ETH Zürich · Institute for Theoretical Physics

General Relativity

Lecture script

Babis Anastasiou

Lectures of the autumn semester 2021
Typeset from the lectures by Tomas Mari Surkau

Contents

0 **Introduction and literature** 1

1 **Special relativity** 3

Conventions . . . 3

1.1 Time dilation and Doppler effect . . . 8

1.2 Dynamics in special relativity . . . 10

1.3 Relativistic energy and momentum . . . 12

1.4 Vectors and tensors in special relativity . . . 12

1.5 Currents and densities . . . 15

1.6 Energy-momentum tensor . . . 17

1.7 Spin . . . 19

2 **The principle of equivalence** 22

2.1 Newtonian limit . . . 26

2.2 Clocks and light in a gravitational field . . . 28

2.3 Geodesics and the motion of a particle in a gravitational field . . . 30

3 **Tensors and tensor densities in general relativity** 32

3.1 Scalars, vectors, tensors . . . 32

3.2 Tensor densities . . . 35

3.3 Transformation of the affine connection . . . 37

3.4 Covariant derivative . . . 38

3.4.1 Covariant curl . . . 39

3.4.2 Covariant divergence . . . 39

3.4.3 Covariant differentiation along a curve . . . 40

4 **Writing laws of physics consistently in general relativity** 42

4.1 Particle mechanics . . . 42

4.1.1 Acceleration . . . 42

4.1.2 Spin . . . 43

4.2 Electrodynamics . . . 44

5 Einstein field equations **47**

5.1 Riemann curvature tensor 47

5.1.1 Parallel transport and the Riemann tensor 49

5.1.2 Algebraic properties of the Riemann tensor 53

5.1.3 The Ricci tensor and Ricci scalar 54

5.1.4 The Bianchi identities 55

5.2 The energy-momentum tensor 55

5.3 Einstein's field equations 56

5.3.1 Coordinate conditions 60

5.3.2 Energy, momentum and angular momentum 62

6 The Schwarzschild solution of Einstein's field equations **66**

6.1 Birkhoff's theorem 70

6.2 Trajectories 71

6.3 Gravitational redshift in the Schwarzschild field 74

6.4 Light deflection 76

6.5 Time delay of radar echoes 77

6.6 Bound orbits 79

6.7 The Schwarzschild radius and the horizon 79

6.7.1 Coordinate singularities and curvature singularities 79

6.7.2 Radial free fall 80

6.7.3 Eddington–Finkelstein coordinates 81

6.7.4 Kruskal–Szekeres coordinates 81

6.7.5 What is a black hole? 83

6.8 Exercises 83

7 Gravitational waves **85**

7.1 Solutions of the linearised equations 86

7.2 Analogy with the electromagnetic field 90

7.3 Energy and momentum of plane waves 91

7.4 Generation of gravitational waves 94

7.5 Effect of a gravitational wave on test masses 100

7.6 The quadrupole formula 101

7.7 Binary systems . . . 104
7.8 Detection of gravitational waves . . . 105
7.9 Exercises . . . 107

8 Action principle for gravity 108

8.1 The gravitational part . . . 108
8.2 Adding a matter field . . . 111
8.3 Adding the electromagnetic field . . . 113

9 Symmetries 115

9.1 Homogeneity and isotropy . . . 117
9.2 Curvature of maximally symmetric spaces . . . 122
9.3 Construction of maximally symmetric spaces . . . 123
9.4 Tensors in maximally symmetric spaces . . . 125
9.5 Spaces with maximally symmetric subspaces . . . 127

0

Introduction and literature

The goal of this lecture series is to understand the phenomena of gravitation at a deeper level. Gravity has been described successfully by Newton's theory. However, Newton's laws are theoretically incomplete. For instance, they predict a gravitational force which acts instantaneously at a distance, in contrast to the expectations from special relativity.

In this lecture series, we will introduce Einstein's theory of general relativity, which extends Newtonian mechanics beyond the limit of small velocities. It is also valid beyond the limit of weak gravitational fields, which are our typical experience within our planetary system.
Remark. More precisely, it is already special relativity that extends Newtonian mechanics to large velocities. General relativity extends Newtonian gravity (and special relativity) to strong and time-dependent gravitational fields, and it reduces to Newton's law for weak static fields and slow sources (Weinberg, Ch. 3; Carroll, Ch. 4). General relativity is still an approximate theory, as it ignores quantum effects ($\hbar \to 0$).

