrm{sech}^2\psi(\tau)\frac{\mathrm{d}\psi}{\mathrm{d}\tau} = \frac{c}{\gamma_u^2}\frac{\mathrm{d}\psi}{\mathrm{d}\tau} \implies c\frac{\mathrm{d}\psi}{\mathrm{d}\tau} = f(\tau) \implies \boxed{c\psi(\tau) = \int_0^\tau f(\tau')\,\mathrm{d}\tau'} \tag{2.18}$$

The worldline of the particle in S is parametrised by

$$\frac{\mathrm{d}t}{\mathrm{d}\tau} = \gamma_u = \cosh\psi(\tau), \quad \frac{\mathrm{d}x}{\mathrm{d}\tau} = u\,\mathrm{d}t = u\gamma_u = c\sinh\psi(\tau) \tag{2.19}$$

3 Manifolds and coordinates

3.1 Fundamentals

An N-dim manifold is a set of objects that locally resembles N-dim Euclidean space $\mathbb{R}^N$. To 'locally resemble' means there exists a map ϕ from the N-dim manifold $\mathcal{M}$ to an open subset U of $\mathbb{R}^N$ that is 1-to-1 and onto (i.e. the mapping is invertible). The whole manifold cannot be mapped to U by a single map so $\mathcal{M}$ is subdivided and each subdivision is mapped *locally*. For $\mathcal{M}$ to be *differentiable* the subdivisions must join up smoothly.

Coordinates are denoted x^a with $a = 1, \ldots, N$ for an N-dim $\mathcal{M}$. Coordinate transformations are *passive*.

$$\mathrm{d}x'^a = \frac{\partial x'^a}{\partial x'^b}\mathrm{d}x^b = J^a{}_b \mathrm{d}x^b \tag{3.1}$$

where $J^a{}_b$ is the Jacobian matrix, an $N \times N$ *invertible* transformation matrix, and Einstein summation convention is used. The determinant $J \equiv \det(J^a{}_b)$ is the Jacobian of the transformation. Provided $J \neq 0$ for some range of coordinates, the transformation can be inverted locally to give $x^a(x'^1, x'^2, \ldots, x'^n)$ using the inverse of $J^a{}_b$ as the transformation matrix, which follows from

$$\frac{\partial x'^a}{\partial x^b}\frac{\partial x^b}{\partial x'^c} = \frac{\partial x'^a}{\partial x'^c} = \delta^a_c$$

The determinant of the inverse transformation is $1/J$.

Subsets of points in a manifold define curves and surfaces (submanifolds), defined parametrically in some coordinate system, e.g. $x^a = x^a(u)$. For an $(M < N)$-dim submanifold we need M parameters:

$$x^a = x^a(u^1, u^2, \ldots, u^M), \quad a = 1, 2, \ldots, N \tag{3.2}$$

If $M = N - 1$ then the submanifold is a *hypersurface*. Then the $N - 1$ parameters can be eliminated from N equations to give $f(x^1, x^2, \ldots, x^N) = 0$ for some function f. Similarly, points in an M-dim surface can be specified by $N - M$ independent constraints

$$f(x^1, x^2, \ldots, x^N) = 0, \ldots, f_{N-M}(x^1, x^2, \ldots, x^M) = 0 \tag{3.3}$$

Example. For S^2, a 2-sphere (manifold) can be parametrised by $2 < 3$ coordinates $u^1 = \theta, u^2 = \phi$ to give a hypersurface specified by the constraint $(x^1)^2 + (x^2)^2 + (x^3)^2 = 1$ (there is $3 - 2 = 1$ independent constraint).

3.2 Local geometry of Riemannian manifolds

The *local geometry* near a point P is defined by the invariant interval $\mathrm{d}s$ between P at x^a and a neighbouring point at $x^a + \mathrm{d}x^a$, which satisfies (in a Riemannian manifold)

$$\mathrm{d}s^2 = g_{ab}(x)\,\mathrm{d}x^a\mathrm{d}x^b \tag{3.4}$$

The invariant is quadratic in the coordinate differentials and is an *intrinsic* property of the manifold. If $\mathrm{d}s^2 > 0$ then the geometry is Riemannian; otherwise it is pseudo-Riemannian.

Metric functions $g_{ab}(x)$ relate infinitesimal changes in the coordinates to invariantly-defined 'distances' in the manifold, which in GR are proper distances and proper times. Metric functions can always be chosen to be symmetric. So it follows that in an N-dim Riemannian manifold, there are $N(N+1)/2$ independent metric functions at each point. Metric functions transform as

$$\boxed{g'_{cd}(x') = g_{ab}(x(x'))\frac{\partial x^a}{\partial x'^c}\frac{\partial x^b}{\partial x'^d}} \tag{3.5}$$

which follows from using the chain rule on Eq. 3.4, in order to preserve the invariance of $\mathrm{d}s^2$. It follows that the determinant of the metric transforms as

$$g' = g/J^2 \quad \text{with} \quad J = \det\left(\frac{\partial x'^a}{\partial x^b}\right) \tag{3.6}$$

We can make N arbitrary coordinate transformations, so there are $N(N-1)/2$ independent functional degrees of freedom associated with $g_{ab}(x)$. This comes from $N(N+1)/2 - N = N(N-1)/2$.

For a Riemannian manifold it follows that g_{ab} is positive definite at every point and all eigenvalues are positive. In a pseudo-Riemannian manifold the sign of $\mathrm{d}s^2$ depends on $\mathrm{d}x^a$.

Examples. In $\mathbb{R}^2$, $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2$ and

$$g_{ab} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

In *plane polar coordinates* $x = \rho\cos\phi$ and $y = \rho\sin\phi$

$$\mathrm{d}x = \cos\phi\,\mathrm{d}\rho - \rho\sin\phi\,\mathrm{d}\phi$$
$$\mathrm{d}y = \sin\phi\,\mathrm{d}\rho + \rho\cos\phi\,\mathrm{d}\phi$$

then $x\,\mathrm{d}x + y\,\mathrm{d}y = \rho\,\mathrm{d}p$ and $\mathrm{d}x^2 + \mathrm{d}y^2 = \mathrm{d}\rho^2 + \rho^2\,\mathrm{d}\phi^2$ and

$$g_{\rho\phi} = \begin{pmatrix} 1 & 0 \\ 0 & \rho^2 \end{pmatrix}$$

Consider the 2-sphere of radius a embedded in $\mathbb{R}^3$. The embedding space has Euclidean line element $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2$ in Cartesian. The constraint on $\mathrm{d}z$ that keeps us on the spherical surface (for displacements $\mathrm{d}x$ and $\mathrm{d}y$) is

$$\mathrm{d}z = -\frac{(x\,\mathrm{d}x + y\,\mathrm{d}y)}{\sqrt{a^2 - x^2 - y^2}}$$

The *induced line element* is obtained by substituting for $\mathrm{d}z$

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \frac{(x\,\mathrm{d}x + y\,\mathrm{d}y)^2}{a^2 - (x^2 + y^2)} \tag{3.7}$$

Near the north/south poles $x^2 + y^2 \ll a^2$, and the induced line element is approximately the Euclidean form $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 = \mathrm{d}\rho^2 + \rho^2\,\mathrm{d}\phi^2$ in plane polars. As $x\,\mathrm{d}x + y\,\mathrm{d}y = \rho\,\mathrm{d}p$, substituting into Eq. 3.7 gives

$$\mathrm{d}s^2 = \frac{a^2\,\mathrm{d}\rho^2}{(a^2 - \rho^2)} + \rho^2\,\mathrm{d}\phi^2 \tag{3.8}$$

Consider now the 3-sphere in $\mathbb{R}^4$. The 3-sphere is defined by $x^2 + y^2 + z^2 + w^2 = a^2$ with line element $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 + \mathrm{d}w^2$. Differentiating gives

$$\mathrm{d}w = -\frac{(x\,\mathrm{d}x + y\,\mathrm{d}y + z\,\mathrm{d}z)}{\sqrt{a^2 - (x^2 + y^2 + z^2)}}$$

and the induced line element

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 + \frac{(x\,\mathrm{d}x + y\,\mathrm{d}y + z\,\mathrm{d}z)^2}{a^2 - (x^2 + y^2 + z^2)}$$

Using spherical polars, $x^2 + y^2 + z^2 = r^2$ and $x\,\mathrm{d}x + y\,\mathrm{d}y + z\,\mathrm{d}z = r\,\mathrm{d}r$ so that

$$\mathrm{d}s^2 = \frac{a^2}{(a^2 - r^2)}\mathrm{d}r^2 + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\,\mathrm{d}\phi^2 \tag{3.9}$$

Note:

- In the limit $a \to \infty$, we recover 3-dim Euclidean space in spherical polar coordinates;
- More generally, for $r \ll a$ we recover $\mathbb{R}^3$ locally.

Making the substitution $r = a \sin\chi$ transforms this into

$$\mathrm{d}s^2 = a^2 \left[\mathrm{d}\chi^2 + \sin^2\chi \left(\mathrm{d}\theta^2 + \sin^2\theta \, \mathrm{d}\phi^2\right)\right] \tag{3.10}$$

Lengths

Consider a curve $x^a(u)$ between points A and B. The invariant length along the curve L_{AB} is

$$L_{AB} = \int_{u_A}^{u_B} \mathrm{d}s \quad \text{with} \quad \mathrm{d}s = \left| g_{ab} \frac{\mathrm{d}x^a}{\mathrm{d}u} \frac{\mathrm{d}x^b}{\mathrm{d}u} \right|^{1/2} \mathrm{d}u \tag{3.11}$$

Note that the modulus sign is not required for a Riemmanian manifold.

Volumes

Consider the simple case where the metric is *diagonal*. If that is so then the associated coordinate system is *orthogonal*. Then

$$\mathrm{d}s^2 = g_{11}(\mathrm{d}x^1)^2 + g_{22}(\mathrm{d}x^2)^2 + \cdots + g_{NN}(\mathrm{d}x^N)^2 \tag{3.12}$$

The invariant *volume element* is defined by coordinate increments $\mathrm{d}x^1$ etc. with invariant lengths $\sqrt{g_{11}}\mathrm{d}x^1$ etc. so that for an N-dim manifold

$$\mathrm{d}V = \sqrt{|g_{11}g_{22}\ldots g_{NN}|}\, \mathrm{d}x^1 \mathrm{d}x^2 \ldots \mathrm{d}x^N$$

For a diagonal metric, we note that its determinant $g \equiv g_{ab} = g_{11}g_{22}\ldots g_{NN}$. We generalise to an arbitrary coordinate system with a metric not necessarily diagonal:

$$\boxed{\mathrm{d}V = \sqrt{|g|}\, \mathrm{d}x^1 \mathrm{d}x^2 \ldots \mathrm{d}x^N} \tag{3.13}$$

Check for invariance: consider a coordinate transformation $x^a \to x'^a$. Then using Eq. 3.6, it follows that

$$\sqrt{|g'|}\, \mathrm{d}x'^1 \mathrm{d}x'^2 \ldots \mathrm{d}x'^N = \frac{\sqrt{|g|}}{J} J \, \mathrm{d}x^1 \mathrm{d}x^2 \ldots \mathrm{d}x^N$$

Example. Surface of the 2-sphere in $\mathbb{R}^3$ The line element is as in Eq. 3.8 so the metric is diagonal with components

$$g_{11} = \frac{a^2}{a^2 - \rho^2} \quad \text{and} \quad g_{22} = \rho^2 \tag{3.14}$$

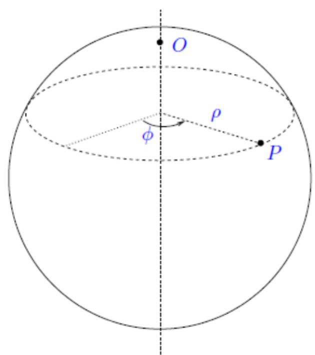


Figure 2: Circle $\rho = R$

The distance D from centre O to the perimeter along the curve $\phi = \text{const.}$ is

$$D = \int_0^R \frac{a}{(a^2-\rho^2)^{1/2}}\,\mathrm{d}\rho = a\sin^{-1}\left(\frac{R}{a}\right) \tag{3.15}$$

In 2-dim the enclosed area is the volume; using Eq. 3.13

$$A = \int_0^{2\pi}\int_0^R \frac{a}{(a^2-\rho^2)^{1/2}}\rho\,\mathrm{d}\rho\,\mathrm{d}\phi = 2\pi a^2\left[1-\left(1-\frac{R^2}{a^2}\right)^{1/2}\right] \tag{3.16}$$

From Eq. 3.15, $R = a\sin(D/a)$. Then the circumference C and enclosed area A are

$$C = 2\pi a\sin\left(\frac{D}{a}\right) \tag{3.17}$$

$$A = 2\pi a^2\left[1-\cos\left(\frac{D}{a}\right)\right] \tag{3.18}$$

Note:

- For $D \ll a$ we recover the Euclidean results $C = 2\pi D$ and $A = \pi D^2$;
- The coordinates (ρ, ϕ) are degenerate beyond the equator; g_{11} implies poor behaviour at $\rho = a$;
- If we switch to coordinates (D, ϕ), the system is well-defined beyond the equator and only degenerate at $D = \pi a$.

The metric then becomes

$$\mathrm{d}s^2 = \mathrm{d}D^2 + a^2\sin^2\left(\frac{D}{a}\right)\mathrm{d}\phi^2 \tag{3.19}$$

3.3 Local Cartesian coordinates

On a Riemannian manifold it is generally not possible to choose coordinates such that the line element takes the Euclidean form at *every* point: $g_{ab}(x)$ has $N(N+1)/2$ independent functions but there are only N functions involved in coordinate transformations. However it is always possible to adopt coordinates such that *in the neighbourhood of some point* P, the line element takes the Euclidean form. We can always find coordinates such that at P

$$g_{ab} = \delta_{ab} \quad \text{and} \quad \left.\frac{\partial g_{ab}}{\partial x^c}\right|_P = 0 \tag{3.20}$$

In the neighbourhood of P, in these *local Cartesian coordinates*

$$g_{ab}(x) = \delta_{ab} + \mathcal{O}[(x-x_P)^2] \tag{3.21}$$

In GR the generalisation of local Cartesian coordinates to spacetime corresponds to coordinates defined by locally-inertial observers. Their existence can be proved by showing that a coordinate transformation $x^a \to x'^a$ has enough degrees of freedom to bring the metric to the form in Eq. 3.21. Under the coordinate transformation, the metric and its derivatives transform as

$$g'_{ab} = \frac{\partial x^c}{\partial x'^a}\frac{\partial x^d}{\partial x'^b}g_{cd} \tag{3.22}$$

$$\frac{\partial g'_{ab}}{\partial x'^e} = \frac{\partial}{\partial x'^e}\left(\frac{\partial x^c}{\partial x'^a}\frac{\partial x^d}{\partial x'^b}\right)g_{cd} + \frac{\partial x^c}{\partial x'^a}\frac{\partial x^d}{\partial x'^b}\frac{\partial x^f}{\partial x'^e}\frac{\partial g_{cd}}{\partial x^f} \tag{3.23}$$

We now try to construct the relation $x^a(x')$ such that in the primed coordinates, $g'_{ab} = \delta_{ab}$ and $\partial g'_{ab}/\partial x'^c = 0$ at P. There are more than enough degrees of freedom in the transformation matrices to do so:

- Each transformation matrix at P has N^2 d.f. and its derivative $N^2(N+1)/2$;
- The metric and its derivatives at P have $N(N+1)/2$ and $N^2(N+1)/2$ d.f. respectively;
- We are left with $N(N-1)/2$ 'unused' degrees of freedom after N coordinate transformations in the $\partial x^a/\partial x'^b$.

For $N = 4$ spacetime, this corresponds to the six d.f. (three boosts, three rotations) associated with the homogeneous Lorentz transformations that preserve the Minkowski form of the metric.

The second derivative of the metric cannot be set to zero, as the second derivative of the metric has $N^2(N+1)/2 \times (N+1)/2 = N^2(N+1)^2/4$ degrees of freedom while the third derivatives of the transformation matrices have $N^2(N+1)(N+2)/6$. The difference of $N^2(N^2-1)/12$ degrees of freedom corresponds to the curvature of the manifold and in GR the physical degrees of freedom associated with gravity.

In a *pseudo-Riemannian manifold* there exists coordinates such that at point P

$$g_{ab} = \eta_{ab}, \quad \eta_{ab} = \mathrm{diag}(\pm 1, \pm 1, \ldots, \pm 1) \tag{3.24}$$

The *signature* of the manifold is the number of positive entries minus the number of negative entries in η_{ab}. For example, Minkowski spacetime is a pseudo-Riemannian manifold with $\eta_{ab} = \mathrm{diag}(+1, -1, -1, -1)$ and signature -2.

3.4 Topology of manifolds

The *global geometry* or *topology* is defined crudely by coincident points with different coordinates. It is an intrinsic but non-local property of a manifold.

4 Vector and tensor algebra and calculus

4.1 Scalar and vector fields on manifolds

Scalar fields

A complex scalar field defined on (some subset of) a manifold $\mathcal{M}$ assigns a complex number to each point P in (the subset of) $\mathcal{M}$. The value of the scalar field $\phi(x^a)$ at P is independent of the chosen coordinate system:

$$\phi'(x'^a) = \phi(x^a) \tag{4.1}$$

Vector fields and tangent spaces

Displacement vectors (in the traditional sense) between infinitesimally-separated points are *local vectors*. On a general manifold only local vectors can be defined. Displacement vectors are permissible if we specify an embedding of $\mathcal{M}$ in some higher-dimensional Euclidean space. Let us first consider this case, but focus on local vectors, e.g. the velocity of a particle confined to $\mathcal{M}$. This local velocity vector lies tangent to the manifold at P. For an N-dim manifold $\mathcal{M}$, the set of all possible local vectors at any point P lie in an N-dim (vector) subspace of the Euclidean embedding space. This subspace is the *tangent space* at P, $T_P(\mathcal{M})$. There are different T_Ps at different Ps so local vectors can only be added at the same P.

We generalise to remove any reference to embedding: at each point P of a general N-dim manifold $\mathcal{M}$, we can construct an N-dim tangent (vector) space $T_P(\mathcal{M})$ with local vectors as elements.

Vectors as differential operators

We now consider a vector at a point P as a *differential operator* at P which maps scalar fields $\phi(x^a)$ on $\mathcal{M}$ to a number. Then by extension a vector field is associated with a differential operator at every point. Vector fields map scalar fields to scalar fields.

With differential operators, we can explicitly construct the tangent space $T_P(\mathcal{M})$. Consider the operator at P

$$\mathbf{v} = v^a \frac{\partial}{\partial x^a} \tag{4.2}$$

Adding two operators/multiplying by a scalar returns a differential operator. So the space of all such operators at P is closed and forms a vector space.

The N partial derivative operators $\{\partial/\partial x^1, \ldots, \partial/\partial x^N\}$ at P form a set of *basis vectors* for $T_P(\mathcal{M})$. Then the v^a are associated components. Under a coordinate change these basis vectors transform as

$$\frac{\partial}{\partial x'^a} = \frac{\partial x^b}{\partial x'^a}\frac{\partial}{\partial x^b} \qquad \Longrightarrow \qquad v'^a = \frac{\partial x'^a}{\partial x^b} v^b \tag{4.3}$$

in order for $\mathbf{v}$ to remain invariant (they transform inversely), as

$$\mathbf{v} \to v'^a \frac{\partial}{\partial x'^a} = \frac{\partial x'^a}{\partial x^b}\frac{\partial x^c}{\partial x'^a} v^b \frac{\partial}{\partial x^c} = \delta^c_b v^b \frac{\partial}{\partial x^c} = v^b \frac{\partial}{\partial x^b} = \mathbf{v} \tag{4.4}$$

Any N-tuple that transforms according to Eq. 4.3 forms the components of a vector, e.g. coordinate differentials $\mathrm{d}x^a$ are the components of the infinitesimal "displacement" vector as they transform like

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

The tangent vector to a curve $x^a(u)$ has components $\mathrm{d}x^a/\mathrm{d}u$:

$$x^a(u) = \frac{\mathrm{d}x^a}{\mathrm{d}u}\frac{\partial}{\partial x^a} = \frac{\mathrm{d}}{\mathrm{d}u} \tag{4.5}$$

Dual vector fields

Under a coordinate change, objects that transform (linearly) as

$$X'_a = \frac{\partial x^b}{\partial x'^a} X_b \tag{4.6}$$

are the components of a *dual vector*.

Example. Consider the gradient of a scalar field, i.e. N-tuples such as $X_a = \partial\phi/\partial x^a$. Under a coordinate change

$$X'_a = \frac{\partial \phi'}{\partial x'^a} = \frac{\partial x^b}{\partial x'^a}\frac{\partial \phi}{\partial x^b} = \frac{\partial x^b}{\partial x'^a} X_b$$

Note that the X'_a transform inversely (relative to the components of a vector), the same way as the basis vector $\partial/\partial x^a$ - see Eq. 4.3.

Dual vectors at P span the *dual vector space* $T^*_P(\mathcal{M})$. Dual vectors are dual to vectors in the sense that the *contraction* of a dual vector X_a and a vector v^a, defined by the summation $X_a v^a$ (or the result of a linear map X_a acting on vector v^a), is *invariant* under coordinate transformations:

$$X'_a v'^a = \frac{\partial x^b}{\partial x'^a}\frac{\partial x'^a}{\partial x^c} X_b v^c = X_b v^b \tag{4.7}$$

In general, there is no invariant way to relate vectors and their duals, i.e. given a vector v^a we cannot construct a dual vector, unless there exists a *metric*, as in (pseudo-)Riemannian manifolds.

4.2 Tensor fields

A tensor at point P is a multi-linear map on tensor products of $T_P(\mathcal{M})$ and $T^*_P(\mathcal{M})$ that maps k dual vectors and l vectors at P to $\mathbb{C}$. Such a tensor is said to be of *type* (k,l) and to have *rank* $k+l$. There are k *contravariant* indices and l *covariant* indices. Offsetting the indices defines an order. Rank-0 tensors are scalar fields, type-$(1,0)$ tensors are vectors, and type-$(0,1)$ tensors are duals. Tensors are invariant under changes of coordinates. A tensor field assigns a tensor of the same type to every point in the manifold.

The components of a type-(k,l) tensor transform under a coordinate change as

$$T'^{a\ldots b}{}_{c\ldots d} = \frac{\partial x'^a}{\partial x^p}\cdots\frac{\partial x'^b}{\partial x^q}\frac{\partial x^r}{\partial x'^c}\cdots\frac{\partial x^s}{\partial x'^d} T^{p\ldots q}{}_{r\ldots s} \tag{4.8}$$

Elementary operations

Tensors under addition and scalar multiplication must be *of the same type*. A tensor with components $T_{ab} + S_{ab}$ is a tensor; under a coordinate transformation

$$T'_{ab} + S'_{ab} = \frac{\partial x^c}{\partial x'^a}\frac{\partial x^d}{\partial x'^b}(T_{cd} + S_{cd}) \tag{4.9}$$

The *tensor product* of a type-(p,q) tensor $S^{a_1\ldots a_p}{}_{b_1\ldots b_q}$ and a type-(r,s) tensor $T^{c_1\ldots c_r}{}_{d_1\ldots d_s}$ is the type-$(p+r, q+s)$ tensor

$$\boldsymbol{S}\otimes\boldsymbol{T} = S^{a_1\ldots a_p}{}_{b_1\ldots b_q} T^{c_1\ldots c_r}{}_{d_1\ldots d_s} \tag{4.10}$$

In general, the outer product does not commute, i.e. $\boldsymbol{S}\otimes\boldsymbol{T} \neq \boldsymbol{T}\otimes\boldsymbol{S}$.

Consider two vectors u^a and v^a and the outer product $T^{ab} = v^a v^b$. Under a coordinate transformation

$$T'^{ab} = \frac{\partial x'^a}{\partial x^c}\frac{\partial x'^b}{\partial x^d} T^{cd} \tag{4.11}$$

Contraction consists of summing over the same contravariant and covariant index. For a type-(k,l) tensor, contraction returns a type-$(k-1,l-1)$ tensor.

Consider $T^{ab}{}_c$; contracting over b and c returns a type-$(0,1)$ tensor (a vector) $S^a \equiv T^{ab}{}_b$. We can show that S^a is a vector by contracting $T^{ab}{}_c$ under a coordinate transformation:

$$T'^{ab}{}_c = \frac{\partial x'^a}{\partial x^p}\frac{\partial x'^b}{\partial x^q}\frac{\partial x^r}{\partial x'^c}T^{pq}{}_r \quad \text{then} \quad S'^a = \frac{\partial x'^a}{\partial x^p}\frac{\partial x'^b}{\partial x^q}\frac{\partial x^r}{\partial x'^b}T^{pq}{}_r = \frac{\partial x'^a}{\partial x^p}S^p \tag{4.12}$$

A type of inner product can be formed from combining a tensor product and a contraction: for tensors T^{ab} and S_{ab}

$$\begin{aligned} \text{outer product} \quad &\rightarrow \quad T^{ab}S_{cd} \\ \text{contracting over } b \text{ and } c \quad &\rightarrow \quad T^{ab}S_{bc} \end{aligned}$$

For a vector v^a and a dual vector X_a composition reduces to contraction: $v^a X_a$.

Symmetrisation. A type-$(0,2)$ [or type-$(2,0)$] tensor can always be decomposed into a sum of symmetric and antisymmetric parts

$$S_{ab} = \frac{1}{2}(S_{ab}+S_{ba}) + \frac{1}{2}(S_{ab}-S_{ba})$$

The *operation of [anti]symmetrising* is denoted as (c.f. [anti]commutators)

$$S_{(ab)} \equiv \frac{1}{2}(S_{ab}+S_{ba}), \qquad S_{[ab]} \equiv \frac{1}{2}(S_{ab}-S_{ba}) \tag{4.13}$$

[Anti]symmetrising on subsets of indices, for example

$$S_{(ab)c} = \frac{1}{2}(S_{abc}+S_{bac})$$

Extend to arbitrary numbers of indices n; for $S_{ab\ldots c}$

$$S_{(ab\ldots c)} = \frac{1}{n!}(\text{sum over all perms}), \qquad S_{[ab\ldots c]} = \frac{1}{n!}(\text{alternating sum over all perms}^\dagger) \tag{4.14}$$

†*Even permutations are positive and odd permutations are negative.*

The *quotient theorem* is used to test whether a set of quantities are the components of a tensor.

> If a set of quantities when contracted with an arbitrary tensor produces another tensor, the original set of quantities form the components of a tensor.

Suppose v^a are the components of an *arbitrary* vector, and denote the set of quantities as $T^a{}_{bc}$. Further suppose that the contraction $T^a{}_{bc}v^c$ transforms as the components of a type-$(1,1)$ tensor. So we require

$$T'^a{}_{bc}v'^c = T'^a{}_{bc}\frac{\partial x'^c}{\partial x^f}v^f = \frac{\partial x'^a}{\partial x^d}\frac{\partial x^e}{\partial x'^b}T^d{}_{ef}v^f$$

As v^f is arbitrary, it follows that

$$T'^a{}_{bc} = \frac{\partial x'^a}{\partial x^d}\frac{\partial x^e}{\partial x'^b}\frac{\partial x^f}{\partial x'^c}T^d{}_{ef} \tag{4.15}$$

4.3 Metric tensor

The transformation law for the metric in Eq. 3.5 implies that g_{ab} are the components of a type-$(0,2)$ tensor. This is called the *metric tensor*. The metric defines a symmetric, bilinear map from pairs of vectors to real numbers. It therefore defines a natural inner product between vectors, and provides map between vectors and their duals (a dual $v_a \equiv g_{ab}v^b$ is a contraction between the metric tensor and a vector, called 'lowering an index'):

$$\mathbf{g}(\mathbf{u},\mathbf{v}) = g_{ab}u^a v^b = u^a v_a \tag{4.16}$$

Although v^a and v_a are the components of distinct mathematical objects, we always work with manifolds equipped with a metric, so we can view them as two ways of representing the same physical object.

More generally, we can 'lower the indices' of tensors by contracting with the metric:

$$T_{ab} \equiv g_{ac}T^c{}_b, \qquad T_{abc} \equiv g_{ap}g_{bq}T^{pq}{}_c \tag{4.17}$$

Inverse metric

The matrix inverse of the metric functions $g^{ab} := (g^{-1})^{ab}$ (and where $g^{ab}g_{bc} = \delta^a_c$) transforms as a type-$(2,0)$ tensor under a coordinate change. Transforming g_{ab} using Eq. 3.5 then computing the inverse gives

$$g'^{ab} = g^{cd}\frac{\partial x'^a}{\partial x^c}\frac{\partial x'^b}{\partial x^d} \tag{4.18}$$

then

$$g'^{ab}g'_{bc} = g^{pq}\frac{\partial x'^a}{\partial x^p}\frac{\partial x'^b}{\partial x^q}g_{rs}\frac{\partial x^r}{\partial x'^b}\frac{\partial x^s}{\partial x'^c} = g^{pq}g_{qs}\frac{\partial x'^a}{\partial x^p}\frac{\partial x^s}{\partial x'^c} = \frac{\partial x'^a}{\partial x^p}\frac{\partial x^p}{\partial x'^c} = \delta^a_c$$

The inverse metric provides a map from duals to vectors; given a dual X_a, $X^a \equiv g^{ab}X_a$, called 'raising the index'. This is the inverse of the map from vectors to duals provided by the metric.

Raising one index on the metric gives the components of a type-$(1,1)$ tensor

$$\boxed{g^a{}_b = g_b{}^a = \delta^a_b} \tag{4.19}$$

which follows from $g^{ab}g_{bc} = \delta^a_c$. $g^a{}_b$ is the only rank-2 tensor whose components are the same in all coordinate systems

$$g'^a{}_b = \frac{\partial x'^a}{\partial x^c}\frac{\partial x^d}{\partial x'^b}g^c{}_d = \frac{\partial x'^a}{\partial x^c}\frac{\partial x^c}{\partial x'^b} = \delta^a_b = g^a{}_b \tag{4.20}$$

Combining the metric and its inverse, we can raise and lower indices on general tensors, for example

$$T_a{}^{bc} \equiv g_{ad}g^{ce}T^{db}{}_e \tag{4.21}$$

4.4 Scalar product of vectors

Equivalent forms of the scalar product between $\mathbf{u}$ and $\mathbf{v}$ in terms of components are

$$g_{ab}u^a v^b = g^{ab}u_a v_b = u^a v_a = u_a v^a \tag{4.22}$$

The 'length' of a vector $|\mathbf{v}|$ and generalised 'angle' θ between two non-null vectors are defined by

$$|\mathbf{v}| \equiv |g_{ab}v^a v^b|^{1/2} \quad \text{and} \quad \cos\theta \equiv \frac{u_a v^a}{|u_b u^b|^{1/2}|v_c v^c|^{1/2}} \tag{4.23}$$

Two vectors are orthogonal if their scalar product vanishes. On a Riemannian manifold $g_{ab}v^a v^b \geq 0$ with equality only if $\mathbf{v} = 0$. If pseudo-Riemannian there exists non-zero *null vectors* for which $g_{ab}v^a v^b = 0$ and $|\mathbf{v}| = 0$.

4.5 Covariant derivatives

Derivatives of scalar fields

The gradient (covariant derivative) of a scalar field ϕ is a dual vector and its components are the partial derivatives $\partial\phi/\partial x^a$ because they transform like duals (see Section 4.1)

$$\frac{\partial \phi'}{\partial x'^a} = \frac{\partial x^b}{\partial x'^a}\frac{\partial \phi}{\partial x^b} \tag{4.24}$$

It is a linear map from vectors to real numbers because it maps an infinitesimal displacement (vector with components δx^a) into

$$\delta\phi = \frac{\partial \phi}{\partial x^a}\delta x^a \tag{4.25}$$

Covariant derivatives of tensor fields

Derivatives should preserve the tensorial nature of the object. In Euclidean space, in global Cartesian coordinates, the partial derivatives of *Cartesian* components of a tensor transform as a Cartesian tensor under orthogonal coordinate transformations. But:

- Even in Euclidean space, this does not work for a general coordinate system;
- Global Cartesian coordinates do not exist on manifolds in general.

Consider a vector field $v^a(x) \mapsto v'^a(x')$ under a coordinate transformation

$$\frac{\partial v'^b}{\partial x'^a} = \frac{\partial}{\partial x'^a}\left(\frac{\partial x'^b}{\partial x^c}v^c\right) = \frac{\partial x^d}{\partial x'^a}\frac{\partial}{\partial x^d}\left(\frac{\partial x'^b}{\partial x^c}v^c\right) = \boxed{\frac{\partial x^d}{\partial x'^a}\frac{\partial x'^b}{\partial x^c}\frac{\partial v^c}{\partial x^d}} + \frac{\partial x^d}{\partial x'^a}\frac{\partial^2 x'^b}{\partial x^d \partial x^c}v^c \tag{4.26}$$

The boxed term is the usual transformation law for a type-$(1,1)$ tensor. To resolve the existence of the second term we introduce the *covariant derivative*. The covariant derivative of a type-(k,l) tensor is a type-$(k,l+1)$ tensor denoted by $\nabla_c T^{a_1\ldots a_k}{}_{b_1\ldots b_l}$ which satisfies (c.f. a derivative)

- Action on scalar fields returns the gradient: $\nabla_a\phi = \partial_a\phi$
- Linearity (usual definition)
- Leibniz rule: the covariant derivative of the *outer product* satisfies the product rule

$$\nabla_f T^{a_1\ldots a_k}{}_{b_1\ldots b_l} S^{c_1\ldots c_m}{}_{d_1\ldots d_n} = (\nabla_f T^{a_1\ldots a_k}{}_{b_1\ldots b_l}) S^{c_1\ldots c_m}{}_{d_1\ldots d_n} + T^{a_1\ldots a_k}{}_{b_1\ldots b_l}(\nabla_f S^{c_1\ldots c_m}{}_{d_1\ldots d_n}) \tag{4.27}$$

We define (for a *vector* field)

$$\boxed{\nabla_a v^b = \partial_a v^b + \Gamma^b_{ac}v^c} \tag{4.28}$$

where the *connection coefficients* Γ^b_{ac} are *not* the components of a tensor but must transform so that $\nabla_a v^b$ transforms as a type-$(1,1)$ tensor

$$\begin{aligned}
\nabla'_a v'^b &= \frac{\partial v'^b}{\partial x'^a} + \Gamma'^b_{ac}v'^c \\
&= \boxed{\frac{\partial x^d}{\partial x'^a}\frac{\partial x'^b}{\partial x^c}\frac{\partial v^c}{\partial x^d}} + \frac{\partial x^d}{\partial x'^a}\frac{\partial^2 x'^b}{\partial x^d \partial x^c}v^c + \Gamma'^b_{ac}\frac{\partial x'^c}{\partial x^d}v^d \\
&= \boxed{\frac{\partial x^d}{\partial x'^a}\frac{\partial x'^b}{\partial x^c}\nabla_d v^c - \frac{\partial x^d}{\partial x'^a}\frac{\partial x'^b}{\partial x^c}\Gamma^c_{de}v^e} + \frac{\partial x^d}{\partial x'^a}\frac{\partial^2 x'^b}{\partial x^d \partial x^c}v^c + \Gamma'^b_{ac}\frac{\partial x'^c}{\partial x^d}v^d
\end{aligned}$$

where the last three terms must vanish for *arbitrary* $\mathbf{v}$. Note that the indices should correspond to the Γ and the v indices are dummy. We relabel the Γ to Γ'^{a}_{bc} and Γ^{d}_{ef} so that (OK I have no idea what is actually going on here)

$$\Gamma'^{a}_{bc} = \frac{\partial x'^{a}}{\partial x^{d}} \frac{\partial x^{e}}{\partial x'^{b}} \frac{\partial x^{f}}{\partial x'^{c}} \Gamma^{d}_{ef} - \frac{\partial x^{d}}{\partial x'^{b}} \frac{\partial x^{e}}{\partial x'^{c}} \frac{\partial^2 x'^{a}}{\partial x^{d} \partial x^{e}} \tag{4.29}$$

The connection is generally unique up to a type-$(1, 2)$ tensor, because the difference of two connections transforms as such a tensor (the last term on the right cancels). Any coefficients that satisfy Eq. 4.29 will give a valid covariant derivative, but where there is a metric this is generally clear. It follows that the antisymmetric part of the connection transforms as a tensor (called the torsion tensor).

We now expand to more general tensor fields. For type-$(0, 1)$ dual vector fields, impose the requirement that the covariant derivative commutes with contraction. To see this, consider that for the mixed metric tensor $g^{a}{}_{b}$ the covariant derivative vanishes

$$\nabla_c g^{a}{}_{b} = \partial_c \delta^{a}_{b} + \Gamma^{a}_{cd} \delta^{d}_{b} - \Gamma^{d}_{cb} \delta^{a}_{d} = \Gamma^{a}_{cb} - \Gamma^{a}_{cb} = 0 \tag{4.30}$$

Now for a contraction of X_a and v^a, using the Leibniz rule (LHS) and the definition of action on a scalar (returns the gradient; RHS)

$$(\nabla_a X_b) v^b + X_b (\nabla_a v^b) = \nabla_a (X_b v^b) = (\partial_a X_b) v^b + X_b (\partial_a v^b)$$

The second term expands to $X_b(\partial_a v^b + \Gamma^{b}_{ac} v^c)$, then

$$(\nabla_a X_b) v^b = (\partial_a X_b) v^b + X_b (\partial_a v^b) - X_b (\partial_a v^b) - X_b \Gamma^{b}_{ac} v^c \implies \boxed{\nabla_a X_b = \partial_a X_b - \Gamma^{c}_{ab} X_c} \tag{4.31}$$

Rank-2 tensors: Consider first a type-$(2, 0)$ tensor $T^{ab} = u^a v^b$. The covariant derivative is linear, then we use the *Leibniz rule* to give

$$\begin{aligned} \nabla_a (u^b v^c) &= (\nabla_a u^b) v^c + u^b (\nabla_a v^c) \\ &= \left(\partial_a u^b + \Gamma^{b}_{ad} u^d\right) v^c + u^b \left(\partial_a v^c + \Gamma^{c}_{ad} v^d\right) \\ &= \partial_a (u^b v^c) + \Gamma^{b}_{ad} u^d v^c + \Gamma^{c}_{ad} u^b v^d \end{aligned} \tag{4.32}$$

so that for the general rank-2 tensor

$$\begin{aligned} \nabla_c T^{ab} &= \partial_c T^{ab} + \Gamma^{a}_{cd} T^{db} + \Gamma^{b}_{cd} T^{ad} \\ \nabla_c T^{a}{}_{b} &= \partial_c T^{a}{}_{b} + \Gamma^{a}_{cd} T^{d}{}_{b} - \Gamma^{d}_{cb} T^{a}{}_{d} \\ \nabla_c T_{ab} &= \partial_c T_{ab} - \Gamma^{d}_{ca} T_{db} - \Gamma^{d}_{cb} T_{ad} \end{aligned} \tag{4.33}$$

The first covariant index on each connection coefficient matches that on the covariant derivative. Each index of T must appear in one but not both connection coefficients, maintaining its contra/covariancy.

Metric connection

On a manifold equipped with a metric, two further conditions are imposed

- Metric compatibility:

$$\boxed{\nabla_a g_{bc} = \nabla_a g^{bc} = 0} \tag{4.34}$$

- Commutative action on scalar fields:

$$\boxed{\nabla_a \nabla_b \phi = \nabla_b \nabla_a \phi} \tag{4.35}$$

  The connection must be symmetric in its lower indices, $\Gamma^{a}_{[bc]} = 0$, as

$$\nabla_a \nabla_b \phi = \partial_a \partial_b \phi - \Gamma^{c}_{ab} \partial_c \phi \implies \nabla_{[a} \nabla_{b]} \phi = 0 \ \ \forall \ \phi$$

  We are concerned only with these torsion-free connections.

Then considering metric compatibility $0 = \nabla_c g_{ab} = \partial_c g_{ab} - \Gamma^d_{ca} g_{db} - \Gamma^d_{cb} g_{ad}$ and its cyclic permutations and contracting with the inverse metric

$$\boxed{\Gamma^a_{bc} = \frac{1}{2} g^{ad} (\partial_b g_{dc} + \partial_c g_{db} - \partial_d g_{bc})} \tag{4.36}$$

The order of raising/lowering indices and covariant differentiation can be exchanged as

$$\nabla_c T^{ab} = \nabla_c (g^{bd} T^a{}_d) = (\nabla_c g^{bd}) T^a{}_d + g^{bd} (\nabla_c T^a{}_d) = g^{bd} (\nabla_c T^a{}_d) \tag{4.37}$$

Connection coefficients can be summed over the upper and a lower index, e.g. Γ^a_{ab}, which is related to the derivative of the determinant of the metric functions. We use metric compatibility (Eq. 4.34):

$$g^{ab} \partial_c g_{ab} = g^{ab} (\Gamma^d_{ca} g_{db} + \Gamma^d_{cb} g_{ad}) = 2 g^{ab} g_{db} \Gamma^d_{ca} = 2 \Gamma^a_{ac}$$

where the last equality comes from contracting $g^{ab} g_{db} = \delta^a_d$ and symmetry $\Gamma^a_{[bc]} = 0$. Using Jacobi's formula for an invertible matrix M: $(\det \mathrm{M})^{-1} \partial_c \det \mathrm{M} = \mathrm{Tr}(\mathrm{M}^{-1} \partial_c \mathrm{M})$. Then

$$\boxed{\Gamma^a_{ac} = \frac{1}{2} g^{-1} \partial_c g = |g|^{-1/2} \partial_c |g|^{1/2}} \tag{4.38}$$

The metric-compatible covariant derivative reduces to partial differentiation in local (and global) Cartesian coordinates. Recall that at any point P we can find local Cartesian coordinates such that

$$g_{ab}(P) = \eta_{ab}, \quad \partial_c g_{ab}|_P = 0 \tag{4.39}$$

where the RH result implies that the metric connection also vanishes at P. Then for a tensor, the components of its covariant derivative reduce to the partial derivatives of its components.

Divergence, curl, and the Laplacian

For a vector field v the divergence is the scalar field (to see this, apply chain rule)

$$\boxed{\nabla_a v^a = \partial_a v^a + \Gamma^a_{ab} v^b = |g|^{-1/2} \partial_a \left(|g|^{1/2} v^a \right)} \tag{4.40}$$

The curl of a dual field $\boldsymbol{X}$ is the antisymmetric part of its covariant derivative, the type-$(0,2)$ tensor

$$(\mathrm{curl}\, \boldsymbol{X})_{ab} \equiv \nabla_a X_b - \nabla_b X_a = \partial_a X_b - \partial_b X_a \tag{4.41}$$

which is independent of the connection if it is symmetric (last equality). Also for a symmetric connection, the curl of a gradient $\nabla_b \phi$ vanishes by construction (Eq. 4.35).

Laplacian acting on a scalar field. If $\mathbf{v} = \nabla \phi$, then $\boxed{\nabla_a v^a = \nabla_a \nabla^a \phi \equiv \nabla^2 \phi}$. Also $v_a = \nabla_a \phi = \partial_a \phi \implies v^a = g^{ab} \partial_b \phi$. Using Eq. 4.40, the LHS becomes

$$\boxed{\nabla^2 \phi \equiv \nabla_a (g^{ab} \nabla_b \phi) = |g|^{-1/2} \partial_a (|g|^{1/2} g^{ab} \partial_b \phi)} \tag{4.42}$$

Form of Laplacian for tensor fields. Divergence of gradient, e.g.

$$\nabla^2 T^{ab} = \nabla_c \nabla^c T^{ab} = g^{cd} \nabla_c \nabla_d T^{ab} \tag{4.43}$$

4.6 Intrinsic derivatives of vectors along a curve

The intrinsic derivative of $\mathbf{v}(u)$ along the curve $x^a(u)$ is the vector, defined along $x^a(u)$, obtained by contracting the tangent vector to the curve with the covariant derivative of $\mathbf{v}$:

$$\boxed{\frac{\mathrm{D}v^a}{\mathrm{D}u} \equiv \frac{\mathrm{d}x^b}{\mathrm{d}u}\nabla_b v^a} = \frac{\mathrm{d}x^b}{\mathrm{d}u}(\partial_b v^a + \Gamma^a_{bc}v^c) = \frac{\mathrm{d}v^a}{\mathrm{d}u} + \frac{\mathrm{d}x^b}{\mathrm{d}u}\Gamma^a_{bc}v^c \tag{4.44}$$

It follows from Eq. 4.31 that the intrinsic derivative of a dual, akin to lowering an index, is

$$\frac{\mathrm{D}X_a}{\mathrm{D}u} = \frac{\mathrm{d}X_a}{\mathrm{d}u} - \frac{\mathrm{d}x^b}{\mathrm{d}u}\Gamma^c_{ba}X_c \tag{4.45}$$

4.7 Parallel transport

$\mathcal{C}$ is a curve defined in $\mathbb{R}^2$ by $x^a(u)$. The vector field $\mathbf{v}(u)$ is parallel transported along $x^a(u)$ if its vectors preserves length and direction (in Cartesian) at all points along it. For this we require $\mathrm{d}v^a/\mathrm{d}u = 0$. Extending this to a (pseudo-)Riemannian manifold, parallel transport is defined by

$$\boxed{\frac{\mathrm{D}v^a}{\mathrm{D}u} = 0} \tag{4.46}$$

Eq. 4.46 is an ODE for the v^a, with a unique solution if the v^a are specified at some initial point A. Parallel transporting from A to B along $x^a(u)$ is independent of the parametrisation; for an infinitesimal step

$$\delta v^a = \delta u \frac{\mathrm{d}v^a}{\mathrm{d}u} = -\delta u \Gamma^a_{bc}\frac{\mathrm{d}x^b}{\mathrm{d}u}v^c = -\Gamma^a_{bc}\delta x^b v^c$$

i.e. this is independent of u, where the second equality invokes Eq. 4.46.