Guiding principle

The principle of equivalence (Einstein):

- Set up reference frames on freely falling bodies.
- For small time intervals and in small regions of space, the effects of gravity are absent!
  Remark. "Absent" holds only to first order: tidal effects (second derivatives of the potential, i.e. curvature) cannot be transformed away in any finite region (Weinberg §3.1; MTW §1.6).
- In such a frame, we can apply the laws of special relativity.
- Physics in the "lab frame" can be obtained with a coordinate transformation.

A mathematical analogy

Geometry is simple locally.

- The Earth is approximately flat . . . in small regions.
- But its true geometry is spherical.

Literature

The course closely follows the book Gravitation and Cosmology by S. Weinberg.

In addition, it is partially based on material from:

- L. D. Landau and E. M. Lifshitz, The Classical Theory of Fields;
- R. P. Feynman, Six Not-So-Easy Pieces;
- C. W. Misner, K. S. Thorne and J. A. Wheeler, Gravitation (MTW);
- S. M. Carroll, Spacetime and Geometry: An Introduction to General Relativity.

1

Special relativity

Conventions

The following conventions are used throughout the script. They are those of Weinberg's Gravitation and Cosmology except for the signature of the metric.

- Signature $(+,-,-,-)$: $\eta_{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)$, $\mathrm{d}\tau^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$, so that $u^\mu u_\mu = g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} = 1$ and $p^\mu p_\mu = m^2$. Units $c = 1$ from Eq. (1.11) on; $x^0 = t$. Greek indices run over $0,1,2,3$, Latin indices over $1,2,3$. Three-vectors carry Latin indices with Euclidean summation ($v_i = v^i$); when a spatial index of a four-tensor is raised, this is done with η (e.g. $h^i{}_j = -h_{ij}$).

- $g \equiv -\det(g_{\mu\nu}) > 0$; the invariant volume element is $\sqrt{g}\,\mathrm{d}^4x$ (other books write $\sqrt{-g}$ for the same quantity). Levi-Civita symbol: $\varepsilon^{0123} = +1$, hence $\varepsilon_{0123} = -1$; the tensors are $\varepsilon^{\mu\nu\rho\sigma}/\sqrt{g}$ and $\sqrt{g}\,\varepsilon_{\mu\nu\rho\sigma}$.

- Affine connection: $\Gamma^\lambda{}_{\mu\nu} = \frac{1}{2}g^{\lambda\kappa}\left(\partial_\mu g_{\kappa\nu} + \partial_\nu g_{\kappa\mu} - \partial_\kappa g_{\mu\nu}\right)$; geodesics $\frac{\mathrm{d}^2x^\lambda}{\mathrm{d}\tau^2} + \Gamma^\lambda{}_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} = 0$; covariant derivative $A_{\mu;\nu} = \partial_\nu A_\mu - \Gamma^\lambda{}_{\mu\nu}A_\lambda$.

- Riemann tensor (Weinberg's sign): $R^\lambda{}_{\mu\nu\kappa} = \partial_\kappa\Gamma^\lambda{}_{\mu\nu} - \partial_\nu\Gamma^\lambda{}_{\mu\kappa} + \Gamma^\beta{}_{\mu\nu}\Gamma^\lambda{}_{\beta\kappa} - \Gamma^\beta{}_{\mu\kappa}\Gamma^\lambda{}_{\beta\nu}$, with $V^\lambda{}_{;\nu;\kappa} - V^\lambda{}_{;\kappa;\nu} = R^\lambda{}_{\mu\nu\kappa}V^\mu$ and $A_{\mu;\nu;\kappa} - A_{\mu;\kappa;\nu} = -R^\lambda{}_{\mu\nu\kappa}A_\lambda$; Ricci tensor $R_{\mu\nu} = R^\lambda{}_{\mu\lambda\nu}$; curvature scalar $R = g^{\mu\nu}R_{\mu\nu}$. With this sign a two-sphere has $R < 0$.