Parallel transport is generally path-dependent unless the manifold is flat.

4.8 Geodesics

Geodesics are defined as curves $x^a(u)$ that, provided u is an affine parameter, parallel transport their tangent vector $\mathbf{t}$ with components (Eq. 4.5)

$$t^a = \frac{\mathrm{d}x^a}{\mathrm{d}u} \tag{4.47}$$

In a pseudo-Riemannian manifold, at a point, a curve is timelike, spacelike or null according to the character of its tangent vector there (via Eq. 4.16; recall the Minkowski interval)

$$\begin{aligned} \mathbf{g}(\mathbf{t},\mathbf{t}) &> 0 \quad \text{timelike;} \\ \mathbf{g}(\mathbf{t},\mathbf{t}) &= 0 \quad \text{null;} \\ \mathbf{g}(\mathbf{t},\mathbf{t}) &< 0 \quad \text{spacelike.} \end{aligned} \tag{4.48}$$

Per Eqs. 3.4 and 3.11, the length of the tangent vector for a non-null curve is

$$|\mathbf{t}| = \left| g_{ab}\frac{\mathrm{d}x^a}{\mathrm{d}u}\frac{\mathrm{d}x^b}{\mathrm{d}u}\right|^{1/2} = \left|\frac{\mathrm{d}s}{\mathrm{d}u}\right| \tag{4.49}$$

where s is the proper path length along the curve. If u is linear in s ($u = as + b$) then $|\mathbf{t}|$ is constant along the curve. Then u is an affine parameter along the curve.

Stationary points

Let $x^a(u)$ be a non-null curve between points A and B. Along the curve,

$$L = \int_A^B \mathrm{d}s = \int_{u(A)}^{u(B)} F\,\mathrm{d}u \quad \text{with} \quad F = \left|g_{ab}\dot{x}^a\dot{x}^b\right|^{1/2} = \dot{s} \tag{4.50}$$

The form of F is invariant under reparametrisation provided such a reparametrisation is monotonic in u in the interval. The *length* is invariant under arbitrary changes in the path x^a to first order provided F is extremized, i.e. the EL equations are satisfied. With such an F

$$\partial_c F = \pm\frac{1}{2F}(\partial_c g_{ab})\dot{x}^a\dot{x}^b \quad \text{and} \quad \frac{\partial F}{\partial \dot{x}^c} = \pm\frac{1}{F}g_{ac}\dot{x}^a$$

$$\implies \frac{\mathrm{d}}{\mathrm{d}u}\left(\frac{\partial F}{\partial \dot{x}^c}\right) - \partial_c F = \frac{\mathrm{d}}{\mathrm{d}u}\left(\frac{1}{\dot{s}}g_{ac}\dot{x}^a\right) - \frac{1}{2\dot{s}}(\partial_c g_{ab})\dot{x}^a\dot{x}^b = 0$$

Noting that $\dot{g}_{ac} = (\partial_b g_{ac})\dot{x}^b$, the LH term becomes

$$\frac{\mathrm{d}}{\mathrm{d}u}\left(\frac{1}{\dot{s}}g_{ac}\dot{x}^a\right) = \frac{1}{\dot{s}}\left[(\partial_b g_{ac})\dot{x}^a\dot{x}^b + g_{ac}\ddot{x}^a - \frac{\ddot{s}}{\dot{s}}g_{ac}\dot{x}^a\right]$$

$$\implies g_{ac}\ddot{x}^a + (\partial_b g_{ac})\dot{x}^a\dot{x}^b - \frac{1}{2}(\partial_c g_{ab})\dot{x}^a\dot{x}^b = \frac{\ddot{s}}{\dot{s}}g_{ac}\dot{x}^a$$

Exchanging dummy indices a and b we note that $(\partial_b g_{ac})\dot{x}^a\dot{x}^b = \frac{1}{2}(\partial_a g_{bc} + \partial_b g_{ac})\dot{x}^a\dot{x}^b$ so

$$g_{ac}\ddot{x}^a + \frac{1}{2}(\partial_a g_{bc} + \partial_b g_{ac} - \partial_c g_{ab})\dot{x}^a\dot{x}^b = \frac{\ddot{s}}{\dot{s}}g_{ac}\dot{x}^a$$

Finally, contracting with g^{dc} and relabelling indices, we note that the second term takes the form of a Christoffel symbol, and the entire LHS is just the intrinsic derivative of t^a, per Eq. 4.44

$$\frac{\mathrm{D}t^a}{\mathrm{D}u} \equiv \ddot{x}^a + \Gamma^a_{bc}\dot{x}^b\dot{x}^c = \left(\frac{\ddot{s}}{\dot{s}}\right)\dot{x}^a \tag{4.51}$$

Note that if u is an affine parameter (recall $u = as + b$) then the RHS is zero. Thus if u is affine then the tangent vector is parallel transported (we recover Eq. 4.46). This is the geodesic equation

$$\boxed{\frac{\mathrm{D}t^a}{\mathrm{D}u} \equiv \ddot{x}^a + \Gamma^a_{bc}\dot{x}^b\dot{x}^c = 0} \tag{4.52}$$

Lagrangian procedure and conserved quantities

Lowering the index on Eq. 4.52 gives, via Eq. 4.45, and explicitly writing the connection coefficients,

$$\frac{\mathrm{D}t_a}{\mathrm{D}u} = \frac{\mathrm{d}t_a}{\mathrm{d}u} - \Gamma^c_{ba}t^b t_c = \frac{\mathrm{d}t_a}{\mathrm{d}u} - \frac{1}{2}g^{cd}(\partial_b g_{ad} + \partial_a g_{bd} - \partial_d g_{ab})t^b t_c = 0$$

$$\implies \dot{t}_a = \frac{1}{2}(\partial_b g_{ad} + \partial_a g_{bd} - \partial_d g_{ab})t^b t^d$$

In the RH (connection) term we note that on exchanging dummy indices b and d, the first and third terms cancel

$$(\partial_b g_{ad} - \partial_d g_{ab})t^b t^d = \frac{1}{2}(\partial_b g_{ad} + \partial_d g_{ab} - \partial_d g_{ab} - \partial_b g_{ad})t^b t^d = 0$$

This gives an alternative form of the geodesic equation

$$\boxed{\dot{t}_a = \frac{1}{2}\partial_a g_{bc} t^b t^c} \tag{4.53}$$

which on inspection is just the EL equations, up to a factor of two, for the Lagrangian

$$L = g_{ab}\dot{x}^a\dot{x}^b \quad \text{with} \quad t_a = g_{ab}\dot{x}^b \tag{4.54}$$

L does not depend explicitly on u so the first integral gives

$$L - \dot{x}^a\frac{\partial L}{\partial \dot{x}^a} = g_{ab}\dot{x}^a\dot{x}^b - 2g_{ab}\dot{x}^a\dot{x}^b = \text{const.} \implies L = g_{ab}\dot{x}^a\dot{x}^b = \text{const.} \tag{4.55}$$

If $\partial_a g_{bc} = 0$ then per Eq. 4.53, $t_a = \text{const.}$, i.e. if the metric has no dependence on a coordinate x^a then the ath component of the dual tangent vector is conserved. Conversely if conjugate momentum $\pi_a \equiv \partial L/\partial\dot{x}^a$ is conserved then $\partial_a g_{bc} = 0$; L has no dependence on x^a.

For null curves the path length vanishes so the stationary property cannot be used to define a null geodesic. Null geodesics are defined as curves with null tangent vector satisfying Eq. 4.52. More generally, given an initial vector, a unique geodesic in an affine parametrisation can be generated by solving Eqs. 4.47 and 4.52. Parallel transport preserves character so such curves are timelike, spacelike or null *everywhere*.

5 Minkowski spacetime and particle dynamics

Special relativity revisited.

5.1 Minkowski spacetime in Cartesian coordinates

Minkowski spacetime is a 4-dim pseudo-Euclidean manifold. We adopt a *global* system of Cartesian coordinates x^μ $(\mu = 0, 1, 2, 3)$ such that the line element is everywhere

$$\mathrm{d}s^2 = \eta_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu \tag{5.1}$$

where $\eta_{\mu\nu} = \mathrm{diag}(+1, -1, -1, -1)$ is the Minkowski metric. The components of the inverse metric are also $\eta^{\mu\nu} = \eta_{\mu\nu}$. The metric connection $\Gamma^\mu_{\nu\sigma}$ vanishes in Cartesian coordinates.

Lorentz transformations

Lorentz transformations correspond to coordinate transformations $x^\mu \to x'^\mu$ that leaves the Minkowski metric invariant. The requirement for a Lorentz transformation is thus

$$\eta_{\mu\nu} = \frac{\partial x'^\rho}{\partial x^\mu}\frac{\partial x'^\sigma}{\partial x^\nu}\eta_{\rho\sigma} \tag{5.2}$$

which implies (by differentiation) that Lorentz transforms are linear. The *inhomogeneous Lorentz transform / Poincaré transformation* is

$$\boxed{x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu} \quad \text{with} \quad \eta_{\mu\nu} = \Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu \eta_{\rho\sigma} \tag{5.3}$$

a^μ is a constant which corresponds to changing the spacetime origin. If $a^\mu = 0$ then Eq. 5.3 is a *homogeneous Lorentz transformation*:

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \tag{5.4}$$

Homogeneous (and proper) Lorentz transformations

The $\Lambda^\mu{}_\nu$ depend on the relative velocity and orientation between the two inertial frames. For a standard Lorentz boost with $v = \beta c$

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.5}$$

with inverse $\Lambda_\mu{}^\nu := (\Lambda^{-1})^\nu{}_\mu = \partial x^\nu / \partial x'^\mu$. From Eq. 5.3

$$\Lambda_\mu{}^\nu = \eta_{\mu\rho}\eta^{\nu\sigma}\Lambda^\rho{}_\sigma \tag{5.6}$$

These are NOT raising and lowering indices on $\Lambda^\mu{}_\nu$ as they are not components of a tensor!

We consider only proper Lorentz transformations, which form a subgroup of the full LT: includes transformations between inertial frames with the same spatial-handedness but excludes time reversal. Properties:

$$\boxed{\det(\Lambda^\mu{}_\nu) = 1, \quad \Lambda^0{}_0 \geq 1} \tag{5.7}$$

On a general manifold, the set of basis vectors ∂_a that span T_P are represented as arrow tangents to the associated coordinate curve. The inner product between two basis vectors ∂_a and ∂_b returns g_{ab}. In Minkowski space the basis vectors $\mathbf{e}_\mu \equiv \partial_\mu$ are orthonormal as

$$\mathbf{g}(\mathbf{e}_\mu, \mathbf{e}_\nu) = \eta_{\mu\nu} \tag{5.8}$$

Basis vectors Lorentz-transform with the inverse transformation matrix

$$\mathbf{e}'_\mu = \Lambda_\mu{}^\nu \mathbf{e}_\nu \tag{5.9}$$

4-vectors are vectors in 4-dim spacetime. Under a LT, the coordinate components transform as

$$v'^\mu = \Lambda^\mu{}_\nu v^\nu \tag{5.10}$$

Components of the dual vector associated with v^μ are $v_\mu = \eta_{\mu\nu} v^\nu$ and transform like basis vectors.

Per Eq. 4.48, the character of $\mathbf{g}(\mathbf{v}, \mathbf{v})$ characterises the *vector* $\mathbf{v}$; in Cartesian coordinates

$$\begin{aligned} \eta_{\mu\nu} v^\mu v^\nu &> 0 \quad \text{timelike;} \\ \eta_{\mu\nu} v^\mu v^\nu &= 0 \quad \text{null;} \\ \eta_{\mu\nu} v^\mu v^\nu &< 0 \quad \text{spacelike.} \end{aligned} \tag{5.11}$$

In an inertial frame, $\mathbf{e}_0$ is timelike while the $\mathbf{e}_i$ are spacelike. A timelike/null vector is future-pointing if $v^0 > 0$ and past-pointing if $v^0 < 0$. These characterisations are invariant under *proper* LTs, which is the same as saying that the temporal ordering of causally-connected events is Lorentz-invariant.

The set of *all* null vectors at a point P defines the lightcone.

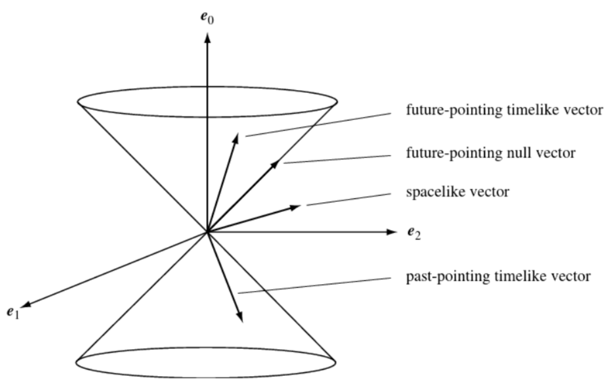


Figure 3: The lightcone separates timelike and spacelike vectors.

5.2 Particle dynamics

Velocity and acceleration 4-vectors

The worldline is parametrised by proper time τ which is related to the invariant path length by $\mathrm{d}s^2 = c^2\,\mathrm{d}\tau^2$. τ is affine parameter for the worldline. The tangent vector to the worldline is the *4-velocity* of the particle with components

$$u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}t}{\mathrm{d}\tau}\left(c, \frac{\mathrm{d}x}{\mathrm{d}t}, \frac{\mathrm{d}y}{\mathrm{d}t}, \frac{\mathrm{d}z}{\mathrm{d}t}\right) = \boxed{\gamma_u(c, \vec{u})} \tag{5.12}$$

which for a massive particle is future-pointing and timelike. For notation simplicity we write the 3-velocity as $\vec{u} = (\vec{u}^i)$ with $\vec{u}^i = \mathrm{d}x^i/\mathrm{d}t$. To 'derive' $\mathrm{d}t/\mathrm{d}\tau$ first consider that the length of the 4-velocity is constant along the worldline. This follows from τ being an affine parameter

$$|\mathbf{u}|^2 = \eta_{\mu\nu} u^\mu u^\nu = \left(\frac{\mathrm{d}s}{\mathrm{d}\tau}\right)^2 = c^2 \tag{5.13}$$

We note that the 4-velocity is a *future-pointing, timelike* vector. From this 'normalisation' we also recover Eq. 2.9:

$$c^2 = \eta_{\mu\nu} u^\mu u^\nu = \left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 (c^2 - |\vec{u}|^2) \implies \frac{\mathrm{d}t}{\mathrm{d}\tau} = \left(1 - \frac{|\vec{u}|^2}{c^2}\right)^{-1/2} = \gamma_u \tag{5.14}$$

Components of the 4-velocity transform as Eq. 5.10.

Free massive particles (constant $\vec{u}$) move on timelike geodesics with the EoM

$$0 = \frac{\mathrm{D}u^\mu}{\mathrm{D}\tau} \tag{5.15}$$

Under external forces the particle accelerates. The 4-acceleration is

$$\boxed{a^\mu = \frac{\mathrm{D}u^\mu}{\mathrm{D}\tau}} \tag{5.16}$$

In both cases the RHS reduces to $\mathrm{d}u^\mu/\mathrm{d}\tau$ in Cartesian coordinates. (Inertial frames are local Cartesian frames.)

a is always orthogonal to the 4-velocity. In Cartesian inertial coordinates, for example

$$\eta_{\mu\nu} a^\mu u^\nu = \eta_{\mu\nu} \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} u^\nu = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\tau}(\eta_{\mu\nu} u^\mu u^\nu) = 0 \implies \mathbf{g}(\mathbf{a}, \mathbf{u}) = 0 \tag{5.17}$$

Per Eqs. 5.12 and 5.16 the components of $\mathbf{a}$ are

$$a^\mu = \gamma_u \frac{\mathrm{d}}{\mathrm{d}t}(\gamma_u c, \gamma_u \vec{u}) = \gamma_u^2 \left(\frac{\gamma_u^2}{c}\,\vec{u}\cdot\vec{a},\ \vec{a} + \frac{\gamma_u^2}{c^2}(\vec{u}\cdot\vec{a})\vec{u}\right) \tag{5.18}$$

where $\vec{a} \equiv \mathrm{d}\vec{u}/\mathrm{d}t$ is the 3-acceleration and per Eq. 5.14

$$\frac{\mathrm{d}\gamma_u}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(1 - \frac{\vec{u}\cdot\vec{u}}{c^2}\right)^{-1/2} = \frac{\gamma_u^3}{c^2}\vec{u}\cdot\vec{a} \tag{5.19}$$

In the IRF of the particle, $\vec{u} = 0$, so simply have $a^\mu = (0, \vec{a}_{\mathrm{IRF}})$. Similar to (but not exactly like) the 4-velocity (see Eq. 5.13), the length/magnitude of the 4-acceleration is invariant (but not constant)

$$|\mathbf{a}|^2 = \eta_{\mu\nu} a^\mu a^\nu = -|\vec{a}_{\mathrm{IRF}}|^2 \tag{5.20}$$

Thus 4-acceleration is a *spacelike* vector.

Momentum and force 4-vectors

The 4-momentum of a massive particle is the *future-pointing, timelike* vector (as per $\mathbf{u}$)

$$\mathbf{p} = m\mathbf{u} \tag{5.21}$$

where m is the rest mass. In some inertial frame the components are

$$p^\mu = (\gamma_u mc, \gamma_u m\vec{u}) = \boxed{(E/c, \vec{p})} \tag{5.22}$$

where $\vec{p}$ is the *relativistic generalisation* of the 3-momentum. Note:

- $\vec{p}$ reduces to the non-relativistic limit $\vec{p} \approx m\vec{u}$ for $|\vec{u}| \ll c$;

- $\vec{p}$ is constant for a free particle (for which $\vec{u}$ is constant);
- For a system of point particles undergoing collisional interactions, the *sum of the individual 3-momenta* of all particles is conserved;
- The time component of $\mathbf{p}$ is the total energy $\boxed{E = \gamma_u m c^2}$. To see this consider the rate of work done when $\vec{f} = \mathrm{d}\vec{p}/\mathrm{d}t$ accelerates a particle, recalling Eqs. 5.18 and 5.19

$$\vec{u} \cdot \vec{f} = \vec{u} \cdot \frac{\mathrm{d}}{\mathrm{d}t}(\gamma_u m \vec{u}) = \gamma_u m \left(\vec{u} \cdot \vec{a} + \gamma_u^2 (\vec{u} \cdot \vec{a}) \frac{|\vec{u}|^2}{c^2} \right) = \gamma_u^3 m \, \vec{u} \cdot \vec{a} = mc^2 \frac{\mathrm{d}\gamma_u}{\mathrm{d}t}$$

  Then the result follows from $\mathrm{d}E/\mathrm{d}t = \vec{u} \cdot \vec{f}$.

It follows from Eq. 5.15 that for a *free particle* the total 4-momentum is constant

$$\frac{\mathrm{D}p^\mu}{\mathrm{D}\tau} = 0 \tag{5.23}$$

and from Eq. 5.13 that along the particle's worldline

$$|\mathbf{p}|^2 = m^2 c^2 \tag{5.24}$$

The energy-momentum (E-p) invariant follows from considering the inner product $\eta_{\mu\nu} p^\mu p^\nu$ in an inertial frame, invoking Eq. 5.22

$$E^2 - |\vec{p}|^2 c^2 = |\mathbf{p}|^2 c^2 = m^2 c^4 \tag{5.25}$$

Under external forces the 4-momentum is not constant. The 4-force $\mathbf{f}$ is defined as

$$\boxed{f^\mu = \frac{\mathrm{D}p^\mu}{\mathrm{D}\tau}} \tag{5.26}$$

and is related to the 4-acceleration via $\mathbf{f} = m\mathbf{a}$. In some inertial frame, using $\mathrm{d}E/\mathrm{d}t = \vec{f} \cdot \vec{u}$

$$f^\mu = \gamma_u \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{E}{c}, \vec{p} \right) = \boxed{\gamma_u \left(\frac{\vec{f} \cdot \vec{u}}{c}, \vec{f} \right)} \tag{5.27}$$

Similar to Eq. 5.17, p^μ is orthogonal to $f^\mu \equiv \mathrm{D}p^\mu/\mathrm{D}\tau$ so the 4-velocity and 4-force are orthogonal

$$\mathbf{g}(\mathbf{f}, \mathbf{u}) = 0 \tag{5.28}$$

This result is also obtained by considering $\eta_{\mu\nu} f^\mu u^\nu$ explicitly using component forms Eqs. 5.12 and 5.27.

4-momentum of a photon

From the E-p invariant, the limit for zero rest mass is $E = |\vec{p}|c$. Then the 4-momentum becomes a future-pointing *null* vector

$$p^\mu = (|\vec{p}|, \vec{p}) \implies \mathbf{g}(\mathbf{p}, \mathbf{p}) = 0 \tag{5.29}$$

Let the photon worldline be $x^\mu(\lambda)$ for arbitrary parameter λ. We cannot use τ as photons travel at c and the worldline is a null curve ($\mathrm{d}\tau = 0$). Instead we adopt a parametrisation such that the tangent vector to the path is the 4-momentum:

$$\boxed{p^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}} \tag{5.30}$$

This works because p^μ is always parallel to the tangent vector $\mathrm{d}x^\mu/\mathrm{d}\lambda$ for any λ. To see this consider in an inertial frame, the second term is just a unit vector in the direction of motion

$$p^\mu = \frac{E}{c}\left(1, \frac{\vec{p}}{|\vec{p}|}\right) = \frac{E}{c^2}\left(c\frac{\mathrm{d}t}{\mathrm{d}t}, \frac{\mathrm{d}x}{\mathrm{d}t}, \frac{\mathrm{d}y}{\mathrm{d}t}, \frac{\mathrm{d}z}{\mathrm{d}t}\right) = \frac{E}{c^2}\frac{\mathrm{d}x^\mu}{\mathrm{d}t}$$

With this tangent vector construction it follows that $x^\mu(\lambda)$ is an affinely-parametrised *null* geodesic

$$\frac{\mathrm{D}p^\mu}{\mathrm{D}\lambda} \equiv \frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\nabla_\nu p^\mu = p^\nu \nabla_\nu p^\mu = 0 \tag{5.31}$$

We introduce the 4-wavevector $\mathbf{k}$ with $\mathbf{p} = \hbar\mathbf{k}$ with components in some inertial frame

$$k^\mu = \left(\frac{2\pi}{\lambda}, \vec{k}\right) \tag{5.32}$$

where λ is the wavelength in that frame and $\vec{k} = (2\pi/\lambda)\vec{n}$ where $\vec{n}$ is a unit 3-vector in the direction of propagation.

Doppler effect revisited

Consider an observer $\mathcal{O}$ at rest in inertial frame S observing light with wavelength λ propagating at an angle θ to the x-axis. In S

$$k^\mu = \frac{2\pi}{\lambda}(1, \cos\theta, \sin\theta, 0) \tag{5.33}$$

The light is emitted by a source moving at speed $v = \beta c$ along the x-axis. The rest frame of the source is S'. In S', $k'^\mu = \Lambda^\mu{}_\nu k^\nu$, and

$$k'^0 = \frac{2\pi}{\lambda'} = \frac{2\pi}{\lambda}\gamma(1-\beta\cos\theta) \implies \frac{\lambda}{\lambda'} = \gamma(1-\beta\cos\theta) \tag{5.34}$$

from considering the inner product of the first row of the transformation matrix (Eq. 5.5) with k^ν. For $\theta = 0$ we recover the result derived previously (Eq. 2.10).

5.3 Compton scattering

Describes scattering of a photon from a charged particle. We consider this as a collision between a photon with initial 4-momentum $\mathbf{p}$ and an electron with initial 4-momentum $\mathbf{q}$. In the final state, the particles have 4-momentum $\bar{\mathbf{p}}$ and $\bar{\mathbf{q}}$ respectively. We consider the collision in the inertial frame in which the electron is initially at rest. The photon is propagating along the positive x-direction with frequency ν. After the collision, the photon scatters through an angle θ with final frequency $\bar{\nu}$. The electron recoils. So we have

$$\begin{aligned} p^\mu &= \left(\frac{h\nu}{c}, \frac{h\nu}{c}, 0, 0\right) \\ q^\mu &= (m_e c, 0, 0, 0) \\ \bar{p}^\mu &= \left(\frac{h\bar{\nu}}{c}, \frac{h\bar{\nu}}{c}\cos\theta, \frac{h\bar{\nu}}{c}\sin\theta, 0\right) \end{aligned} \tag{5.35}$$

Conservation of total 4-momentum:

$$\mathbf{p} + \mathbf{q} - \bar{\mathbf{p}} = \bar{\mathbf{q}} \tag{5.36}$$

where the LH terms are evaluated in the initial rest frame of the electron and the RH terms are evaluated in the final rest frame. The squared magnitudes of each side should be equivalent. The LHS expands to

$$|\mathbf{p}|^2 + |\mathbf{q}|^2 + |\bar{\mathbf{p}}|^2 + 2(\mathbf{g}(\mathbf{p},\mathbf{q}) + \mathbf{g}(\mathbf{p},\bar{\mathbf{p}}) + \mathbf{g}(\mathbf{q},\bar{\mathbf{p}}))$$

We have $|\mathbf{p}|^2 = |\bar{\mathbf{p}}|^2 = 0$ and $|\mathbf{q}|^2 = |\bar{\mathbf{q}}|^2 = m_e^2 c^2$ as both frames are on the electron's worldline. Then

$$0 = \eta_{\mu\nu} p^\mu q^\nu - \eta_{\mu\nu} p^\mu \bar{p}^\nu - \eta_{\mu\nu} q^\mu \bar{p}^\nu = h\nu m_e - h\bar{\nu} m_e - \frac{h^2}{c^2}\nu\bar{\nu}(1-\cos\theta)$$

using the components defined above. Rearranging,

$$\bar{\nu} = \frac{\nu}{1 + (h\nu/m_e c^2)(1-\cos\theta)} \tag{5.37}$$

Generally $\bar{\nu} < \nu$; energy is transferred to that of the recoiling electron.

5.4 Local reference frame of a general observer

Consider a general observer $\mathcal{O}$ following a worldline $x^\mu(\tau)$, with 4-velocity and 4-acceleration corresponding to Eqs. 5.12 and 5.16 respectively. At proper time τ, the coordinate basis vectors of the IRF at the observer's position constitute an orthonormal set of timelike basis vectors $\mathbf{e}_\mu(\tau)$. By construction, $\mathbf{e}_0(\tau)$ is equal (up to a factor of c) to the instantaneous 4-velocity $\mathbf{u}(\tau)$. The three spacelike basis vectors $\mathbf{e}_i(\tau)$ are therefore orthogonal to the observer's 4-velocity.

At some later time τ', $\mathbf{e}_0(\tau')$ is uniquely determined by $\mathbf{u}(\tau')$ but $\mathbf{e}_i(\tau')$ are only determined up to a spatial rotation. We require that they point along directions specified by three orthogonal gyroscopes carried by $\mathcal{O}$. These basis vectors, called an *orthonormal tetrad*, form the idealisation of a local laboratory for an arbitrary observer. They satisfy

$$\mathbf{g}(\mathbf{e}_\mu, \mathbf{e}_\nu) = \eta_{\mu\nu} \quad \text{and} \quad c\mathbf{e}_0(\tau) = \mathbf{u}(\tau) \tag{5.38}$$

If the observer is non-accelerating then the $\mathbf{e}_i(\tau)$ are parallel transported along $x^\mu(\tau)$. If $\mathcal{O}$ is accelerating the $\mathbf{e}_i(\tau)$ *cannot* be parallel transported since they must remain orthogonal to $\mathbf{u}(\tau)$. Instead a 'Fermi-Walker transport' is applied where the $\mathbf{e}_i(\tau)$ are parallel transported to the observer's new position at time $\tau + \mathrm{d}\tau$ then a pure Lorentz boost is applied to boost the parallel-transported $\mathbf{e}_0$ onto $\mathbf{u}(\tau + \mathrm{d}\tau)$.

5.5 Minkowski space in a non-inertial coordinate system

Let $X^\mu = (cT, X, Y, Z)$ be Cartesian coordinates of an *inertial* frame S. Introduce new coordinates $x^\mu = (ct, x, y, z)$ where points with fixed x, y, z *rotate* with angular speed ω around the Z axis in S

$$\begin{aligned} X &= x\cos\omega t - y\sin\omega t \\ Y &= x\sin\omega t + y\cos\omega t \\ Z &= z, \ \ T = t \end{aligned} \tag{5.39}$$

Evaluating the differentials

$$\begin{aligned} \mathrm{d}X &= \mathrm{d}x\cos\omega t - \mathrm{d}y\sin\omega t - \omega\,\mathrm{d}t\,(x\sin\omega t + y\cos\omega t) \\ \mathrm{d}Y &= \mathrm{d}x\sin\omega t + \mathrm{d}y\cos\omega t + \omega\,\mathrm{d}t\,(x\cos\omega t - y\sin\omega t) \end{aligned}$$

$$\mathrm{d}X^2 + \mathrm{d}Y^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \omega^2(x^2+y^2)\,\mathrm{d}t^2 + 2\omega\,\mathrm{d}t\,(x\mathrm{d}y - y\mathrm{d}x)$$

The line element in S becomes, in terms of the x^μ coordinates,

$$\mathrm{d}s^2 = [c^2 - \omega^2(x^2+y^2)]\,\mathrm{d}t^2 + 2\omega y\,\mathrm{d}t\,\mathrm{d}x - 2\omega x\,\mathrm{d}t\,\mathrm{d}y - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2 \tag{5.40}$$

For a free particle travelling on a timelike geodesic, this gives the Lagrangian $L = g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$ to be (parametrised by proper time τ)

$$L = [c^2 - \omega^2(x^2+y^2)]\dot{t}^2 + 2\omega y\dot{t}\dot{x} - 2\omega x\dot{t}\dot{y} - \dot{x}^2 - \dot{y}^2 - \dot{z}^2 \tag{5.41}$$

For $t(\tau)$, L has no explicit dependence on t, so the EL equations give

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left([c^2 - \omega^2(x^2+y^2)]\dot{t} + \omega(y\dot{x} - x\dot{y})\right) = 0$$
$$\implies [c^2 - \omega^2(x^2+y^2)]\ddot{t} - 2\omega^2(x\dot{x} + y\dot{y})\dot{t} + \omega(y\ddot{x} - x\ddot{y}) = 0 \tag{5.42}$$

and

$$\begin{aligned}
\ddot{x} &= \omega y\ddot{t} + \omega^2 x\dot{t}^2 + 2\omega\dot{t}\dot{y} \\
\ddot{y} &= -\omega x\ddot{t} + \omega^2 y\dot{t}^2 - 2\omega\dot{t}\dot{x} \\
\ddot{z} &= 0
\end{aligned}$$

Substituting into Eq. 5.42 we find that $\ddot{t} = 0$. Parametrising in terms of t with $\mathrm{d}\tau/\mathrm{d}t = \mathrm{const.}$ (just force a variable change and apply chain rule generously), we recover the EoMs for a free particle in a rotating frame

$$\begin{aligned}
\frac{\mathrm{d}^2x}{\mathrm{d}t^2} &= \omega^2 x + 2\omega\frac{\mathrm{d}y}{\mathrm{d}t} \\
\frac{\mathrm{d}^2y}{\mathrm{d}t^2} &= \omega^2 y - 2\omega\frac{\mathrm{d}x}{\mathrm{d}t} \\
\frac{\mathrm{d}^2z}{\mathrm{d}t^2} &= 0
\end{aligned} \tag{5.43}$$

6 Electromagnetism

Maxwell's equations in free space, expressed in terms of Cartesian coordinates in some inertial frame S are

$$\vec{\nabla}\cdot\vec{E} = \frac{\rho}{\epsilon_0} \tag{6.1}$$

$$\vec{\nabla}\times\vec{E} = -\frac{\partial\vec{B}}{\partial t} \tag{6.2}$$

$$\vec{\nabla}\cdot\vec{B} = 0 \tag{6.3}$$

$$\vec{\nabla}\times\vec{B} = \mu_0\vec{J} + \mu_0\epsilon_0\frac{\partial\vec{E}}{\partial t} \tag{6.4}$$

ρ and $\vec{J}$ are the charge and current densities in S respectively.

Charge conservation: take the divergence of Eq. 6.4 and using Eq. 6.1

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\vec{J} = 0 \tag{6.5}$$

6.1 Lorentz force law

$$\vec{f} = q(\vec{E} + \vec{u}\times\vec{B}) \tag{6.6}$$

$\vec{f}$ is the electromagnetic 3-force. This depends linearly on the 3-velocity $\vec{u}$. Let us extend this property to the 4-vector equivalents. We introduce a type-$(0,2)$ Maxwell field-strength tensor $\mathbf{F}$ with components $F_{\mu\nu}$ so that

$$f_\mu = qF_{\mu\nu}u^v \implies f^\mu = qF^\mu{}_\nu u^\nu \tag{6.7}$$

In the presence of this force the EoM becomes

$$\frac{\mathrm{D}u^\mu}{\mathrm{D}\tau} = \frac{q}{m}F^\mu{}_\nu u^\nu \tag{6.8}$$

As the 4-force must be orthogonal to the 4-velocity ($f_\mu u^\mu = 0$), we require $\mathbf{F}$ to be antisymmetric

$$0 = f_\mu u^\mu = qF_{\mu\nu}u^\mu u^\nu = \frac{1}{2}q(F_{\mu\nu} + F_{\nu\mu})u^\mu u^\nu \implies F_{\mu\nu} = -F_{\nu\mu} \tag{6.9}$$

This antisymmetry is preserved even after raising both indices.

The components in any inertial frame encode the E- and B-fields there. The $F_{\mu\nu}$ can be found by comparing Eq. 6.6 with Eq. 6.7: for f_0 we have

$$f_0 = q\sum_i F_{0i}u^i = \gamma_u\frac{\vec{f}\cdot\vec{u}}{c} = \frac{q}{c}\gamma_u\vec{E}\cdot\vec{u}$$

as $\vec{f}\cdot\vec{u} = q\vec{E}\cdot\vec{u}$, so $\boxed{F_{0i} = \vec{E}^i/c}$. For the spatial components

$$f_i = qF_{i0}u^0 + q\sum_j F_{ij}u^j = -\gamma_u\vec{f} = -q\gamma_u\left(\vec{E}^i + (\vec{u}\times\vec{B})^i\right)$$

Note the minus sign due to contracting with the Minkowski metric. Recall $u^0 = \gamma_u c$, and that $F_{\mu\nu}$ is antisymmetric, so that $qF_{i0}u^0 = -q\gamma_u\vec{E}^i$ so that the term cancels, and we are left with

$$qF_{ij}u^j = -q\gamma_u\epsilon_{ijk}u^jB^k \implies \boxed{F_{ij} = -\epsilon_{ijk}B^k} \tag{6.10}$$

So that

$$[F_{\mu\nu}] = \begin{pmatrix} 0 & \vec{E}^1/c & \vec{E}^2/c & \vec{E}^3/c \\ -\vec{E}^1/c & 0 & -\vec{B}^3 & \vec{B}^2 \\ -\vec{E}^2/c & \vec{B}^3 & 0 & -\vec{B}^1 \\ -\vec{E}^3/c & -\vec{B}^2 & \vec{B}^1 & 0 \end{pmatrix} \tag{6.11}$$

Contracting with $\eta_{\mu\nu}$

$$[F^{\mu}{}_{\nu}] = \begin{pmatrix} 0 & \vec{E}^1/c & \vec{E}^2/c & \vec{E}^3/c \\ \vec{E}^1/c & 0 & \vec{B}^3 & -\vec{B}^2 \\ \vec{E}^2/c & -\vec{B}^3 & 0 & \vec{B}^1 \\ \vec{E}^3/c & \vec{B}^2 & -\vec{B}^1 & 0 \end{pmatrix}, \quad [F^{\mu\nu}] = \begin{pmatrix} 0 & -\vec{E}^1/c & -\vec{E}^2/c & -\vec{E}^3/c \\ \vec{E}^1/c & 0 & -\vec{B}^3 & \vec{B}^2 \\ \vec{E}^2/c & \vec{B}^3 & 0 & -\vec{B}^1 \\ \vec{E}^3/c & -\vec{B}^2 & \vec{B}^1 & 0 \end{pmatrix} \tag{6.12}$$

Under a Lorentz transform the type-$(2,0)$ tensor $F^{\mu\nu}$ transforms as

$$F'^{\mu\nu} = \Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}F^{\rho\sigma} \implies [F'^{\mu\nu}] = [\Lambda^{\mu}{}_{\rho}][F^{\rho\sigma}][\Lambda^{\nu}{}_{\sigma}] \tag{6.13}$$

which gives

$$[F'^{\mu\nu}] = \begin{pmatrix} 0 & -\vec{E}^1/c & -\gamma(\vec{E}^2/c - \beta\vec{B}^3) & -\gamma(\vec{E}^3/c + \beta\vec{B}^2) \\ \vec{E}^1/c & 0 & -\gamma(\vec{B}^3 - \beta\vec{E}^2/c) & \gamma(\vec{B}^2 + \beta\vec{E}^3/c) \\ \gamma(\vec{E}^2/c - \beta\vec{B}^3) & \gamma(\vec{B}^3 - \beta\vec{E}^2/c) & 0 & -\vec{B}^1 \\ \gamma(\vec{E}^3/c + \beta\vec{B}^2) & -\gamma(\vec{B}^2 + \beta\vec{E}^3/c) & \vec{B}^1 & 0 \end{pmatrix} \tag{6.14}$$

and in the form of a Faraday tensor

$$\vec{E}' = \begin{pmatrix} \vec{E}^1 \\ \gamma(\vec{E}^2 - v\vec{B}^3) \\ \gamma(\vec{E}^3 + v\vec{B}^2) \end{pmatrix}, \quad \vec{B}' = \begin{pmatrix} \vec{B}^1 \\ \gamma(\vec{B}^2 + v\vec{E}^3/c^2) \\ \gamma(\vec{B}^3 - v\vec{E}^2/c^2) \end{pmatrix} \tag{6.15}$$

6.2 Maxwell's equations and the field-strength tensor

Current 4-vector

A static charge density in some inertial frame S will transform to a *moving* charge density (i.e. possess a current) in a different inertial frame. Consider a *current distribution* $\vec{J}$ formed from a charge density ρ moving with 3-velocity $(v,0,0)$ in S, so that $\vec{J} = \rho(v,0,0)$. Let S' (in standard configuration with S) be the rest frame of the charges. In S', the charge density is ρ_0 and $\vec{J}' = 0$.

Consider a given volume V'. In S the same amount of charges occupy a smaller volume V'/γ by length contraction. So

$$\rho = \gamma\rho_0 \quad \text{and} \quad \vec{J} = \gamma\rho_0(v,0,0) \tag{6.16}$$

We can define a current 4-vector

$$\boxed{j^{\mu} = (c\rho, \vec{J})} \tag{6.17}$$

which would allow us to recover Eq. 6.16, as $j'^{\mu} = (c\rho_0, \vec{0})$, and so

$$j^0 = \gamma(j'^0 + \beta j'^1) \implies \rho = \gamma\rho_0$$

$$\text{and} \quad j^1 = \gamma(j'^1 + \beta j'^0) \implies \vec{J}^1 = \gamma\beta c\rho_0 = \gamma v\rho_0$$

Relativistic field equations

We expect the relation between $\mathbf{F}$ and $\mathbf{j}$ to be linear in spacetime derivatives. Forming the covariant divergence of the type-$(2,0)$ tensor $F^{\mu\nu}$ results in a 4-vector. We consider, for some constant scalar k, a (source-ful) tensor equation of the form

$$\nabla_\mu F^{\mu\nu} = kj^\nu \tag{6.18}$$

In Cartesian inertial coordinates this transforms to $\partial_\mu F^{\mu\nu} = kj^\nu$. Taking the divergence and using the antisymmetry of $F^{\mu\nu}$ gives (why?)

$$k\partial_\nu j^\nu = \partial_\nu\partial_\mu F^{\mu\nu} = 0 \implies \partial_\mu j^\mu = 0, \qquad \boxed{\nabla_\mu j^\mu = 0 \text{ in general coordinates}} \tag{6.19}$$

i.e. we recover the continuity equation (charge conservation)

$$\frac{\partial j^0}{\partial(ct)} + \sum_i \frac{\partial j^i}{\partial x^i} = 0 \implies \frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\vec{J} = 0 \tag{6.20}$$

Considering each component in turn $\nu = (0,1,2,3)$, using Eq. 6.12

$$\frac{\partial F^{00}}{\partial(ct)} + \sum_i \frac{\partial F^{i0}}{\partial x^i} = kj^0 \implies \vec{\nabla}\cdot\vec{E} = kc^2\rho = \frac{k}{\epsilon_0\mu_0}\rho$$

and so we recover M1 (Eq. 6.1) if $k = \mu_0$. For e.g. $\nu = 1$ we have

$$\frac{\partial F^{01}}{\partial(ct)} + \sum_i \frac{\partial F^{i1}}{\partial x^1} = kj^1 \implies -\frac{1}{c^2}\partial_t\vec{E}^1 + \partial_2\vec{B}^3 - \partial_3\vec{B}^2 = \mu_0\vec{J}^1$$

Repeating for other indices we recover M4 (Eq. 6.4).

For the two source-free equations we introduce a further homogeneous tensor equation with four independent components. A totally antisymmetric type-$(0,3)$ tensor in a 4-dim space works because it has ${}_4\mathrm{C}_3 = 4$ independent components as required. So we consider

$$\nabla_{[\mu}F_{\nu\rho]} = 0 \tag{6.21}$$

$F_{\mu\nu}$ is antisymmetric so this can be written explicitly (in Cartesian inertial coordinates) as

$$\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} = 0$$

Considering each set of (μ,ν,ρ) in turn, for $(\mu,\nu,\rho) = (0,1,2)$

$$\partial_0 F_{12} + \partial_1 F_{20} + \partial_2 F_{01} = -\frac{1}{c}\partial_t\vec{B}^3 - \frac{1}{c}(\partial_1\vec{E}^2 - \partial_2\vec{E}^1) \implies \partial_t\vec{B}^3 + (\vec{\nabla}\times\vec{E})^3 = 0$$

with $(0,1,3)$ and $(0,2,3)$ we recover M2 (Eq. 6.2). For the remaining case $(\mu,\nu,\rho) = (1,2,3)$ we recover M3 (Eq. 6.3) $\vec{\nabla}\cdot\vec{B} = 0$. *Summary:* the four Maxwell's equations in any inertial frame are the components in Cartesian inertial coordinates of the two tensor equations

$$\nabla_\mu F^{\mu\nu} = \mu_0 j^\nu \quad \text{sourced} \tag{6.22}$$

$$\nabla_\mu F_{\nu\rho} + \nabla_\nu F_{\rho\mu} + \nabla_\rho F_{\mu\nu} = 0 \quad \text{sourceless} \tag{6.23}$$

4-vector potential

Consider a (dual-)vector potential A_μ. If we take

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{6.24}$$

then $\partial_{[\mu} F_{\nu\rho]} = 0$ (Eq. 6.23) is identically satisfied. Adding a gradient of a scalar $\partial_\nu \psi$ to A_μ leaves $F_{\mu\nu}$ unaltered; there is so-called residual *gauge freedom*. We choose the Lorentz-invariant Lorenz gauge, which takes A_μ to be divergence-free

$$\partial_\mu A^\mu = 0 \tag{6.25}$$

Note that Eq. 6.24 is equivalent to the more general form $F_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu$ due to the symmetry of the metric connection (see Eq. 4.41).