- Field equations:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = -8\pi G\,T_{\mu\nu}, \quad \text{or, equivalently,}$$
$$R_{\mu\nu} = -8\pi G\left(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T^\lambda{}_\lambda\right) + \Lambda g_{\mu\nu}.$$

  Newtonian limit: $g_{00} \approx 1 + 2\phi$, $\nabla^2\phi = 4\pi G\rho$; $\Lambda > 0$ is a positive vacuum energy density $\Lambda/(8\pi G)$.

- Action: $S = \int \mathrm{d}^4x\sqrt{g}\left[\frac{R-2\Lambda}{16\pi G} + \mathcal{L}_M\right]$, and the energy-momentum tensor is defined by $\delta S_M = -\frac{1}{2}\int \mathrm{d}^4x\sqrt{g}\,T^{\mu\nu}\delta g_{\mu\nu}$ (so that $T^{00} > 0$).

- Weak fields: $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, $\Box = \eta^{\mu\nu}\partial_\mu\partial_\nu = \partial_t^2 - \nabla^2$; in harmonic gauge $\Box h_{\mu\nu} = -16\pi G S_{\mu\nu}$ with $S_{\mu\nu} = T_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}T^\lambda{}_\lambda$, and the retarded solution is $h_{\mu\nu}(\vec{x},t) = -4G\int \mathrm{d}^3x'\, S_{\mu\nu}(\vec{x}\,', t - |\vec{x} - \vec{x}\,'|)/|\vec{x} - \vec{x}\,'|$ (Newtonian check: $h_{00} = 2\phi$).
- Geodesic deviation: $\frac{\mathrm{D}^2\xi^\lambda}{\mathrm{D}\tau^2} = R^\lambda{}_{\nu\sigma\rho}\, u^\nu \xi^\sigma u^\rho$; in the Newtonian limit $R^i{}_{0j0} = -\partial_i\partial_j\phi$.

Notation. $r_S \equiv 2GM$ (Schwarzschild radius); both r_S and $2GM$ appear. Λ without indices is the cosmological constant; $\Lambda^\mu{}_\nu$ is a Lorentz transformation. κ denotes the curvature constant of a maximally symmetric space (Chapter 9; $\kappa = 0, \pm 1$ in the Robertson-Walker metric, whose scale factor $R(t)$ should not be confused with the curvature scalar R); in the derivation of the Schwarzschild metric κ is an integration constant (set to 1). The gravitational coupling is always written as $8\pi G$. ϕ is the Newtonian potential and φ the azimuthal angle. Electrodynamics: $E_i = F^{0i}$, $B_k = \frac{1}{2}\varepsilon_{ijk}F^{ij}$, $F^{\mu\nu}{}_{;\mu} = -J^\nu$.

Comparison with the literature.

- Weinberg uses $\eta = \mathrm{diag}(-1,1,1,1)$ with $\mathrm{d}\tau^2 = -g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$, and the same Γ, Riemann tensor, Ricci contraction and $g = -\det g_{\mu\nu}$ as here. Hence $R_{\mu\nu}$, $R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$ and $-8\pi G T_{\mu\nu}$ agree, while every explicit η or g, every lowered index, and the Λ term change sign: Weinberg's field equation reads $R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R - \Lambda g_{\mu\nu} = -8\pi G T_{\mu\nu}$, his $\Box$ is $\nabla^2 - \partial_t^2$, and his retarded solution is $h_{\mu\nu} = +4G\int S_{\mu\nu}/|\vec{x} - \vec{x}\,'|$.
- Misner–Thorne–Wheeler and Carroll (also Wald) use $(-,+,+,+)$ and the opposite sign of the Riemann tensor, $R^{\mathrm{MTW}\lambda}{}_{\mu\nu\kappa} = -R^\lambda{}_{\mu\nu\kappa}$, hence also of $R_{\mu\nu}$; their curvature scalar of a given spacetime equals ours (both signs flip). Their equations read $R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = +8\pi G T_{\mu\nu}$, $u^\mu u_\mu = -1$, $T_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta S_M}{\delta g^{\mu\nu}}$, and the geodesic deviation equation carries a minus sign, $\frac{\mathrm{D}^2\xi^\lambda}{\mathrm{D}\tau^2} = -R^\lambda{}_{\nu\sigma\rho}u^\nu\xi^\sigma u^\rho$ (their Riemann). The Einstein-Hilbert action has the same form, $\frac{1}{16\pi G}\int\sqrt{-g}\,(R - 2\Lambda)$.

In special relativity we postulate that all inertial observers, say O and O', measure the same speed of light c. Consider a light signal which travels from the spacetime point $(\vec{x},t)$ to the infinitesimally close point $(\vec{x} + \mathrm{d}\vec{x}, t + \mathrm{d}t)$, as sketched in Fig. 1.1.

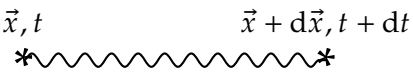


Figure 1.1. A light signal travelling between two infinitesimally close spacetime points.

Both observers measure the same speed of light,

$$c^2 = \frac{\mathrm{d}\vec{x}^{\,2}}{\mathrm{d}t^2} = \frac{\mathrm{d}\vec{x}'^{\,2}}{\mathrm{d}t'^2}. \tag{1.1}$$

Hence

$$c^2\mathrm{d}t^2 - \mathrm{d}\vec{x}^{\,2} = c^2\mathrm{d}t'^2 - \mathrm{d}\vec{x}'^{\,2} = 0, \tag{1.2}$$

if the points are infinitesimally separated by a light ray. For a general infinitesimal separation we define the proper time interval $\mathrm{d}\tau$ by

$$c^2\mathrm{d}\tau^2 \equiv c^2\mathrm{d}t^2 - \mathrm{d}\vec{x}^2. \tag{1.3}$$

For two events connected by a light signal, $\mathrm{d}\tau^2 = 0$ in every inertial frame. If we assume that spacetime is homogeneous and isotropic, it follows that $\mathrm{d}\tau^2$ between any two events is the same in all inertial frames. To show this, we write

$$\mathrm{d}\tau^2 = \mathrm{d}\tau'^2\lambda(\cdots), \tag{1.4}$$

where the dots represent the dependence on the spacetime point and on the velocity of the factor λ introduced by the transformation between inertial frames. Intuitively, the dependence on the spacetime point is eliminated by the homogeneity of spacetime and the dependence on the velocity by its isotropy. To see the independence of the velocity more explicitly, consider transformations between three frames

$$O \xrightarrow{\vec{v}_1} O' \xrightarrow{\vec{v}_{21}} O''. \qquad O \xrightarrow{\vec{v}_2} O'' \tag{1.5}$$

The proper time intervals are related by

$$\mathrm{d}\tau^2 = \lambda(\vec{v}_1)\mathrm{d}\tau'^2 = \lambda(\vec{v}_2)\mathrm{d}\tau''^2, \quad \mathrm{d}\tau'^2 = \lambda(\vec{v}_{21})\mathrm{d}\tau''^2, \tag{1.6}$$

and thus

$$\lambda(\vec{v}_{21}) = \frac{\lambda(\vec{v}_2)}{\lambda(\vec{v}_1)}. \tag{1.7}$$

Because of isotropy, λ can only depend on the magnitudes of the velocities,

$$\lambda(|\vec{v}_{21}|) = \frac{\lambda(|\vec{v}_2|)}{\lambda(|\vec{v}_1|)}. \tag{1.8}$$

If $\vec{v}_{21} \approx \vec{v}_2 - \vec{v}_1$, the magnitudes of the velocities are related by

$$|\vec{v}_{21}|^2 = |\vec{v}_1|^2 + |\vec{v}_2|^2 - 2|\vec{v}_1||\vec{v}_2|\cos\theta. \tag{1.9}$$

The left-hand side depends on the angle θ between $\vec{v}_1$ and $\vec{v}_2$, while the right-hand side depends only on $|\vec{v}_1|$ and $|\vec{v}_2|$. Hence λ has to be a constant. **Remark.** The approximation $\vec{v}_{21} \approx \vec{v}_2 - \vec{v}_1$ is not needed: the exact relativistic $|\vec{v}_{21}|$ also depends on the angle between $\vec{v}_1$ and $\vec{v}_2$ (Landau–Lifshitz, Vol. 2, §2). Then (1.7) gives $\lambda = \lambda/\lambda = 1$. This means that the proper time intervals

$$\boxed{\mathrm{d}\tau^2 = \mathrm{d}\tau'^2} \tag{1.10}$$

are preserved, with $c^2\mathrm{d}\tau^2 = c^2\mathrm{d}\tau'^2 = c^2\mathrm{d}t^2 - \mathrm{d}\vec{x}^2$. From now on we will use units such that $c = 1$. We can then write