In Cartesian inertial coordinates, lower the index on the sourced 4-dim Maxwell equation (Eq. 6.22)

$$\nabla^\mu F_{\mu\nu} = \eta^{\mu\rho} \partial_\rho (\partial_\mu A_\nu - \partial_\nu A_\mu) = \mu_0 j_\nu$$

Applying the Lorenz gauge, the second term dies so this simplifies to

$$\nabla^2 A_\nu = \eta^{\mu\rho} \partial_\mu \partial_\rho A_\nu = \mu_0 j_\nu \tag{6.26}$$

where the Laplacian in Minkowski spacetime is the wave operator (wavespeed c)

$$\nabla^2 = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 \tag{6.27}$$

In the absence of charges and currents Eq. 6.26 admits wavelike solutions that travel at c. If variable source instead it describes the generation of EM fields that, asymptotically far from the source, describe radiation fields. The components of A^μ are

$$A^\mu = \left(\frac{\phi}{c}, \vec{A} \right) \tag{6.28}$$

Invoking Eq. 6.24 and Eq. 6.11, for $i = (1, 2, 3)$

$$F_{0i} = \frac{\vec{E}^i}{c} = -\frac{1}{c} \partial_t \vec{A}^i - \frac{1}{c} \partial_i \phi \implies \vec{E} = -\frac{\partial \vec{A}}{\partial t} - \vec{\nabla} \phi \tag{6.29}$$

Note the minus sign from lowering the index of A. The spatial components of $F_{\mu\nu}$ give, for example

$$F_{12} = -\vec{B}^3 = -\partial_1 \vec{A}^2 + \partial_2 \vec{A}^1 \implies \vec{B} = \vec{\nabla} \times \vec{A} \tag{6.30}$$

7 Spacetime curvature

7.1 Gravity as spacetime curvature

In curved spacetime, the equivalence principle tells us that the EoM of a massive particle is $\mathrm{D}u^\mu/\mathrm{D}\tau = 0$ as this reduces to the special-relativistic form $\mathrm{d}^2X^\mu/\mathrm{d}\tau^2 = 0$ in local-inertial coordinates, i.e. the worldline of a free-falling (under gravity) particle is a geodesic.

Local-inertial coordinates

Correspond to a local free-falling, non-rotating Cartesian reference frame. They are not unique at P; instead there are infinitely many related by Lorentz transformations. At P the metric connection vanishes and the basis vectors $\mathbf{e}_\mu \equiv \partial/\partial X^\mu$ are orthonormal. Close to any point P

$$g_{\mu\nu}(P) = \eta_{\mu\nu} + \frac{1}{2}\left(\frac{\partial^2 g_{\mu\nu}}{\partial X^\rho \partial X^\sigma}\right)_P [X^\rho - X^\rho(P)]\,[X^\sigma - X^\sigma(P)] \quad \text{and} \quad (\partial_\rho g_{\mu\nu})_P = 0 \tag{7.1}$$

to second order; the size of the second derivative determines the extent of the 'local' region.

Fermi-normal coordinates

Constructed c.f. the orthonormal tetrad (Eq. 5.38 and corresponding subsection). The evolution of the spacelike vectors along the worldline reflects motion of the local laboratory. At each τ the observer constructs a family of spacelike geodesics $\{\mathcal{C}\}$ that have unit tangent vectors at the observer constructed as linear combinations of the spacelike $\mathbf{e}_i(\tau)$. Any point within this 'local' region lies on exactly one such $\mathcal{C}$. The coordinates are $T = \tau$, $X^i =$ projection of the unit tangent vector *at the observer* onto the proper distance from the observer to the point of interest. Quantities measured in this laboratory correspond to projections of relevant 4-tensors onto this orthonormal frame.

Newtonian limit for a free-falling particle; weak-field approximation

In the absence of gravity, spacetime is Minkowski space. In the presence of a weak gravitational field, *globally* the metric approximates Minkowski:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \quad \text{with} \quad |h_{\mu\nu}| \ll 1 \quad \text{and} \quad \partial_0 h_{\mu\nu} = 0 \tag{7.2}$$

i.e. the metric is stationary in these global coordinates. For slow-moving particles, chain rule gives

$$\left|\frac{\mathrm{d}x^i}{\mathrm{d}t}\right| \ll c \quad \text{with} \quad t = \frac{x^0}{c} \implies \left|\frac{\mathrm{d}x^i}{\mathrm{d}\tau}\right| \ll \frac{\mathrm{d}x^0}{\mathrm{d}\tau} \tag{7.3}$$

In the geodesic equation we can ignore the $\mathrm{d}x^i/\mathrm{d}\tau$ terms relative to the $\mathrm{d}x^0/\mathrm{d}\tau$ term so that

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu_{00}c^2\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 \approx 0 \tag{7.4}$$

The relevant connection coefficients (Eq. 4.36) to first order in $h_{\mu\nu}$ are

$$\Gamma^\mu_{00} = \frac{1}{2}g_{\mu\nu}\left(2\frac{\partial g_{\nu 0}}{\partial x^0} - \frac{\partial g_{00}}{\partial x^\nu}\right) \approx -\frac{1}{2}\eta^{\mu\nu}\frac{\partial g_{00}}{\partial x^\nu} \approx -\frac{1}{2}\sum_i \eta^{\mu i}\frac{\partial h_{00}}{\partial x^i} \tag{7.5}$$

It follows that

$$\Gamma^0_{00} \approx 0 \quad \text{and} \quad \Gamma^i_{00} \approx \frac{1}{2}\frac{\partial h_{00}}{\partial x^i} \tag{7.6}$$

the first because the only non-zero Minkowski metric component is when $i = 0$ but we claimed in Eq. 7.3 that $\partial_0 h_{\mu\nu} = 0$. Then the geodesic equation becomes, for each component

$$\frac{\mathrm{d}^2 x^0}{\mathrm{d}\tau^2} \approx 0 \implies \boxed{\frac{\mathrm{d}^2 t}{\mathrm{d}\tau^2} \approx 0} \quad \text{and} \quad \frac{\mathrm{d}^2 x^i}{\mathrm{d}\tau^2} \approx -\frac{c^2}{2}\frac{\partial h_{00}}{\partial x^i}\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 \implies \boxed{\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} \approx -\frac{c^2}{2}\frac{\partial h_{00}}{\partial x^i}} \tag{7.7}$$

The latter has the form of the Newtonian EoM

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} = -\frac{\partial \Phi}{\partial x^i} \implies h_{00} \approx \frac{2\Phi}{c^2} \quad \text{or} \quad g_{00} \approx \left(1 + \frac{2\Phi}{c^2}\right) \tag{7.8}$$

So the small perturbation holds provided that $\Phi/c^2 \ll 1$.

7.2 Intrinsic curvature of a manifold

Riemann curvature tensor

A manifold (or some extended region of one) is flat if there exist Cartesian coordinates X^a such that the line element takes the pseudo-Euclidean form. If a region of spacetime is flat, then we can construct global inertial coordinates (metric $\eta_{\mu\nu}$) and we recover SR globally. To test for curvature in arbitrary coordinate systems without knowing the coordinate transform to Cartesian, we construct the Riemann curvature tensor.

Consider a dual-vector field v_a

$$\begin{aligned}\nabla_a \nabla_b v_c &= \partial_a(\nabla_b v_c) - \Gamma^d_{ab}\nabla_d v_c - \Gamma^d_{ac}\nabla_b v_d \\ &= \partial_a(\partial_b v^c - \Gamma^d_{bc} v_d) - \Gamma^d_{ab}(\partial_d v^c - \Gamma^e_{dc} v_e) - \Gamma^d_{ac}(\partial_b v_d - \Gamma^e_{bd} v_e)\end{aligned} \tag{7.9}$$

Then we define the Riemann curvature tensor $R_{abc}{}^d$; it is a type-$(1,3)$ tensor (from the quotient theorem)

$$\boxed{(\nabla_a \nabla_b - \nabla_b \nabla_a)\, v_c = \left(-\partial_a \Gamma^d_{bc} + \partial_b \Gamma^d_{ac} + \Gamma^e_{ac}\Gamma^d_{be} - \Gamma^e_{bc}\Gamma^d_{ae}\right) v_d = R_{abc}{}^d v_d} \tag{7.10}$$

$R_{abc}{}^d$ vanishes for a flat manifold as we can always choose Cartesian coordinates for which the connection vanishes; the converse is also true. It is antisymmetric in its first two indices and has cyclic symmetry

$$R_{abc}{}^d = -R_{bac}{}^d \quad \text{and} \quad R_{abc}{}^d + R_{cab}{}^d + R_{bca}{}^d = 0 \implies R_{[abc]}{}^d = 0 \tag{7.11}$$

Note that symmetries are preserved under general coordinate transformations. Any symmetry of, say, the components R_{abcd} taken at an arbitrary point implies a general symmetry of the Riemann tensor.

Lowering to form the type-$(0,4)$ tensor R_{abcd}; if we take this at P in local Cartesian coordinates on a general curved manifold, where the connection vanishes,

$$(R_{abcd})_P = (g_{de} R_{abc}{}^d)_P = -(g_{de}\partial_a \Gamma^e_{bc} - g_{de}\partial_b \Gamma^e_{ac})_P \tag{7.12}$$

where, per Eq. 4.36,

$$(g_{de}\partial_a \Gamma^e_{bc})_P = \frac{1}{2}(\partial_a\partial_b g_{cd} + \partial_a\partial_c g_{bd} - \partial_a\partial_d g_{bc})_P$$

and so

$$(R_{abcd})_P = \frac{1}{2}(\partial_a\partial_d g_{bc} + \partial_b\partial_c g_{ad} - \partial_a\partial_c g_{bd} - \partial_b\partial_d g_{ac})_P \tag{7.13}$$

We note two further symmetries: antisymmetric on the third and fourth indices and symmetric under swapping pairs of indices

$$R_{abcd} = -R_{abdc} \quad \text{and} \quad R_{abcd} = R_{cdab} \tag{7.14}$$

To summarise the antisymmetries

$$R_{abcd} = R_{[ab]cd} = R_{ab[cd]} \tag{7.15}$$

What do these symmetries mean for the number of independent components of $\mathbf{R}$?

- 1-dim: $\mathbf{R}$ necessarily vanishes as R_{1111} (the one possible component) vanishes by antisymmetry;
- 2-dim: only one independent component, R_{1212};
- 3-dim: *six* independent components;
- 4-dim: 20 independent components; there are 21 from considering the [anti]symmetries above, and we lose one from invoking cyclic symmetry.

Bianchi identity

The Riemann curvature tensor satisfies the differential Bianchi identity (a tensor identity).

$$\nabla_a R_{bcd}{}^e + \nabla_b R_{cad}{}^e + \nabla_c R_{abd}{}^e = 0 \quad \text{or equivalently} \quad \boxed{\nabla_{[a} R_{bc]d}{}^e} = 0 \tag{7.16}$$

Proof. We show that the Bianchi identity holds in one coordinate system. Then it will hold in all. At an arbitrary point P in local Cartesian coordinates (the connection vanishes), using Eq. 7.10

$$(\nabla_a R_{bcd}{}^e)_P = (-\partial_a \partial_b \Gamma^e_{cd} + \partial_a \partial_c \Gamma^e_{bd})_P \tag{7.17}$$

Summing through the cyclic permutations causes the RHS to vanish.

Ricci tensor and Ricci scalar

Contracting once gives the Ricci tensor R_{ab}; contracting once again gives the Ricci scalar R. For flat manifolds R_{ab} and R vanish as the curvature tensor R_{abcd} vanishes.

Given Eq. 7.15 we contract the first and last indices to form the *symmetric* Ricci tensor

$$R_{ab} \equiv R_{cab}{}^c = -\partial_c \Gamma^c_{ab} + \partial_a \Gamma^c_{cb} + \Gamma^d_{cb} \Gamma^c_{ad} - \Gamma^d_{ab} \Gamma^c_{cd} \tag{7.18}$$

Symmetry follows from contracting the cyclic identity (Eq. 7.11).

Contracting once again gives the Ricci scalar

$$R \equiv g^{ab} R_{ab} \tag{7.19}$$

It is possible for R_{ab} to vanish even if the full curvature tensor is non-zero. Contracting the Bianchi identity once

$$0 = \delta^d_e (\nabla_a R_{bcd}{}^e + \nabla_c R_{abd}{}^e + \nabla_b R_{cad}{}^e) = \nabla_a R_{cd} - \nabla_c R_{ad} + \nabla_b R_{cadb} \tag{7.20}$$

Contracting once more over a and d

So we obtain the *contracted Bianchi identity*

$$\nabla^a (G_{ab}) = \nabla^a \left(R_{ab} - \frac{1}{2} g_{ab} R \right) = 0 \tag{7.21}$$

G_{ab} is the symmetric and divergence-free Einstein tensor.

7.3 Parallel transport

We relate the curvature of the manifold to the path dependence of parallel transport by considering the parallel transport of a vector $\mathbf{v}$ around an infinitesimal loop defined by a curve $\mathcal{C}$, parametrised by $x^a(u)$. Let the start and endpoint be P.

Recall Eqs. 4.44 and 4.46 so that at some point along the curve $x^a(u)$, the result of parallel transporting from P is

$$v^a(u) = v^a_P - \int_{u_P}^{u} \Gamma^a_{bc} \frac{\mathrm{d}x^b}{\mathrm{d}u'} v^c \,\mathrm{d}u' = v^a_P - \int_{x^b(u_P)}^{x^b(u)} \Gamma^a_{bc} v^c \,\mathrm{d}x^b \tag{7.22}$$

The closed loop is small so we can expand about x^a_P. To first order

$$\Gamma^a_{bc}(u) = (\Gamma^a_{bc})_P + (\partial_d \Gamma^a_{bc})\left(x^d(u) - x^d_P\right) + \dots \tag{7.23}$$

$$v^c(u) = v^c_P - (\Gamma^c_{ef})_P\, v^f_P\, (x^e(u) - x^e_P) + \dots \tag{7.24}$$

The expansion of $v^c(u)$ is cyclical and follows from Eq. 7.22. In zeroth order taking the integral from u_P to u_P returns zero; in first order the derivative of v^a_P (a constant) is just zero; differentiating the integral w.r.t x^b returns just $-\Gamma^a_{bc}v^c$. Keeping terms to first order and noting that the connection at P is zero,

$$v^a(u) - v^a_P = -\int_{x^b(u_P)}^{x^b(u)} (\partial_d \Gamma^a_{bc} - \Gamma^a_{be}\Gamma^e_{dc})_P\, v^c_P (x^d(u) - x^d_P)\,\mathrm{d}x^b$$

$$\implies \Delta v^a = -(\partial_d \Gamma^a_{bc} - \Gamma^a_{be}\Gamma^e_{dc})_P\, v^c_P \oint x^d \,\mathrm{d}x^b$$

Noting that $\oint \mathrm{d}(x^b x^d) = 0 \implies \oint x^b \,\mathrm{d}x^d = -\oint x^d \,\mathrm{d}x^b$, and antisymmetrising over the indices b and d

$$\partial_d \Gamma^a_{bc} - \Gamma^a_{be}\Gamma^e_{dc} = \frac{1}{2}(\partial_d \Gamma^a_{bc} - \partial_b \Gamma^a_{dc} - \Gamma^a_{be}\Gamma^e_{dc} + \Gamma^a_{de}\Gamma^a_{bc}) = -\frac{1}{2} R_{dbc}{}^a \tag{7.25}$$

$$\boxed{\Delta v^a = (R_{bcd}{}^a)_P\, v^d_P \oint x^{[b}\,\mathrm{d}x^{c]}} \tag{7.26}$$

This is zero if the Riemann curvature tensor is zero. For an infinitesimal loop the contour integral is also a type-$(2,0)$ tensor at P containing the planar area of the loop.

7.4 Geodesic deviation

Depending on the curvature, two initially parallel geodesics will converge or diverge; let these be $\mathcal{C}$ and $\bar{\mathcal{C}}$ affinely-parametrised by $x^a(u)$ and $\bar{x}^a(u)$ respectively. We choose initial u so that the coordinate difference $\xi^a(u) \equiv \bar{x}^a(u) - x^a(u)$ is infinitesimal; if so then the ξ^a form the components of a type-$(1,0)$ tensor, called the connecting vector. Initially parallel implies $\mathrm{D}\xi^a/\mathrm{D}u = 0$. The geodesic equation gives, for each of $\mathcal{C}$ and $\bar{\mathcal{C}}$ (not shown), with the derivatives evaluated at their respective points

$$\ddot{x}^a + \Gamma^a_{bc}\dot{x}^b\dot{x}^c = 0 \tag{7.27}$$

Taylor expanding to first order in ξ^a,

$$\bar{\Gamma}^a_{bc} = \Gamma^a_{bc} + \partial_d \Gamma^a_{bc}\xi^d \tag{7.28}$$

Taking the difference of Eq. 7.27

$$\frac{\mathrm{d}^2\xi^a}{\mathrm{d}u^2} + \bar{\Gamma}^a_{bc}\dot{\bar{x}}^b\dot{\bar{x}}^c - \Gamma^a_{bc}\dot{x}^b\dot{x}^c = \frac{\mathrm{d}^2\xi^a}{\mathrm{d}u^2} + \Gamma^a_{bc}(\dot{\bar{x}}^b\dot{\bar{x}}^c - \dot{x}^b\dot{x}^c) + \partial_d\Gamma^a_{bc}\dot{x}^b\dot{x}^c\xi^d = 0$$

where using $\dot{\bar{x}}^a = \dot{\xi}^a + \dot{x}^a \implies \dot{\bar{x}}^b\dot{\bar{x}}^c = \dot{\xi}^b\dot{x}^c + \dot{\xi}^c\dot{x}^b + \dot{x}^b\dot{x}^c$ to first order in ξ^a. The bracketed term becomes $\Gamma^a_{bc}(\dot{\xi}^b\dot{x}^c + \dot{\xi}^c\dot{x}^b) = 2\Gamma^a_{bc}\dot{x}^b\dot{\xi}^c$ by symmetry in the lower indices b and c. So

$$\frac{\mathrm{d}^2\xi^a}{\mathrm{d}u^2} + 2\Gamma^a_{bc}\dot{x}^b\dot{\xi}^c + \partial_d\Gamma^a_{bc}\dot{x}^b\dot{x}^c\xi^d = 0 \tag{7.29}$$

Consider (why are we allowed to set $\dot{\xi}^b \equiv \dot{x}^b$?)

$$\begin{aligned}
\frac{\mathrm{D}}{\mathrm{D}u}\left(\frac{\mathrm{D}\xi^a}{\mathrm{D}u}\right) &= \frac{\mathrm{d}}{\mathrm{d}u}\left(\dot{\xi}^a + \Gamma^a_{bc}\dot{x}^b\xi^c\right) + \Gamma^a_{bc}\dot{x}^b\left(\dot{\xi}^c + \Gamma^c_{de}\dot{x}^d\xi^e\right) \\
&= \frac{\mathrm{d}^2\xi^a}{\mathrm{d}u^2} + \partial_d\Gamma^a_{bc}\dot{x}^b\dot{x}^d\xi^c + \Gamma^a_{bc}\ddot{x}^b\xi^c + \Gamma^a_{bc}\dot{x}^b\dot{\xi}^c + \Gamma^a_{bc}\dot{x}^b\dot{\xi}^c + \Gamma^a_{bc}\Gamma^c_{de}\dot{x}^b\dot{x}^d\xi^e \\
&= \frac{\mathrm{d}^2\xi^a}{\mathrm{d}u^2} + 2\Gamma^a_{bc}\dot{x}^b\dot{\xi}^c + \partial_d\Gamma^a_{bc}\dot{x}^b\dot{x}^d\xi^c - \Gamma^a_{bc}\Gamma^b_{de}\dot{x}^d\dot{x}^e\xi^c + \Gamma^a_{bc}\Gamma^c_{de}\dot{x}^b\dot{x}^d\xi^e
\end{aligned}$$

where $\ddot{x}^b$ is eliminated by the geodesic equation. The orange term is just $-\partial_d\Gamma^a_{bc}\dot{x}^b\dot{x}^c\xi^d$ by Eq. 7.29 and so, relabelling indices,

$$\frac{\mathrm{D}}{\mathrm{D}u}\left(\frac{\mathrm{D}\xi^a}{\mathrm{D}u}\right) - \left(-\partial_d\Gamma^a_{bc} + \partial_b\Gamma^a_{cd} + \Gamma^a_{be}\Gamma^e_{cd} - \Gamma^a_{de}\Gamma^e_{bc}\right)\dot{x}^b\dot{x}^c\xi^d = 0$$

This is the geodesic deviation equation which describes the evolution of ξ^a.

$$\implies \boxed{\frac{\mathrm{D}}{\mathrm{D}u}\left(\frac{\mathrm{D}\xi^a}{\mathrm{D}u}\right) - R_{dbc}{}^a\dot{x}^b\dot{x}^c\xi^d = 0} \tag{7.30}$$

In Minkowski spacetime Eq. 7.30 describes the relative acceleration of neighbouring free-falling particles due to tidal gravitational effects. Consider free-falling particles with geodesics $x^\mu(\tau)$ and $\bar{x}^\mu(\tau)$. The connecting vector ξ^μ evolves according to

$$\frac{\mathrm{D}}{\mathrm{D}\tau}\left(\frac{\mathrm{D}\xi^a}{\mathrm{D}\tau}\right) = R_{\nu\alpha\beta}{}^\mu u^\alpha u^\beta\xi^\nu = S_\nu{}^\mu\xi^\nu \tag{7.31}$$

u^μ is the 4-velocity and $S_\nu{}^\mu$ is the symmetric tidal tensor. This is analogous to tidal acceleration in Newtonian gravity: let the two free-falling particles have trajectories $x^i(t)$ and $\bar{x}^i(t)$ with connecting vector $\xi^i(t)$. Each responds to a gravitational potential Φ as $\ddot{x}^i = -(\partial_i\Phi)_{x(t)}$. Taking the difference and expanding to first order in ξ^i

$$\frac{\mathrm{d}^2\xi^i}{\mathrm{d}t^2} \approx -\left(\frac{\partial^2\Phi}{\partial x^i\partial x^j}\right)\xi^j \tag{7.32}$$

In free space $\vec{\nabla}^2\Phi = 0$ so the tidal tensor $-\partial_i\partial_j\Phi$ is symmetric and traceless. It generates a volume-preserving distortion of a set of free-falling particles (what?) In the weak-field limit Eqs. 7.32 and 7.30 are equivalent.