$$\mathrm{d}\tau^2 = \mathrm{d}t^2 - \mathrm{d}\vec{x}^2 = \eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu, \quad \eta_{\mu\nu} = \mathrm{diag}(1,-1,-1,-1), \tag{1.11}$$

where $\eta_{\mu\nu}$ is called the Minkowski metric and we employ the Einstein summation convention. **Remark.** We use the signature $(+,-,-,-)$ with $\mathrm{d}\tau^2 = +\eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu$. Weinberg's book uses $\eta = \mathrm{diag}(-1,1,1,1)$ with $\mathrm{d}\tau^2 = -\eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu$ (Weinberg eq. 2.1.4 ff.); formulas containing an explicit η, lowered indices, or $\varepsilon_{\alpha\beta\gamma\delta}$ therefore differ in sign from the book.

Combining (1.10) and (1.11) yields

$$\eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu = \eta_{\mu\nu}\mathrm{d}x'^\mu\mathrm{d}x'^\nu = \eta_{\mu\nu}\frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x'^\nu}{\partial x^\sigma}\mathrm{d}x^\rho\mathrm{d}x^\sigma. \tag{1.12}$$

Therefore

$$\boxed{\eta_{\rho\sigma} = \frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x'^\nu}{\partial x^\sigma}\eta_{\mu\nu}} \tag{1.13}$$

This implies that the (Lorentz) transformation $x \longrightarrow x'$ is a linear transformation. To show this, we take the derivative of the above equation

$$\begin{aligned} 0 = \frac{\partial}{\partial x^\alpha}\eta_{\rho\sigma} &= \frac{\partial}{\partial x^\alpha}\left(\eta_{\mu\nu}\frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x'^\nu}{\partial x^\sigma}\right) \\ &= \eta_{\mu\nu}\frac{\partial^2 x'^\mu}{\partial x^\alpha \partial x^\rho}\frac{\partial x'^\nu}{\partial x^\sigma} + \eta_{\mu\nu}\frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial^2 x'^\nu}{\partial x^\alpha \partial x^\sigma} \\ &\equiv \langle\alpha\rho|\sigma\rangle + \langle\alpha\sigma|\rho\rangle. \end{aligned} \tag{1.14}$$

Adding up three permutations of this relation,

$$\left.\begin{aligned} \cancel{\langle\alpha\rho|\sigma\rangle} + \langle\alpha\sigma|\rho\rangle &= 0 \\ \cancel{\langle\sigma\rho|\alpha\rangle} + \langle\alpha\sigma|\rho\rangle &= 0 \\ -\left(\cancel{\langle\alpha\rho|\sigma\rangle} + \cancel{\langle\rho\sigma|\alpha\rangle}\right) &= 0 \end{aligned}\right\} \Rightarrow 2\langle\alpha\sigma|\rho\rangle = 0, \tag{1.15}$$

we find

$$\langle\alpha\sigma|\rho\rangle = \eta_{\mu\nu}\frac{\partial^2 x'^\nu}{\partial x^\alpha \partial x^\sigma}\frac{\partial x'^\mu}{\partial x^\rho} = 0. \tag{1.16}$$

We can remove the metric by multiplying with the inverse metric, and the Jacobian of the transformation by multiplying with its inverse $\frac{\partial x}{\partial x'}$. The inverse exists, because physically a transformation between inertial frames must be invertible. Thus, the second derivatives of the transformation must vanish,

$$\boxed{\frac{\partial^2 x'^\mu}{\partial x^\rho \partial x^\sigma} = 0} \tag{1.17}$$

or, in other words, the transformation must be linear,

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu + \rho^\mu. \tag{1.18}$$

It is called the Poincaré transformation or inhomogeneous Lorentz transformation. We can now rewrite (1.13) as

$$\Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma\eta_{\mu\nu} = \eta_{\rho\sigma}. \tag{1.19}$$

In matrix form, the above equation gives

$$\begin{aligned} \Lambda^T\eta\Lambda = \eta \quad \Rightarrow \quad \det\left(\Lambda^T\eta\Lambda\right) &= \det(\eta) \\ (\det\Lambda)^2\cancel{\det\eta} &= \cancel{\det\eta} \\ \det\Lambda &= \pm 1. \end{aligned} \tag{1.20}$$

Also, for $\rho = \sigma = 0$ (1.19) becomes

$$\begin{aligned} &\Lambda^\mu{}_0\Lambda^\nu{}_0\eta_{\mu\nu} = 1 \\ \rightarrow \ &\left(\Lambda^0{}_0\right)^2 - \sum_{i=1}^{3}\left(\Lambda^i{}_0\right)^2 = 1 \ \Rightarrow \left(\Lambda^0{}_0\right)^2 \geq 1. \end{aligned} \tag{1.21}$$

Thus, from (1.19) we find that

- $\det \Lambda = \pm 1$,
- ${\Lambda^0}_0 \geq +1$ or ${\Lambda^0}_0 \leq -1$.

There is a subgroup of Lorentz transformations with $\det \Lambda = 1$, ${\Lambda^0}_0 \geq +1$ which is called the group of proper Lorentz transformations. It contains the identity transformation ${\Lambda^\mu}_\nu = {\delta^\mu}_\nu$. This is the group of physical transformations, and from now on we will only consider these.

Let us assume that an observer O sees a particle at rest and that another inertial observer O' sees the same particle moving with velocity $\vec{v}$ (Fig. 1.2).

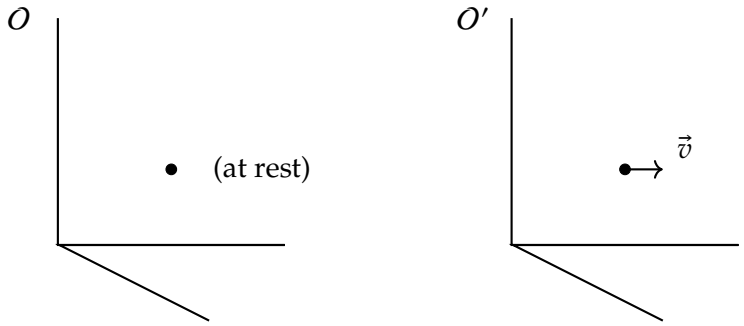


Figure 1.2. A particle at rest in the frame O moves with velocity $\vec{v}$ in the frame O'.

The two frames are related by a coordinate transformation

$$x'^\mu = {\Lambda^\mu}_\nu x^\nu + \rho^\mu. \tag{1.22}$$

Let us observe the particle in the frame O at two spacetime points x^μ and $x^\mu + \mathrm{d}x^\mu$. We have

$$\begin{aligned} \mathrm{d}x^0 &= \mathrm{d}t \neq 0, \\ \mathrm{d}x^1 &= \mathrm{d}x^2 = \mathrm{d}x^3 = 0. \end{aligned} \tag{1.23}$$

Analogously, in the frame O', we have

$$\begin{aligned} \mathrm{d}x'^0 &= {\Lambda^0}_\nu \mathrm{d}x^\nu = {\Lambda^0}_0 \mathrm{d}x^0 = {\Lambda^0}_0 \mathrm{d}t \\ \mathrm{d}x'^i &= {\Lambda^i}_0 \mathrm{d}x^0 + {\Lambda^i}_j \underbrace{\mathrm{d}x^j}_{=0} = {\Lambda^i}_0 \mathrm{d}t. \end{aligned} \tag{1.24}$$

Then

$$\frac{\mathrm{d}x'^i}{\mathrm{d}x'^0} \equiv v_i = \frac{{\Lambda^i}_0}{{\Lambda^0}_0}, \tag{1.25}$$

where $\vec{v}$ is the velocity of the particle. The proper time intervals in the two frames are the

same,

$$\begin{aligned} \mathrm{d}\tau^2 = \mathrm{d}t^2 &= \mathrm{d}t'^2 - \mathrm{d}\vec{x}'^2 \\ &= \left(\Lambda^0{}_0\right)^2 \mathrm{d}t^2 - \left(\Lambda^i{}_0\right)^2 \mathrm{d}t^2 \\ &= \mathrm{d}t^2 \left[\left(\Lambda^0{}_0\right)^2 - \left(\Lambda^i{}_0\right)^2\right] \\ \Rightarrow 1 &= \left(\Lambda^0{}_0\right)^2 - \left(\Lambda^i{}_0\right)^2 \\ &= \left(\Lambda^0{}_0\right)^2 - (v_i)^2 \left(\Lambda^0{}_0\right)^2 \\ \Rightarrow \left(\Lambda^0{}_0\right)^2 &= \frac{1}{1-\vec{v}^2}, \end{aligned} \tag{1.26}$$

where a sum over i is implied. We thus define

$$\gamma = \gamma(\vec{v}) \equiv \Lambda^0{}_0 = \frac{1}{\sqrt{1-\vec{v}^2}}. \tag{1.27}$$