8 Gravitational field equations

8.1 Energy-momentum tensor

c.f. Poisson equation relates the Laplacian of the gravitational potential to the mass density in Newtonian gravity. We introduce the type-$(2,0)$ *energy-momentum tensor* $T^{\mu\nu}$ which acts as the source of spacetime curvature and is *always symmetric*.

Dust, densities, and fluxes

Consider the case of dust: non-interacting particles each of rest mass m with *no* velocity dispersion. At each event P, all particles present there have the same 4-velocity $u^\mu(x)$. Measured in some local inertial frame S, all particles have 3-velocity $\vec{u}$ and the dust has energy density ρc^2. It is possible to find a local inertial frame in which the particles are at rest - the IRF. In this IRF the number density of particles is n_0 so the energy density is the *scalar field* $\rho_0 c^2 = n_0 m c^2$.

Transforming to S, the number density is $\gamma_u n_0$ and the energy of each particle is $\gamma_u m c^2$. Then the energy density is

$$\rho c^2 = \gamma_u^2 n_0 m c^2 = \gamma_u^2 \rho_0 c^2 \tag{8.1}$$

i.e. it transforms like the 00 component of the type-$(2,0)$ tensor (per Eq. 5.12, $u^\mu = \gamma_u(c, \vec{u})$)

$$T^{\mu\nu}(x) = \rho_0(x) u^\mu(x) u^\nu(x) \implies T^{00} = (\gamma_u c)^2 \rho_0 \tag{8.2}$$

The T^{i0} are the 3-momentum density (up to a factor of c; number density $\times$ 3-momentum) or alternatively the energy flux in the ith direction, as, rearranging (energy density $\times$ 3-velocity)

$$T^{i0} = \rho_0(\gamma_u \vec{u}^i)(\gamma_u c) = c(\gamma_u n_0)(m\gamma_u \vec{u}^i); \quad \text{energy flux} = (\gamma_u^2 n_0 m c^2)\vec{u}^i = cT^{i0} \tag{8.3}$$

The T^{ij} with $j = (1,2,3)$ are the flux of the i-component of 3-momentum along the j-direction (i-component of 3-momentum *density* $\times$ j-component of 3-velocity)

$$T^{ij} = n_0 m(\gamma_u \vec{u}^i)(\gamma_u \vec{u}^j) = (\gamma_u^2 n_0 m \vec{u}^i)\vec{u}^j = \frac{1}{c} T^{i0} \vec{u}^j \tag{8.4}$$

This generalises to other sources (e.g. electromagnetic field). The symmetry $T^{\mu\nu} = T^{\nu\mu}$ conserves angular momentum in all inertial frames. The energy density T^{00} must include all sources of energy (e.g. kinetic energy from velocity dispersion, interaction energies); similarly, the energy flux/3-momentum density T^{i0} includes contributions from heat conduction (non-zero in IRF) and bulk motion; the momentum flux T^{ij} includes effects from velocity dispersion and shear stresses.

Ideal fluid; conservation of energy and momentum

Ideal fluids have the property that there is an IRF in which $T^{i0} = 0$ and $T^{ij} \propto \delta^{ij}$ (isotropic spatial components). As we require conduction and shear stresses to be negligible, the mean free path must be much smaller than the length scales of temperature/velocity gradients. In the IRF

$$T^{\mu\nu} = \mathrm{diag}(\rho c^2, p, p, p) = \boxed{\left(\rho + \frac{p}{c^2}\right) u^\mu u^\nu - p g^{\mu\nu}} \tag{8.5}$$

where p is the isotropic pressure and the subscript on ρ_0 has been dropped; the RHS is the *tensor form*, valid in any coordinate system, where u^μ is the fluid 4-velocity defined by its IRF. For $p \ll \rho c^2$, Eq. 8.5 reduces to that for dust. (Note that the energy density ρc^2 and isotropic pressure p are scalar fields as they are defined in the IRF.)

The energy-momentum tensor satisfies a covariant continuity equation (c.f. Eq. 6.19 in EM)

$$\boxed{\nabla_\mu T^{\mu\nu} = 0} \tag{8.6}$$

which reduces to, in local-inertial coordinates at some event P (exactly like in EM), and from which conservation of energy and momentum follow respectively

$$\frac{\partial T^{00}}{\partial t} + \sum_i \frac{\partial(cT^{i0})}{\partial x^i} = 0 \implies \frac{\partial}{\partial t}(\text{energy density}) + \vec{\nabla}\cdot(\text{energy flux}) = 0 \tag{8.7}$$

$$\frac{\partial(T^{i0}/c)}{\partial t} + \sum_j \frac{\partial T^{ij}}{\partial x^j} = 0 \implies \frac{\partial}{\partial t}(\text{momentum density}) + \vec{\nabla}\cdot(\text{momentum flux}) = 0 \tag{8.8}$$

Example. Recovering the (Newtonian) Euler equation for an ideal fluid. Conservation of the energy-momentum tensor via Eq. 8.6 gives

$$\nabla_\mu \left[\left(\rho + \frac{p}{c^2}\right) u^\mu u^\nu - p g^{\mu\nu}\right] = 0$$

By chain rule and recalling $\nabla_\mu g^{\mu\nu} = 0$ by metric compatibility (Eq. 4.34)

$$\nabla_\mu \left(\rho + \frac{p}{c^2}\right) u^\mu u^\nu + \left(\rho + \frac{p}{c^2}\right)(\nabla_\mu u^\mu) u^\nu + \left(\rho + \frac{p}{c^2}\right) u^\mu(\nabla_\mu u^\nu) - (\nabla_\mu p) g^{\mu\nu} = 0 \tag{8.9}$$

Contracting Eq. 8.9 with u_ν extracts the component parallel to u^ν; the red term is 0 as $u^\mu \nabla_\mu u^\nu$ is orthogonal to u^ν

$$c^2 \nabla_\mu \left(\rho + \frac{p}{c^2}\right) u^\mu + c^2 \left(\rho + \frac{p}{c^2}\right)(\nabla_\mu u^\mu) + \left(\rho + \frac{p}{c^2}\right) u^\nu u_\mu \nabla_\mu u^\nu - (\nabla_\mu p) u^\mu = 0$$

$$\implies \nabla_\mu(\rho u^\mu) + \frac{p}{c^2}\nabla_\mu u^\mu = 0 \tag{8.10}$$

This is the relativistic continuity equation (energy conservation). To extract the perpendicular component we subtract from Eq. 8.9 the parallel component, which gives momentum conservation

$$\left(\rho + \frac{p}{c^2}\right) u^\mu \nabla_\mu u^\nu = \left(g^{\mu\nu} - \frac{u^\mu u^\nu}{c^2}\right)\nabla_\mu p \tag{8.11}$$

To see how this reduces to Newtonian fluid mechanics for slowly moving fluids and small pressures, we adopt local inertial coordinates at some event P, at which the metric connection vanishes and $g^{\mu\nu} = \eta^{\mu\nu}$. Consider the 4-velocity $u^\mu = \frac{\mathrm{d}t}{\mathrm{d}\tau}(c, \vec{u})$ in the Newtonian limit $|\vec{u}^i| \ll c$ so that $\mathrm{d}t/\mathrm{d}\tau \approx 1$ and that $p \ll \rho c^2$. Then $u^\mu \approx (c, \vec{u}^i)$ and Eq. 8.10 reduces to, at P, $\partial_\mu(\rho u^\mu) \approx 0$

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x^i}(\rho \vec{u}^i) \approx 0 \tag{8.12}$$

Eq. 8.11 reduces to

$$\rho u^\mu \partial_\mu u^\nu = \left(\eta^{\mu\nu} - \frac{u^\mu u^\nu}{c^2}\right)\partial_\mu p \tag{8.13}$$

which vanishes on both sides for $\nu = 0$. For $\nu = i = (1, 2, 3)$ we reduce this further to

$$\rho u^\mu \partial_\mu u^\nu \approx \eta^{\mu\nu}\partial_\mu p = -\delta^{ji}\partial_j p \implies \rho\left(\frac{\partial}{\partial t} + \vec{u}\cdot\vec{\nabla}\right)\vec{u} = -\vec{\nabla} p \tag{8.14}$$

which is the Euler equation for an ideal fluid.

8.2 Einstein field equations

Recall that in the weak-field limit we recover the correct Newtonian EoM from the geodesic equation if Eq. 7.8 holds true. In the last section we showed that for a slow-moving fluid, in local inertial coordinates, $T_{00} \approx \rho c^2$. Combining these with Poisson's equation, on the LHS we have $\vec{\nabla}^2\Phi \approx c^2\vec{\nabla}^2 g_{00}/2$ and on the RHS, $4\pi G\rho \approx 4\pi G T_{00}/c^2$ so

$$\vec{\nabla}^2 g_{00} \approx \frac{8\pi G}{c^4} T_{00} \tag{8.15}$$

Via the Riemann curvature tensor (Eq. 7.10), the second derivative of the metric is a measure of curvature: this suggests we look for a relativistic field equation of the form

$$K_{\mu\nu} = \kappa T_{\mu\nu} \quad \text{with} \quad \kappa = \frac{8\pi G}{c^4} \tag{8.16}$$

$K_{\mu\nu}$ is a symmetric type-$(0, 2)$ tensor related to the curvature of spacetime, so it must be constructed from $R_{\mu\nu\sigma\rho}$ and $g_{\mu\nu}$. As we require $K_{\mu\nu}$ to be linear in the second derivatives of $g_{\mu\nu}$ the most general form of $K_{\mu\nu}$ must have

$$K_{\mu\nu} = aR_{\mu\nu} + bRg_{\mu\nu}$$

(where $R_{\mu\nu}$ is the Ricci tensor and R is the Ricci scalar). It follows from Eqs. 8.6 and 8.16 that $K_{\mu\nu}$ must also satisfy the covariant continuity equation

$$\nabla_\mu K^{\mu\nu} = \nabla_\mu(aR^{\mu\nu} + bRg^{\mu\nu}) = 0$$

The most obvious candidate for $K_{\mu\nu}$ is the contracted Bianchi identity (Eq. 7.21)

$$\nabla^\mu \left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R\right) = 0$$

so we define the symmetric Einstein tensor, $G_{\mu\nu} \equiv R_{\mu\nu} - g_{\mu\nu}R/2$, which satisfies $\nabla_\mu G^{\mu\nu} = 0$. For consistency with the weak-field limit (Newtonian) we require $a = -1$ (see Section 8.3). These are the Einstein field equations (EFEs); $G_{\mu\nu}$ and $T_{\mu\nu}$ have 10 independent components so there are 10 non-linear PDEs for the $g_{\mu\nu}$.

$$\boxed{G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = -\kappa T_{\mu\nu}} \tag{8.17}$$

Expressing Eq. 8.17 in terms of mixed components then contracting by setting $\mu = \nu$ (equivalently just contract with $g^{\mu\nu}$)

$$R^\mu_\nu - \frac{1}{2}\delta^\mu_\nu R = -\kappa T^\mu_\nu \implies R = \kappa T$$

where $T \equiv T^\mu_\mu$ is the trace of the energy-momentum tensor; we obtain an alternative form of Eq. 8.17

$$\boxed{R_{\mu\nu} = -\kappa\left(T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T\right)} \tag{8.18}$$

A region of spacetime is *empty* (i.e. a vacuum) if $T_{\mu\nu} = 0$ there; such a region is not only devoid of matter but also of radiative energy and momentum. It follows from Eq. 8.18 that the gravitational field equations for empty space are

$$R_{\mu\nu} = 0 \tag{8.19}$$

To be fixed

8.3 Weak-field limit

Consider a slowly-moving non-relativistic fluid ($p \ll \rho c^2$) and assume a stationary situation (i.e. the mass distribution and metric are independent of time in these coordinates). Then Eq. 8.5 reduces to

$$T_{\mu\nu} \approx \rho u_\mu u_\nu \implies T \approx g^{\mu\nu} \rho u_\mu u_\nu = \rho c^2$$

Also $u^\mu \approx (c, \vec{u}^i)$ so with $u_0 = g_{0\mu} u^\mu \approx g_{00} c \approx c$ for $|\vec{u}^i| \ll c$ and $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$,

$$T_{00} \approx \rho u_0 u_0 \approx \rho c^2$$

Then from Eq. 8.18

$$R_{00} = -\kappa \left(T_{00} - \frac{1}{2} g_{00} T \right) \approx -\frac{1}{2} \kappa \rho c^2 \tag{8.20}$$

From the definition of the Riemann curvature tensor we find

$$R_{00} = -\partial_\mu \Gamma^\mu_{00} + \partial_0 \Gamma^\mu_{\mu 0} + \Gamma^\nu_{\mu 0} \Gamma^\mu_{0\nu} - \Gamma^\nu_{00} \Gamma^\mu_{\mu\nu} \approx -\partial_i \Gamma^i_{00} \tag{8.21}$$

where the connection coefficients are first-order in $h_{\mu\nu}$ so we can ignore the terms in red. We have implicitly summed over i. From Eq. 7.6

$$\Gamma^i_{00} \approx \frac{1}{2} \frac{\partial h_{00}}{\partial x^i} \implies R_{00} \approx -\frac{1}{2} \vec{\nabla}^2 h_{00} \implies \vec{\nabla}^2 h_{00} \approx \kappa \rho c^2 = \frac{8\pi G}{c^2} \rho$$

As $h_{00} \approx 2\Phi / c^2$, we recover Poisson's equation. This justifies the choice of a in Eq. 8.17.

8.4 Cosmological constant

In the previous section we imposed the requirement that $K_{\mu\nu}$ must be linear in the second derivatives of $g_{\mu\nu}$. If we relax this condition then we find that as $\nabla_\mu g^{\mu\nu} = 0$ we can add any constant multiple of $g_{\mu\nu}$ and still obtain a consistent set of EFEs[1] - the covariant continuity equation is still satisfied. The constant multiple is known as the cosmological constant, Λ. Λ is non-zero but very small.

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = -\kappa T_{\mu\nu} \tag{8.22}$$

Contracting with $g^{\mu\nu}$ we find $R = \kappa T + 4\Lambda$ so that

$$R_{\mu\nu} = -\kappa \left(T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T \right) + \Lambda g_{\mu\nu} \tag{8.23}$$

In the weak-field limit $\vec{\nabla}^2 \Phi = 4\pi G \rho - \Lambda c^2$, which, for a point source of mass M, has solution

$$\vec{g} = -\vec{\nabla} \Phi = -\frac{GM}{|\vec{x}|^3} \vec{x} + \frac{\Lambda c^2}{3} \vec{x} \tag{8.24}$$

Λ provides a gravitational repulsion whose strength increases linearly with $|\vec{x}|$.

Zero-point energy

Recall the energy-momentum tensor for an ideal fluid (Eq. 8.5). Suppose a fluid has a large *negative* pressure $p = -\rho c^2$ so that

$$T_{\mu\nu} = -p g_{\mu\nu} = \rho c^2 g_{\mu\nu} \tag{8.25}$$

[1] Possibly see Lovelock's theorem.

In *any* local inertial frame, *all* observers measure energy density $T^{00} = \rho c^2$. This is a fundamental property of the vacuum with corresponding energy-momentum tensor $T_{\mu\nu}^{\text{vac}}$. We can write Eq. 8.22 as

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = -\kappa \left(T_{\mu\nu} + \frac{\Lambda}{\kappa} g_{\mu\nu} \right) = -\kappa \left(T_{\mu\nu} + T_{\mu\nu}^{\text{vac}} \right) \tag{8.26}$$

$$\text{with} \quad T_{\mu\nu}^{\text{vac}} = \rho_{\text{vac}} c^2 g_{\mu\nu} \quad \text{and} \quad \rho_{\text{vac}} c^2 = \frac{\Lambda}{\kappa} = \frac{\Lambda c^4}{8\pi G} \tag{8.27}$$

Is this expected? In QFT, the Fourier modes of free quantum fields behave like QHOs, each with a zero-point energy. Summing all of these gives an answer 120 orders of magnitude higher than observed cosmologically - this is still unsolved.

9 The Schwarzschild solution

Where symmetries are present in spacetime, some exact solutions to the EFEs are known. We now adopt a passive viewpoint, that we can change our coordinate system without changing the *functional form* of the fields on our coordinates. A spacetime possesses a symmetry if under some $x^\mu \to x'^\mu$, the new components of $g'_{\mu\nu}(x')$ have the same functional dependence as $g_{\mu\nu}(x)$, so that the new line element has the same dependence on x'^μ and $\mathrm{d}x'^\mu$ as the original line element does on x^μ and $\mathrm{d}x^\mu$.

Some definitions of symmetries:

- Stationary: if the spacetime is invariant under constant time shifts $t \to t + \mathrm{const.}$ (or symmetric under such a shift for all constant shifts); $g_{\mu\nu}$ is independent of time;
- Static: the spacetime is stationary but additionally is invariant under time reversal, $t \to -t$; this requires $g_{0i} = 0$.

The Schwarzschild solution represents the spacetime geometry outside a spherically-symmetric matter distribution.

9.1 General static isotropic line element

A good starting point is to construct the most general form of the metric for a spatially-isotropic spacetime, and impose the static constraint afterwards. The metric is isotropic if $\mathrm{d}s^2$ depends only on rotational invariants of the spacelike coordinates x^i and their differentials, which are limited to

$$\vec{x} \cdot \vec{x} = r^2, \quad \vec{x} \cdot \mathrm{d}\vec{x} = r\,\mathrm{d}r, \quad \mathrm{d}\vec{x} \cdot \mathrm{d}\vec{x} = \mathrm{d}r^2 + r^2\,\mathrm{d}\Omega^2$$

where $\mathrm{d}\Omega^2 = \mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2$ in spherical polars. Absorb factors of r into the functions so that the spatially-isotropic line element takes the form

$$\mathrm{d}s^2 = A(t,r)\,\mathrm{d}t^2 - B(t,r)\,\mathrm{d}t\,\mathrm{d}r - C(t,r)\,\mathrm{d}r^2 - D(t,r)\,\mathrm{d}\Omega^2$$

If we define a new radial coordinate by $\bar{r}^2 = D(t,r)$ so that the coordinates are $(t, \bar{r}, \theta, \phi)$, and redefine the arbitrary functions in terms of t and $\bar{r}$,

$$\mathrm{d}s^2 = A(t,\bar{r})\,\mathrm{d}t^2 - B(t,\bar{r})\,\mathrm{d}t\,\mathrm{d}\bar{r} - C(t,\bar{r})\,\mathrm{d}\bar{r}^2 - \bar{r}^2\,\mathrm{d}\Omega^2$$

To remove the cross-term $\mathrm{d}t\,\mathrm{d}\bar{r}$, define a new timelike coordinate $\bar{t}$ with

$$\mathrm{d}\bar{t} = \Phi(t,\bar{r})\,[A(t,\bar{r})\,\mathrm{d}t - B(t,\bar{r})\,\mathrm{d}\bar{r}] \tag{9.1}$$

where $\Phi(t,\bar{r})$ is an integrating factor which makes the RHS an exact differential. Squaring Eq. 9.1 we find

$$A\,\mathrm{d}t^2 - B\,\mathrm{d}t\,\mathrm{d}r = \frac{1}{A\Phi^2}\,\mathrm{d}\bar{t}^2 - \frac{B^2}{A}\,\mathrm{d}\bar{r}^2$$

$$\implies \mathrm{d}s^2 = \frac{1}{A\Phi^2}\,\mathrm{d}\bar{t}^2 - \left(C + \frac{B^2}{A}\right)\mathrm{d}\bar{r}^2 - \bar{r}^2\,\mathrm{d}\Omega^2$$

$$\boxed{\mathrm{d}s^2 = \bar{A}(\bar{t},\bar{r})\,\mathrm{d}\bar{t}^2 - \bar{B}(\bar{t},\bar{r})\,\mathrm{d}\bar{r}^2 - \bar{r}^2\,\mathrm{d}\Omega^2} \tag{9.2}$$

We recover a *diagonal* form for the isotropic line element. For surfaces given by $\bar{t}$, $\bar{r}$ constant, Eq. 9.2 describes the geometry of 2-spheres with area $4\pi\bar{r}^2$. So $\bar{r}$ is an area coordinate.

We can drop the bars now. Imposing the static constraint, the static isotropic line element takes the form, in the coordinate system (t, r, θ, ϕ)

$$\boxed{\mathrm{d}s^2 = A(r)\,\mathrm{d}t^2 - B(r)\,\mathrm{d}r^2 - r^2\,\mathrm{d}\Omega^2} \tag{9.3}$$

9.2 Solution of the empty-space EFEs

$A(r)$ and $B(r)$ are determined by solving the EFEs given some spherical and static matter distribution. For empty space the EFEs reduce to Eq. 8.19

$$R_{\mu\nu} = -\partial_\rho \Gamma^\rho_{\mu\nu} + \partial_\mu \Gamma^\rho_{\rho\nu} + \Gamma^\rho_{\sigma\nu}\Gamma^\sigma_{\mu\rho} - \Gamma^\rho_{\mu\nu}\Gamma^\sigma_{\sigma\rho} = 0 \tag{9.4}$$

With the metric components we can calculate the Ricci tensor

$$g_{tt} = A(r), \quad g_{rr} = -B(r), \quad g_{\theta\theta} = -r^2, \quad g_{\phi\phi} = -r^2 \sin^2\theta$$

The algebra will be left firmly shut in Hobson (or the course notes) - here only useful results are quoted for 'completeness':

- Where primes denote derivatives w.r.t. r,

$$\Gamma^\rho_{\rho\sigma} = \left(\frac{A'}{2A} + \frac{B'}{2B} + \frac{2}{r}\right)\delta^r_\sigma + \cot\theta\,\delta^\theta_\sigma;$$

- The off-diagonal components of $R_{\mu\nu}$ are identically zero.