Returning to (1.25) yields

$$\Lambda^i{}_0 = \gamma v_i = \frac{v_i}{\sqrt{1-\vec{v}^2}}. \tag{1.28}$$

The remaining elements of $\Lambda^\mu{}_\nu$ cannot be determined uniquely. The Lorentz transformations contain, for example, the subgroup of rotations,

$$\Lambda^\mu{}_\nu = \begin{cases} \mu = \nu = 0: & 1, \\ \mu = 0 \textbf{ or } \nu = 0: & 0, \\ \mu = i,\ \nu = j: & R_{ij}, \end{cases} \tag{1.29}$$

with $\det R = 1$ and $R^T R = \mathbb{1}_{3\times 3}$.

For coordinate systems with parallel axes, we have

$$\Lambda(\vec{v}) = \left(\begin{array}{c|c} \gamma & \gamma v_i \\ \hline \gamma v_i & \delta_{ij} + \dfrac{v_i v_j}{\vec{v}^2}(\gamma - 1) \end{array}\right). \tag{1.30}$$

This is called a boost; it involves a relative velocity but no rotation.

Exercise: Check that $\Lambda(\vec{v})$ satisfies (1.19).

1.1 Time dilation and Doppler effect

One consequence of the preservation of proper time in special relativity is that time flows more slowly for moving objects. Consider again a body at rest in O, which is also a clock with period Δt. The clock moves with velocity $\vec{v}$ in the frame O'. We have

$$\mathrm{d}t' = \Lambda^0{}_0 \Delta t = \gamma(\vec{v})\Delta t, \tag{1.31}$$

or

$$\boxed{\mathrm{d}t' = \frac{\Delta t}{\sqrt{1-\vec{v}^2}} \geq \Delta t} \tag{1.32}$$

Doppler effect: Suppose now that our "clock" is a source of light with frequency $\nu = \frac{1}{\Delta t}$. In the frame O the time interval between two crests of the wave is Δt (Fig. 1.3).

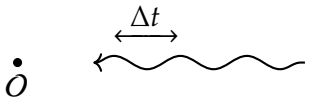


Figure 1.3. In the frame O the source is at rest, and successive wave crests are emitted at time intervals Δt.

In the frame O' the light source moves with velocity $\vec{v}$, whose component along the line of sight we denote by v_r (Fig. 1.4).

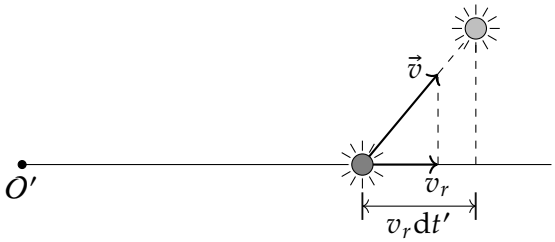


Figure 1.4. In the frame O' the source moves with velocity $\vec{v}$; between the emission of two crests, its distance from the observer grows by $v_r \mathrm{d}t'$.

The time between the arrival of two crests at the observer is

$$\begin{aligned} \Delta t_0' &\equiv \mathrm{d}t' + v_r \mathrm{d}t' = \gamma \mathrm{d}t(1 + v_r) \\ \Rightarrow \Delta t_0' &= \frac{1 + v_r}{\sqrt{1 - \vec{v}^2}} \Delta t, \end{aligned} \tag{1.33}$$

and the new frequency is

$$\nu' = \frac{1}{\Delta t_0'} = \frac{\sqrt{1 - \vec{v}^2}}{1 + v_r} \nu. \tag{1.34}$$

For a source moving towards the observer along the line of sight, we have

$$v_r = -v, \tag{1.35}$$

and

$$\nu' = \sqrt{\frac{1 + v}{1 - v}} \nu \geq \nu. \tag{1.36}$$

For visible light, this is a shift towards the blue end of the spectrum (a blueshift). If the light source moves away from the observer radially, we have

$$v_r = v, \tag{1.37}$$

and

$$\nu' = \sqrt{\frac{1 - v}{1 + v}} \nu \leq \nu, \tag{1.38}$$

which is a shift towards the infrared (a redshift).