The results are

$$A(r) = \alpha\left(1 + \frac{k}{r}\right), \quad B(r) = \left(1 + \frac{k}{r}\right)^{-1} = \frac{\alpha}{A(r)} \tag{9.5}$$

where α and k are integration constants. These are determined by considering the weak-field limit at large r; we require that the line element takes the form

$$\mathrm{d}s^2 \approx \left(1 + \frac{2\Phi}{c^2}\right)\mathrm{d}(ct)^2 + \ldots$$

where $\Phi = -GM/r$ is the Newtonian potential outside a spherical mass M. Then

$$A(r)\,\mathrm{d}t^2 \to c^2\left(1 - \frac{2GM}{c^2 r}\right)\mathrm{d}t^2 \implies \alpha = c^2, \quad k = -\frac{2GM}{c^2}$$

The Schwarzschild metric for the vacuum outside a static, spherically-symmetric body of mass M is

$$\boxed{\mathrm{d}s^2 = c^2\left(1 - \frac{2\mu}{r}\right)\mathrm{d}t^2 - \left(1 - \frac{2\mu}{r}\right)^{-1}\mathrm{d}r^2 - r^2\,\mathrm{d}\Omega^2} \tag{9.6}$$

where $\mu \equiv GM/c^2$. The Schwarzschild metric is valid down to the surface of the spherical distribution; here the EFEs no longer hold. There is a *coordinate singularity* at the Schwarzschild radius $r_s = 2\mu$. As $r \to \infty$, Eq. 9.6 tends towards the Minkowski metric and the spacetime is *asymptotically flat*.

Birkhoff's theorem

If we drop the static constraint then we have Eq. 9.2

$$\mathrm{d}s^2 = \bar{A}(\bar{t}, \bar{r})\,\mathrm{d}\bar{t}^2 - \bar{B}(\bar{t}, \bar{r})\,\mathrm{d}\bar{r}^2 - \bar{r}^2\,\mathrm{d}\Omega^2$$

Substituting this into the empty-space EFEs $R_{\mu\nu} = 0$ leads to the same Schwarzschild metric. Birkhoff's theorem states

> Any spherically-symmetric solution of the Einstein field equations in vacuum is given by the Schwarzschild solution; it is static and asymptotically flat.

For example, a spherical star undergoing radial pulsations has a static external metric; it cannot emit gravitational waves. This is similar to Gauss's theorem in that any mass redistribution in a spherical system that conserves mass has a static external gravitational field. The converse is not true: a matter distribution that gives rise to Schwarzschild geometry outside it need not be spherically symmetric.

9.3 Geodesics in Schwarzschild spacetime

To consider the motion of a free-falling particle in the Schwarzschild solution, we use the Lagrangian approach in Section 4.8

$$L = c^2\left(1-\frac{2\mu}{r}\right)\dot{t}^2 - \left(1-\frac{2\mu}{r}\right)^{-1}\dot{r}^2 - r^2\dot{\theta}^2 - r^2\sin^2\theta\,\dot{\phi}^2 \tag{9.7}$$

Through the EL equations, we obtain for θ

$$\ddot{\theta} + \frac{2}{r}\dot{r}\dot{\theta} - \sin\theta\cos\theta\,\dot{\phi}^2 = 0 \tag{9.8}$$

A possible solution is $\theta = \pi/2$, i.e. planar motion in the equatorial plane. Given spherical symmetry, wlog we can always consider motion in this plane. We choose $\pi/2$ because the last term in the Lagrangian simplifies to $r^2\dot{\phi}^2$.

Next, L has no dependence on t nor ϕ (the first being a consequence of the solution being stationary), so the EL equations for t and ϕ reduce to, respectively,

$$\begin{cases} \frac{\partial L}{\partial \dot{t}} = \text{const.} \quad \Longrightarrow \left(1-\frac{2\mu}{r}\right)\dot{t} = k \\ \frac{\partial L}{\partial \dot{\phi}} = \text{const.} \quad \Longrightarrow r^2\dot{\phi} = h \qquad \text{in the plane } \theta = \frac{\pi}{2} \end{cases} \tag{9.9}$$

with $k, h = \text{const.}$ Lastly, the EL equations for r are too complicated. Instead we use the first integral for L for non-null (massive) and null (massless) geodesics respectively, noting that because we have set $\theta = \pi/2$ we have reduced the Lagrangian to

$$\boxed{L = c^2\left(1-\frac{2\mu}{r}\right)\dot{t}^2 - \left(1-\frac{2\mu}{r}\right)^{-1}\dot{r}^2 - r^2\dot{\phi}^2 = \begin{cases} c^2 & \text{if massive;} \\ 0 & \text{if massless.} \end{cases}} \tag{9.10}$$

How do we know the constant is c^2? h is the specific angular momentum; this arises from symmetry under rotations about the z-axis. k is related to the energy of the particle as measured by a stationary observer. Consider an (obviously massive) observer at rest. In the Schwarzschild metric the 4-velocity of the *observer* is of the form $u^\mu = \mathcal{A}\delta^\mu_0$, where

$$c^2 = g_{\mu\nu}u^\mu u^\nu \implies \mathcal{A} = \left(1-\frac{2\mu}{r}\right)^{-1/2}$$

while the massive *particle* with mass m and energy E has 4-momentum $p^\mu = mu^\mu$; in particular $p^0 = m\dot{t}$ (w.r.t the proper time of the particle).

In local inertial (Minkowski) coordinates the *observer* has 4-velocity $u'^\mu = (c, \vec{0})$ and the *particle* has $p'^0 = E/c$ so that $\boxed{E = \eta_{\mu\nu}u'^\mu p'^\nu = g_{\mu\nu}u^\mu p^\nu = g_{00}u^0p^0}$. Then it follows that

$$E = g_{00}\mathcal{A}p^0 = mc^2\dot{t}\left(1-\frac{2\mu}{r}\right)^{1/2} = \boxed{kmc^2\left(1-\frac{2\mu}{r}\right)^{-1/2}} \quad \text{where} \quad k = \left(1-\frac{2\mu}{r}\right)\dot{t} \tag{9.11}$$

kmc^2 is the energy of the particle measured by a stationary observer as $r \to \infty$ (is r the distance between the particle and the observer?) As mc^2 is the minimum possible energy, we require $k \geq 1$.

For a massless particle, invoking Eq. 5.30, $p^0 = \dot{t}$ so

$$E = c^2\dot{t}\left(1-\frac{2\mu}{r}\right)^{1/2} = \boxed{kc^2\left(1-\frac{2\mu}{r}\right)^{-1/2}} \tag{9.12}$$

We require $k \geq 0$ for the massless particle to reach spatial infinity.

9.4 The energy equation and the effective potential

An energy equation analogous to the Newtonian case is obtained by eliminating $\dot{t}$ and $\dot{\phi}$ (Eq. 9.9) from Eq. 9.10 (for the massive case, $L = c^2$). Then

$$\frac{1}{2}\dot{r}^2 - \frac{GM}{r} + \frac{h^2}{2r^2}\left(1 - \frac{2\mu}{r}\right) = \frac{1}{2}\dot{r}^2 + V_{\rm eff}(r) = \frac{1}{2}c^2(k^2 - 1) \tag{9.13}$$

where $V_{\rm eff}$ is similar to the Newtonian case except that the centrifugal term is modified by the factor $(1 - 2\mu/r)$. At small r (for non-zero h) the sign of the centrifugal barrier is reversed. This differs from Newtonian gravity, in which there is a centrifugal/angular momentum barrier which prevents particles from reaching $r = 0$; in GR an in-falling particle with sufficient energy will spiral into $r = 0$.

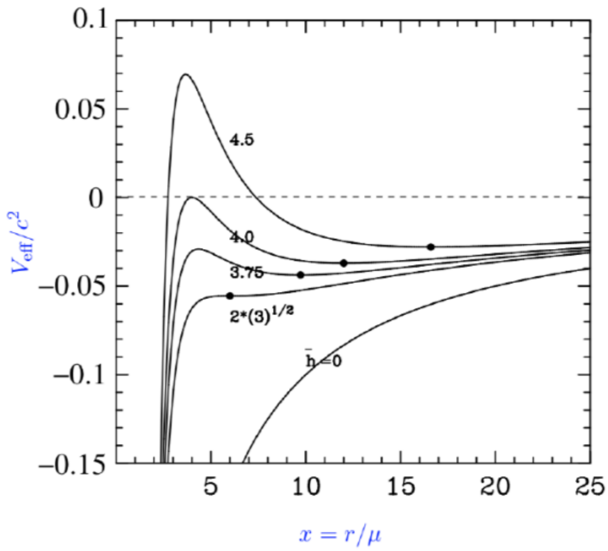


Figure 4: Effective potentials for different values of $\bar{h} = h/(\mu c)$. Dots show locations of stable circular orbits (r_+).

The stationary points and their nature are found from solving, respectively,

$$\frac{\mathrm{d}V_{\rm eff}}{\mathrm{d}r} = \frac{\mu c^2}{r^2} + \frac{h^2}{r^3}\left(\frac{3\mu}{r} - 1\right) = 0 \implies \boxed{r_\pm = \frac{h}{2\mu c^2}\left(h \pm \sqrt{h^2 - 12\mu^2 c^2}\right)} \tag{9.14}$$

$$\frac{\mathrm{d}^2V_{\rm eff}}{\mathrm{d}r^2} = -\frac{2\mu c^2}{r^3} + \frac{3h^2}{r^4}\left(1 - \frac{4\mu}{r}\right) \implies \boxed{\left.\frac{\mathrm{d}V_{\rm eff}}{\mathrm{d}r}\right|_{r_\pm} = \frac{h^2}{r_\pm^5}(r_\pm - 6\mu)} \tag{9.15}$$

r_- corresponds to a local maximum and r_+ to a local minimum. At $r_\pm$ the effective potential is

$$\boxed{V_{\rm eff}(r_\pm) = \frac{h^2}{2r_\pm^3}(4\mu - r_\pm)} \tag{9.16}$$

- Thus circular orbits with $r > 4\mu$ are necessarily bound as $V_{\rm eff}(r) < 0$;
- The innermost stable circular orbit (ISCO) is at $h = \sqrt{12}\mu c$, where $r_\pm = r = 6\mu$;
- For $h > \sqrt{12}\mu c$,
  - Circular orbits at the two stationary points $r_\pm$;
  - As $\bar{h} \to \infty$, $r_+ \to \infty$ and $r_- \to 3\mu$ from above; therefore
  - Unstable innermost orbit has radius in range $3\mu < r_- \leq 6\mu$;
  - Stable outermost orbit has $r_+ > 6\mu$.

Massless particles, $L = 0$

Starting from Eq. 9.10 and eliminating $\dot{t}$ and $\dot{\phi}$ (Eq. 9.9) gives

$$\frac{1}{2}\dot{r}^2 + \frac{h^2}{2r^2}\left(1 - \frac{2\mu}{r}\right) = \frac{1}{2}\dot{r}^2 + V_{\text{eff}}(r) = \frac{1}{2}c^2k^2 \tag{9.17}$$

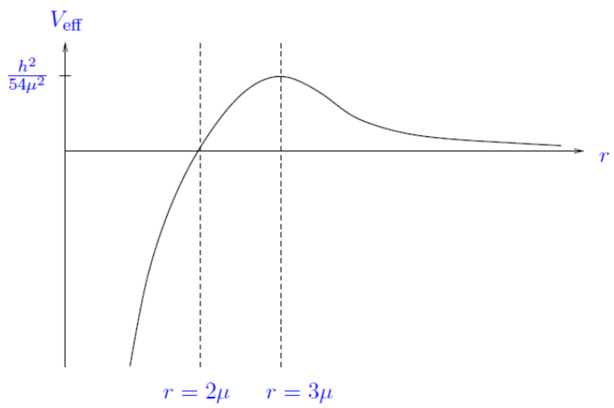


Figure 5: The effective potential has the same shape for all non-zero h.

There is a single stationary point (maximum, so unstable circular orbit) at $r_s = 3\mu$, as

$$\frac{\mathrm{d}V_{\text{eff}}}{\mathrm{d}r} = -\frac{h^2}{r^3}\left(1 - \frac{3\mu}{r}\right) = 0 \implies r_s = 3\mu \quad \text{and} \quad V_{\text{eff}}(3\mu) = \frac{h^2}{54\mu^2} \tag{9.18}$$

Consider a massless particle (e.g. a photon) moving inwards from large radii, with angular momentum h. Rearranging Eq. 9.17 for $r = 3\mu$

$$\frac{1}{2}\dot{r}^2 = \frac{1}{2}c^2k^2 - \frac{h^2}{54\mu^2}$$

we see that if $c^2k^2/2 < h^2/(54\mu^2)$ there are no solutions for $\dot{r}$; the particle cannot cross the barrier at $r = 3\mu$. Instead there is a single turning point of closest approach where $c^2k^2/2 = V_{\text{eff}}(r)$, and the particle subsequently escapes to infinity. If $c^2k^2/2 > h^2/(54\mu^2)$, the particle is captured by the object. Noting that the ratio of h/k is what matters, we derive the *impact parameter* by converting Eq. 9.17 into an orbit-shape equation

$$\left(\frac{\mathrm{d}r}{\mathrm{d}\phi}\right)^2 = \frac{\dot{r}^2}{\dot{\phi}^2} = \frac{r^4\dot{r}^2}{h^2} = r^2\left[\frac{c^2k^2}{h^2}r^2 - 1 + \frac{2\mu}{r}\right] \tag{9.19}$$

invoking $r^2\dot{\phi} = h$ and substituting for $\dot{r}^2$ using Eq. 9.17. At large radii ($r \to \infty$) the metric tends to the Minkowski metric; we expect orbits to approach straight lines described by $r\sin\phi = b$, for impact parameter b, and as $\phi \to 0$ wlog. Then

$$r = \frac{b}{\sin\phi} \implies \left(\frac{\mathrm{d}r}{\mathrm{d}\phi}\right)^2 = \left(-\frac{b}{\sin^2\phi}\cos\phi\right)^2 = \frac{r^4}{b^2}\left(1 - \frac{b^2}{r^2}\right) = r^2\left(\frac{r^2}{b^2} - 1\right)$$

By comparison with Eq. 9.19 in the limit $r \to \infty$, the impact parameter is

$$b = \frac{h}{ck} \tag{9.20}$$

This is because gravitational fields act *achromatically* (independent of energy) on massless particles. As $p^\mu = \mathrm{d}x^\mu/\mathrm{d}\lambda$, different energy = constant scaling of the affine parameter.

Note again that the particle is captured by the object if $c^2k^2/2 > h^2/(54\mu^2) \implies b < \sqrt{27}\mu$.

Gas in an accretion disc

Gas in an accretion disc around a compact object moves in quasi-circular orbits. Angular momentum is lost via viscosity. The packet of gas moves slowly inwards until it can no longer follow a stable circular orbit, at which point it falls into the object. The energy radiated from this process can be estimated relativistically.[2] Consider a particle originally in a large circular orbit, parameter k, perturbed to radius r such that k is conserved. The particle has excess energy as it is moving too fast to enter a circular orbit. We want to work out the k required for a circular orbit at a given r.

$$\text{From Eq. 9.14} \quad \frac{\mathrm{d}V_{\text{eff}}}{\mathrm{d}r} = 0 \implies \mu c^2 r^2 = h^2(r - 3\mu) \tag{9.21}$$

$$\text{From Eq. 9.13} \quad \dot{r} = 0 \implies \frac{1}{2}c^2(k^2 - 1) = V_{\text{eff}}(r) = -\frac{h^2}{2r^2}\left(1 - \frac{4\mu}{r}\right)$$

where the last equality invokes Eq. 9.21. Solving for k gives

$$k = \frac{(1 - 2\mu/r)}{(1 - 3\mu/r)^{1/2}} \tag{9.22}$$

The energy of the particle is $E = kmc^2(1 - 2\mu/r)^{-1/2}$ (Eq. 9.11). For a circular orbit at r we use k given by Eq. 9.22. However, originally the particle starts in a circular orbit at $r \gg \mu$, where $k \approx 1$, so for a local stationary observer at r we take $k = 1$. Thus

$$\frac{\Delta E}{mc^2} = (1 - k)\left(1 - \frac{2\mu}{r}\right)^{-1/2} = \left(1 - \frac{2\mu}{r}\right)^{-1/2} - \left(\frac{1 - 2\mu/r}{1 - 3\mu/r}\right)^{1/2} \tag{9.23}$$

We recover the Newtonian result in the limit $\mu \ll r$. At the ISCO ($r = 6\mu$), $\Delta E \approx 0.07mc^2$; 7% of the rest-mass energy is radiated as the packet of gas moves through the accretion disc to the innermost edge at $r = 6\mu$.

9.5 Gravitational redshift

This generalises to any *static* metric, where $g_{0i} = 0$ and $g_{\mu\nu}$ is independent of t. The EL equation for t gives $g_{00}\dot{t} = bc^2$ with b a constant. The 4-velocity of a stationary observer has $u^0 = c/\sqrt{g_{00}}$. The energy of the photon measured by this stationary observer is, recalling Section 9.3 and $p^0 = \dot{t}$,

$$E = g_{00}u^0 p^0 = \frac{bc^2}{\sqrt{g_{00}/c^2}} \tag{9.24}$$

Then it follows that, where ν is the observed frequency,

$$\frac{\nu_R}{\nu_E} = \sqrt{\frac{g_{00}(E)}{g_{00}(R)}} \tag{9.25}$$

where $_E$ and $_R$ represent emitted and received respectively, and both observers are stationary w.r.t. the mass. In the Schwarzschild case, combining Eq. 9.17 and $E = h\nu$ gives

$$\nu = \frac{kc^2}{h}(1 - 2\mu/r)^{-1/2} \implies \frac{\nu_R}{\nu_E} = \left(\frac{1 - 2\mu/r_E}{1 - 2\mu/r_R}\right)^{1/2} \tag{9.26}$$

as $k = \text{const}$. In terms of redshift z

$$1 + z = \left(\frac{\nu_R}{\nu_E}\right)^{-1} \tag{9.27}$$

For observations (i.e. the receiver) at infinity, the redshift z_∞ tends to infinity as the point of emission approaches $r_s = 2\mu$, as

$$1 + z_\infty = \left(1 - \frac{2\mu}{r_E}\right)^{-1/2} \tag{9.28}$$

[2] The non-relativistic Newtonian result (following from the virial theorem) is $\frac{\Delta E_N}{mc^2} = \frac{\mu}{2r}$.

10 Schwarzschild black holes

10.1 Invariants of singularities

So far we have been concerned with motion in the region $r > r_s = 2\mu$ (region I). Here we deal with the region $r < 2\mu$ (region II) and the hypersurface $r = 2\mu$ and discuss the causal structure of the Schwarzschild solution.

Per Eq. 9.6 the Schwarzschild line element is

$$\mathrm{d}s^2 = c^2\left(1 - \frac{2\mu}{r}\right)\,\mathrm{d}t^2 - \left(1 - \frac{2\mu}{r}\right)^{-1}\,\mathrm{d}r^2 - r^2\,\mathrm{d}\Omega^2$$

This is singular at $r = 0$ and $r_s = 2\mu = 2GM/c^2$. Recall that the Schwarzschild solution is valid down to the surface of the spherical distribution (where the EFEs no longer hold). This is fine so long as the radius of the body exceeds r_s. To investigate the nature of the singularities, we consider invariants of spacetime there. Formed from the Riemann curvature tensor, the Kretschmann scalar is a natural invariant as $R_{\mu\nu} = 0$ by construction. Calculating using the Schwarzschild metric

$$R_{\mu\nu\sigma\rho}R^{\mu\nu\sigma\rho} = \frac{48\mu^2}{r^6} \tag{10.1}$$

The Kretschmann scalar is related to tidal effects (it has the same dependence on M and r as the square of the tidal tensor in Newtonian gravity). It is regular at $r = 2\mu$ - this is a coordinate singularity; but $r = 0$ is an intrinsic singularity, where spacetime curvature is infinite.

In region I, $\mathbf{e}_0 \equiv \partial/\partial t$ is timelike as $g_{00} > 0$, but in region II $\mathbf{e}_0$ is spacelike and the basis vector ∂/∂_r is timelike. In region II a particle cannot stay at fixed (r, θ, ϕ) (otherwise their worldline would be spacelike).

10.2 Radial null geodesics

Lightcones separate timelike curves from spacelike curves hence are indicators of causal structure. They are constructed by considering radial null geodesics, for which

$$0 = \mathrm{d}s^2 = c^2\left(1 - \frac{2\mu}{r}\right)\,\mathrm{d}t^2 - \left(1 - \frac{2\mu}{r}\right)^{-1}\mathrm{d}r^2 \implies \frac{\mathrm{d}(ct)}{\mathrm{d}r} = \pm\left(1 - \frac{2\mu}{r}\right)^{-1} \tag{10.2}$$

which is solved by

$$ct = \pm\left(r + 2\mu\ln\left|\frac{r}{2\mu} - 1\right|\right) + \text{const.} \tag{10.3}$$

In region I, these correspond to families of out/[in]going radial null geodesics, which are related by time reversal ($t \to -t$; this follows from time reversal symmetry of the Schwarzschild metric).

- $r \gg 2\mu$: the linear term dominates over the logarithmic term in Eq. 10.3, so $r \to \pm ct$;
- $r \to 2\mu$: $ct \to \mp\infty$ so the lightcones are squashed in the radial direction; there is a discontinuity at $r = 2\mu$; it seems to take an infinite coordinate time for an *ingoing* photon to reach $r = 2\mu$, but this is from the perspective of an observer at infinity, as
- An ingoing photon can cross $r = 2\mu$ with a finite change in its affine parameter: recall from Eq. 9.9 that, for affine parameter τ,

  $$\frac{\mathrm{d}t}{\mathrm{d}\tau} = k\left(1 - \frac{2\mu}{r}\right)^{-1} \implies \frac{\mathrm{d}r}{\mathrm{d}\tau} = k\frac{\mathrm{d}r}{\mathrm{d}t}\left(1 - \frac{2\mu}{r}\right)^{-1} = \pm kc \implies \boxed{r = \pm ck\tau + \text{const.}}$$

  where the second-last equality follows from Eq. 10.2. So from its perspective the photon crosses r_s in a finite time and is able to reach $r = 0$ with increasing τ.

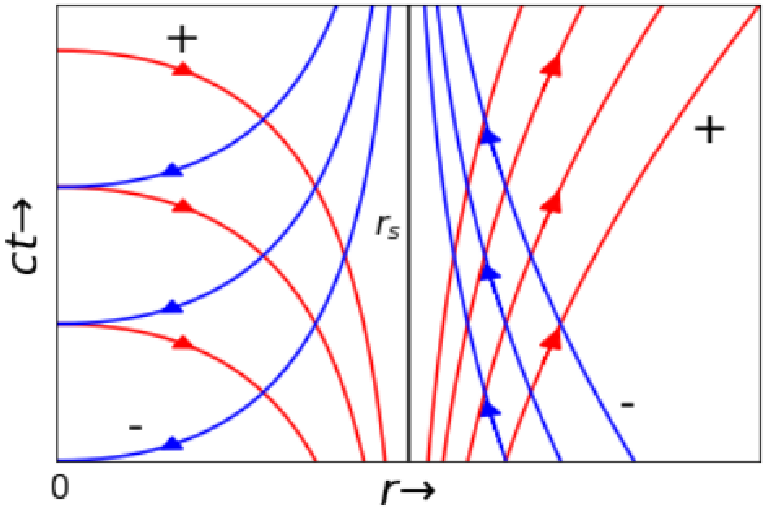


Figure 6: Enter Caption

- Interestingly in region II, as k is necessarily positive, it follows that $\mathrm{d}t/\mathrm{d}\tau < 0$ - ingoing null geodesics move towards $r = 0$ with decreasing t. The forward lightcone is tipped over towards $r = 0$; any particle that enters region II (in the causal future of region I, i.e. a black hole) inevitably falls towards the $r = 0$ intrinsic singularity;
- The hypersurface $r_s = 2\mu$ is an *event horizon*, a one-way membrane;
- Again, r is linear and finite in τ for *outgoing* null geodesics; trying to extend these geodesics beyond r_s results in them joining onto geodesics in region II with increasing r and decreasing τ and t. The forward lightcone is directed *away* from $r = 0$; particles are inevitably expelled to $r > 2\mu$; this region is a *white hole*;
- Hence there are two type-II regions: black holes for ingoing geodesics corresponding to the causal future of region I, and white holes for outgoing geodesics corresponding to the causal past of region I. We show in a later part that there are two type-I regions, and that altogether this constitutes the entire spacetime. (or never...)

10.3 Radial particle geodesics

Setting $h = 0$ we have, from Eq. 9.13,

$$\frac{1}{2}\dot{r}^2 - \frac{GM}{r} = \frac{1}{2}\dot{r}^2 - \frac{\mu c^2}{r} = \frac{1}{2}c^2(k^2 - 1) \tag{10.4}$$

Consider a particle that starts at rest at infinity, so that $k = 1$ (recall that $E = kmc^2$ is the energy of a particle measured by a stationary observer as $r \to \infty$). For an infalling particle,

$$\dot{r} = -\sqrt{\frac{2\mu c^2}{r}} \implies c(\tau - \tau_0) = \frac{2}{3}\left[\left(\frac{r_0^3}{2\mu}\right)^{1/2} - \left(\frac{r^3}{2\mu}\right)^{1/2}\right] \quad \text{with} \quad r(\tau_0) = r_0 \tag{10.5}$$

Coordinate time t as a function of r is determined by

$$\frac{c\dot{t}}{\dot{r}} = -\sqrt{\frac{r}{2\mu}}\left(1 - \frac{2\mu}{r}\right)^{-1} \implies c(t - t_0) = -2\mu\left[\frac{2}{3}x^{3/2} + 2\sqrt{x} + \ln\left|\frac{\sqrt{x} - 1}{\sqrt{x} + 1}\right|\right]_{r_0/(2\mu)}^{r/(2\mu)} \tag{10.6}$$

where $t(r_0) = t_0$, which diverges logarithmically ($t \to \infty$) for $r \to 2\mu$. For $r < 2\mu$, t decreases with increasing τ. For a stationary observer at infinity, light from the particle only reaches the observer if it is emitted at $r > 2\mu$. From their perspective the light signals emitted as $r \to 2\mu$ are infinitely redshifted, and the particle never reaches $r = 2\mu$.

10.4 Eddington-Finkelstein coordinates

In this section we construct a non-singular coordinate system that covers both region I and the type-II region in its causal future.

There are two types, *ingoing* and *outgoing* EF coordinates, which correspond to ingoing and outgoing null geodesics being continuous through $r = 2\mu$ respectively. Recalling Eq. 10.3 for null geodesics

$$ct = \mp \left(r + 2\mu \ln \left| \frac{r}{2\mu} - 1 \right| \right) + \text{const.}$$

we define a new time coordinate t' or t^* by

$$c\tilde{t} \equiv c \pm 2\mu \ln \left| \frac{r}{2\mu} - 1 \right| \tag{10.7}$$

so that the in/outgoing null geodesics are straight lines 45 degrees to the coordinate axes, described by $c\tilde{t} = \mp r + \text{const.}$, and the out/ingoing null geodesics are described by

$$c\tilde{t} = \pm \left(r + 4\mu \ln \left| \frac{r}{2\mu} - 1 \right| \right) + \text{const.} \tag{10.8}$$

which are each discontinuous at $r = 2\mu$. The line element follows from

$$c\,\mathrm{d}\tilde{t} = c\,\mathrm{d}t \pm \left(\frac{r}{2\mu} - 1 \right)^{-1} \mathrm{d}r$$

substituting this for $\mathrm{d}t^2$ in the Schwarzschild line element gives, after simplifying,

$$\mathrm{d}s^2 = c^2 \left(1 - \frac{2\mu}{r} \right) \mathrm{d}t'^2 \mp \frac{4\mu c}{r} \mathrm{d}t' \,\mathrm{d}r - \left(1 + \frac{2\mu}{r} \right) \mathrm{d}r^2 - r^2 \,\mathrm{d}\Omega^2 \tag{10.9}$$

The metric is regular for $0 < r < \infty$ and is non-diagonal, so the inverse metric exists.

It is possible to combine the two to form *Kruskal-Szekeres* coordinates that cover global spacetime in a non-singular way.

10.5 Gravitational collapse

Consider the spherically-symmetric collapse of a cloud of pressure-free dust. Assume that gravity is the only force. We treat the outside edge of the cloud as a massive particle free-falling radially in Schwarzschild spacetime. This is possible because we know by Birkhoff's theorem that the spacetime outside the dust is described by the Schwarzschild metric. We assume that the collapse starts from rest at infinity ($k = 1$) and we take the perspective of a stationary observer at $r \gg \mu$. We never see the surface of the dust cloud pass through $r = 2\mu$. As $r \to 2\mu$ the light gets infinitely redshifted so the frequency observed and the arrival rate of photons fall to zero.

Suppose the edge emits a radially *outgoing* photon at (t_E, r_E), which is received at (t_R, r_R). These coordinates lie on the *same outgoing null geodesic*, so that

$$ct_R - r_R - 2\mu \ln \left(\frac{r_R}{2\mu} - 1 \right) = ct_E - r_E - 2\mu \ln \left(\frac{r_E}{2\mu} - 1 \right) \tag{10.10}$$

Any term containing r_R only is constant. The relation between t_E and r_E is given by the path of the edge, so we consider Eq. 10.6 in the limit $r_E \to 2\mu$, which gives

$$ct_E \approx 2\mu \ln \left(\frac{\sqrt{r_E/2\mu} + 1}{\sqrt{r_E/2\mu} - 1} \right) + \text{const.} \, (\gg r_E) \tag{10.11}$$

Then with some loose algebra

$$\begin{aligned}
ct_R &\to 2\mu \ln\left(\frac{\sqrt{r_E/2\mu}+1}{\sqrt{r_E/2\mu}-1}\right) - 2\mu\ln\left(\frac{r_E}{2\mu}-1\right) + \text{const.} \\
&= 2\mu\ln\left(\sqrt{\frac{r_E}{2\mu}}-1\right)^{-2} + \text{const.} \\
&\approx -4\mu\ln\left(\frac{r_E}{2\mu}-1\right) + \text{const.} \\
&\implies \boxed{r_E(t_R) = 2\mu + a\exp(-ct_R/4\mu)}
\end{aligned} \tag{10.12}$$

where a is some constant. r_E approaches $r = 2\mu$ exponentially with time constant $4\mu/c$.

Redshift

We can't use the result from Section 9.5 as that assumes the emitter is stationary w.r.t. the mass; here the emitter is moving towards the mass. As before we use $E = g_{\mu\nu}p^\mu u^\nu$ (see Section 9.3). Taking the ratio of energies we have

$$\frac{\nu_R}{\nu_E} = \frac{\mathbf{g}(\mathbf{p}_R, \mathbf{u}_R)}{\mathbf{g}(\mathbf{p}_E, \mathbf{u}_E)} \tag{10.13}$$

where $\mathbf{p}_E$ is the 4-momentum of a radial outgoing photon emitted at r_E, and $\mathbf{u}_E$ is the 4-velocity of the edge at r_E. Recall the 4-velocity is given by $u^\mu \equiv \mathrm{d}x^\mu/\mathrm{d}\tau$. For the radially inward-moving emitter, with $k = 1$ we have

$$u_E^\mu = \left(\left(1-\frac{2\mu}{r_E}\right)^{-1}, -\sqrt{\frac{2\mu c^2}{r_E}}, 0, 0\right) \tag{10.14}$$

which comes from Eqs. 9.9 and 10.5 respectively. For the static observer, $\dot{t}$ has the same form, but taking $r \to \infty$ this becomes

$$u_R^\mu = (1, 0, 0, 0) \tag{10.15}$$

The 4-momentum of a radial outgoing photon is, for some general affine parameter λ, where still $\dot{t} = k(1-2\mu/r)^{-1}$,

$$\begin{aligned}
p^\mu = \frac{\mathrm{d}t}{\mathrm{d}\lambda}\left(1, \frac{\mathrm{d}r}{\mathrm{d}t}, 0, 0\right) &= k\left(1-\frac{2\mu}{r}\right)^{-1}\left(1, c\left(1-\frac{2\mu}{r}\right), 0, 0\right) \\
\implies p^\mu &= \left(k\left(1-\frac{2\mu}{r}\right)^{-1}, ck, 0, 0\right)
\end{aligned} \tag{10.16}$$

Then

$$\begin{aligned}
g_{\mu\nu}p_E^\mu u_E^\nu &= c^2\left(1-\frac{2\mu}{r_E}\right)\cdot k\left(1-\frac{2\mu}{r_E}\right)^{-1}\left(1-\frac{2\mu}{r_E}\right)^{-1} - \left(1-\frac{2\mu}{r_E}\right)^{-1}(ck)\cdot\left(-\sqrt{\frac{2\mu c^2}{r_E}}\right) \\
&= kc^2\left(1-\frac{2\mu}{r_E}\right)^{-1}\left(1+\sqrt{\frac{2\mu}{r_E}}\right) = kc^2\left(1-\sqrt{\frac{2\mu}{r_E}}\right)^{-1}
\end{aligned}$$

and $g_{\mu\nu}p_R^\mu u_R^\nu = kc^2$, so

$$\frac{\nu_R}{\nu_E} = 1 - \sqrt{\frac{2\mu}{r_E}} \tag{10.17}$$

which tends to zero as $r_E \to 2\mu$. The time-dependent form follows from Eq. 10.12

$$\frac{\nu_R}{\nu_E} \sim 1 - [1 - a\exp(-ct_R/4\mu)]^{-1/2} \sim \frac{1}{2}a\exp(-ct_R/4\mu) \tag{10.18}$$

11 Classical tests of General Relativity

11.1 Orbit shapes for massive and massless particles

We start with Eqs. 9.13 and 9.17.

$$\frac{1}{2}\dot{r}^2 - \frac{GM}{r} + \frac{h^2}{2r^2}\left(1 - \frac{2\mu}{r}\right) = \frac{1}{2}c^2(k^2-1) \quad \text{massive}$$
$$\frac{1}{2}\dot{r}^2 + \frac{h^2}{2r^2}\left(1 - \frac{2\mu}{r}\right) = \frac{1}{2}c^2k^2 \qquad \text{massless}$$

Following the standard analysis for Newtonian orbits[3], i.e. finding r as a function of ϕ, we make the substitution $r = 1/u$ so that

$$\dot{r} = \dot{\phi}\frac{\mathrm{d}r}{\mathrm{d}\phi} = h\frac{1}{r^2}\frac{\mathrm{d}r}{\mathrm{d}\phi} = -h\frac{\mathrm{d}u}{\mathrm{d}\phi}$$

Then

$$\frac{1}{2}\left(\frac{\mathrm{d}u}{\mathrm{d}\phi}\right)^2 - \frac{GM}{h^2}u + \frac{u^2}{2}(1-2\mu u) = \frac{c^2}{2h}(k^2-1) \quad \text{massive}$$
$$\frac{1}{2}\left(\frac{\mathrm{d}u}{\mathrm{d}\phi}\right)^2 + \frac{u^2}{2}(1-2\mu u) = \frac{c^2k^2}{2h} \qquad \text{massless}$$

Differentiating w.r.t. ϕ gives

$$\frac{\mathrm{d}^2u}{\mathrm{d}\phi^2} + u - 3\mu u^2 = \begin{cases} GM/h^2 & \text{massive} \\ 0 & \text{massless} \end{cases} \tag{11.1}$$

11.2 Precession of planetary orbits

Rewriting Eq. 11.1 (massive case) as

$$\frac{\mathrm{d}^2u}{\mathrm{d}\phi^2} + u = \frac{GM}{h^2} + \frac{3GM}{c^2}u^2 \tag{11.2}$$

We solve this in the weak relativistic limit $GM \ll rc^2$. This can be done by *perturbing* around a Newtonian orbit. Introduce a *dimensionless* inverse radius U satisfying $u = GMU/h^2$. Then Eq. 11.2 becomes

$$\frac{\mathrm{d}^2U}{\mathrm{d}\phi^2} + U = 1 + \frac{3(GM)^2}{c^2h^2}U^2 = 1 + \alpha U^2 \quad \text{with} \quad \alpha = \frac{3\mu}{r_0} \tag{11.3}$$

where $r_0 \equiv h^2/GM$ is the radius of the Newtonian circular orbit ($e = 0$). Just like in PQM time-independent perturbation theory we seek solutions of the form, with U_0 the Newtonian solution satisfying $\mathrm{d}^2U_0/\mathrm{d}\phi^2 + U_0 = 1$,

$$U = U_0 + \alpha U_1 + \alpha^2 U_2 + \dots \quad \text{with} \quad U_0 = 1 + e\cos\phi$$

To first order in α,

$$\frac{\mathrm{d}^2U_1}{\mathrm{d}\phi^2} + U_1 = U_0^2 = (1+e\cos\phi)^2 = 1 + \frac{e^2}{2} + 2e\cos\phi + \frac{e^2}{2}\cos 2\phi \tag{11.4}$$

$$\implies U_1(\phi) = 1 + \frac{e^2}{2} + e\phi\sin\phi - \frac{e^2}{6}\cos 2\phi$$

Only the third term is important to the correction as it depends linearly on ϕ; the others are negligible when multiplied by α (which is necessarily small).

[3] Recall that the Newtonian result is $\mathrm{d}^2u/\mathrm{d}\phi^2 + u = GM/h^2$.

So then to first order in α, the general relativistic orbit is

$$u(\phi) \approx \frac{GM}{h^2}(1 + e\cos\phi + e\alpha\phi\sin\phi) \tag{11.5}$$

Consider the Taylor expansions of $\cos(\alpha\phi)$ and $\sin(\alpha\phi)$ to first order in $\alpha\phi$:

$$\cos(\alpha\phi) = 1 + \mathcal{O}[(\alpha\phi)^2] \quad \text{and} \quad \sin(\alpha\phi) = \alpha\phi + \mathcal{O}[(\alpha\phi)^3]$$

So we can rewrite Eq. 11.5 as

$$u(\phi) \approx \frac{GM}{h^2}\left(1 + e[\cos\phi\cos(\alpha\phi) + \sin\phi\sin(\alpha\phi)]\right) = \frac{GM}{h^2}\left(1 + e\cos\left[\phi(1-\alpha)\right]\right) \tag{11.6}$$

The orbit is *not closed* as r is periodic in ϕ with period $2\pi/(1-\alpha)$; the ellipse precesses. At perihelion, ϕ increases per revolution by

$$\Delta\phi = 2\pi\left(\frac{1}{1-\alpha} - 1\right) \approx 2\pi\alpha = \frac{6\pi GM}{a(1-e^2)c^2} \tag{11.7}$$

where the third equality follows from the Newtonian result $a = h^2/[GM(1-e^2)]$. $\Delta\phi$ is largest when the orbit is small and highly eccentric, e.g. Mercury, where $\Delta\phi = 43''$ per century.

11.3 Gravitational deflection of light

Rewriting Eq. 11.1 (massless case) as

$$\frac{\mathrm{d}^2u}{\mathrm{d}\phi^2} + u = \frac{3GM}{c^2}u^2 \tag{11.8}$$

we solve this perturbatively. For $M = 0$, Eq. 11.8 is solved by $u = \sin\phi/b$ where b is the impact parameter. Then $bu = U_0 = \sin\phi$ becomes our dimensionless unperturbed solution, with Eq. 11.8 becoming

$$\frac{\mathrm{d}^2(bu)}{\mathrm{d}\phi^2} + bu = \frac{3GM}{c^2b^2}(bu)^2 = \beta(bu)^2$$

and we proceed perturbatively, where the U_i are dimensionless

$$bu(\phi) = \sin\phi + \beta U_1(\phi) + \beta^2 U_2(\phi) + \ldots$$

To first order in β,

$$\frac{\mathrm{d}^2U_1}{\mathrm{d}\phi^2} + U_1 = U_0^2 = \sin^2\phi = \frac{1}{2}(1 - \cos 2\phi) \tag{11.9}$$

which is solved by

$$U_1(\phi) = C_1\sin\phi + C_2\cos\phi + \frac{1}{2}\left(1 + \frac{1}{3}\cos 2\phi\right)$$

We require $bu \to \sin\phi$ as $\phi \to \pi$ (why?) so $C_1 = 0$ and $C_2 = 2/3$, so that to first order in β

$$u(\phi) \approx \frac{\sin\phi}{b} + \frac{3GM}{2c^2b^2}\left(\frac{4}{3}\cos\phi + 1 + \frac{1}{3}\cos 2\phi\right) \tag{11.10}$$

As $r \to \infty$, $u \to 0$ and $\phi \to -\Delta\phi$ ($\Delta\phi$ is the deflection); for a small deflection, $\sin\phi \approx \phi \to -\Delta\phi$ and $\cos n\phi \approx 1$ so

$$0 = -\frac{\Delta\phi}{b} + \frac{8}{3}\frac{3GM}{2c^2b^2} = 0$$

So we obtain the gravitational deflection formula[4]

$$\boxed{\Delta\phi = \frac{4GM}{c^2b}} \tag{11.11}$$

[4]This is twice the value obtained with a Newtonian calculation for a particle travelling at c. This reflects the fact that for a massless particle both the g_{tt} and g_{rr} metric perturbations contribute to the particle dynamics, while for a slowly-moving massive particle only g_{tt} is relevant.

12 Cosmology

The Universe looks remarkably symmetric on scales of the order of 100s of Mpcs. Smaller scale structures are treated as perturbations to the highly symmetric background. The symmetry is apparent from e.g. the CMB, which suggests that at *any given time*, the universe looks the same in all directions - isotropic. If the universe has no preferred centre, we can imagine filling space with a whole class of fundamental observers with synchronised clocks, so that they perform measurements simultaneously, and they should agree on what they observe at any given time - isotropy also implies homogeneity. These two conditions form the *cosmological principle*.

It follows that in order to not break isotropy, fundamental observers must comove with matter - if not, they would measure some matter moving away ; they must be free-falling,

As there are no global inertial frames the concept of a globally-valid 'time' parameter is made possible by 'slicing up' spacetime with non-intersecting 3-dim spacelike hypersurfaces labelled by some parameter t. There is generally no preferred way to construct these hypersurfaces,

t is cosmic time.

12.1 Robertson-Walker metric

The evolution of the scale factor $a(t)$ is governed by the EFEs (see Section 12.3).

12.2 Expansion and cosmological redshift

Consider two fundamental observers, one at $\chi = 0$ and another spaced $\Delta\chi$ away, with 4-velocity $u^\mu = \delta^\mu_0$. The proper distance between them is $l(t) = a(t)\Delta\chi$ as $\mathrm{d}t = \mathrm{d}\Omega = 0$. We define the Hubble parameter $H(t)$

$$H(t) \equiv \frac{1}{a}\frac{\mathrm{d}a}{\mathrm{d}t} = \frac{1}{l(t)}\frac{\mathrm{d}l}{\mathrm{d}t} \tag{12.1}$$

For $H > 0$, all fundamental observers see all other fundamental observers (and matter) moving away isotropically with the *same fractional rate*.

Consider a photon emitted by a fundamental observer at $(t_E, \vec{0})$ which is received (later) by a fundamental observer at $(t_R, \chi_R, \theta_R, \phi_R)$. The path of the photon varies with the affine parameter λ in the temporal and radial coordinates only, so its 4-momentum is

$$x^\mu(\lambda) = (t(\lambda), \chi(\lambda), \theta_R, \phi_R) \implies p^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} = (p^0, p^1, 0, 0)$$

and where lowering the index with the metric (for which $g_{00} = c^2$ and $g_{11} = -a^2$) gives $p_0 = c^2 p^0$ and $p_1 = -a^2 p^1$. The photon travels on a null geodesic, so per Eq. 4.53,

$$\frac{\mathrm{d}p_\mu}{\mathrm{d}\lambda} = \frac{1}{2}\frac{\partial g_{\nu\rho}}{\partial x^\mu}p^\nu p^\rho \implies \frac{\mathrm{d}p_1}{\mathrm{d}\lambda} = \frac{1}{2}\left(\frac{\partial g_{00}}{\partial \chi}p^0 p^0 + \frac{\partial g_{11}}{\partial \chi}p^1 p^1\right) = 0 \implies \boxed{p_1 = \text{const.}}$$

as both g_{00} and g_{11} are independent of χ. Using the null condition,

$$g^{\mu\nu}p_\mu p_\nu = 0 \implies c^{-2}(p_0)^2 - a^{-2}(p_1)^2 = 0 \implies \boxed{ap_0 = cp_1 = \text{const.}}$$

To relate this to energy, recall from Section 9.3 that the energy of the photon with 4-momentum p^μ measured by a fundamental observer with 4-velocity u^μ is

$$E = g_{\mu\nu}p^\mu u^\nu = g_{00}p^0 = p_0 \implies \boxed{Ea(t) = \text{const.}}$$

Therefore

$$1 + z \equiv \frac{\lambda_R}{\lambda_E} = \frac{E_E}{E_R} = \frac{a(t_R)}{a(t_E)} \tag{12.2}$$

12.3 Cosmological field equations