1.2 Dynamics in special relativity

Consider an inertial observer, who observes a particle with velocity $\vec{v}$ to be subjected to a force. The velocity is an instantaneous one: it changes due to the acceleration caused by the force.

For $\vec{v} = 0$ the observer knows that they can apply the laws of classical mechanics. So, to calculate what happens to the particle for general $\vec{v}$, the observer can do the following trick (in principle):

boost to a frame where $\vec{v} = 0$ at time t, calculate the acceleration and the new velocity after a time $\mathrm{d}t$ in this instantaneous rest frame, boost back to the original frame with a Lorentz transformation, and start again. But all of this is overkill. We can directly generalise the equations of classical mechanics so that they are consistent with Lorentz transformations. We introduce a relativistic force

$$f^\mu = m\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2}, \tag{1.39}$$

where m is the mass of a particle and τ is proper time. Let us see how f^μ is related to the Newtonian force. We first work in the instantaneous rest frame of the particle ($\vec{v} = 0$). Then

$$f^0 = m\frac{\mathrm{d}^2 x^0}{\mathrm{d}\tau^2} = m\frac{\mathrm{d}^2 x^0}{\mathrm{d}x^{0\,2}} = 0, \tag{1.40}$$

and

$$f^i = m\frac{\mathrm{d}^2 x^i}{\mathrm{d}\tau^2} = m\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} = F^i_{\text{Newton}}. \tag{1.41}$$

Let us now look at what happens in the frame where the particle has an instantaneous velocity $\vec{v}$. Then

$$f^\mu = \underbrace{\Lambda^\mu{}_\nu(\vec{v})}_{\text{boost}}\, f^\nu\big|_{\text{rest}}, \tag{1.42}$$

and

$$\begin{aligned} f^0 &= \Lambda^0{}_\nu(\vec{v})\, f^\nu\big|_{\text{rest}} \\ &= \Lambda^0{}_0 \underbrace{f^0\big|_{\text{rest}}}_{=0} + \Lambda^0{}_i\, f^i\big|_{\text{rest}} \\ &= \gamma v_i F^i_{\text{Newton}}. \\ \Rightarrow \boxed{f^0} &= \boxed{\frac{\vec{v}\cdot\vec{F}_{\text{Newton}}}{\sqrt{1-\vec{v}^{\,2}}}} \end{aligned} \tag{1.43}$$

For the spatial components we have

$$\begin{aligned} f^i &= \Lambda^i{}_\nu(\vec{v})\, f^\nu\big|_{\text{rest}} \\ &= \Lambda^i{}_0 \underbrace{f^0\big|_{\text{rest}}}_{=0} + \Lambda^i{}_j\, f^j\big|_{\text{rest}} \\ &= \left(\delta_{ij} + \frac{v_i v_j}{v^2}(\gamma - 1)\right) F^j_{\text{Newton}}. \\ \Rightarrow \boxed{\underbrace{\vec{f}}_{rel.}} &= \boxed{\vec{F}_{\text{Newton}} + \frac{\vec{v}\cdot\vec{F}_{\text{Newton}}}{\vec{v}^{\,2}}(\gamma - 1)\vec{v}} \end{aligned} \tag{1.44}$$

Exercise: Derive the Lorentz force starting from the electrostatic force.

We now know how to use the Newtonian limit (zero velocity) to construct the relativistic force in (1.39). Equation (1.39) is a second-order differential equation (or rather, four of them, for $\mu = 0, 1, 2, 3$). To solve it, we also need boundary conditions, but here we must be careful! In Newtonian physics,

$$\vec{F} = m\frac{\mathrm{d}^2\vec{x}}{\mathrm{d}t^2}. \tag{1.45}$$

We need $\vec{x}(0)$ and $\dot{\vec{x}}(0)$, i.e. six boundary conditions. For special relativity we may be tempted to require eight boundary conditions $x^\mu(\tau_0)$ and $\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}(\tau_0)$. However, the relativistic "velocities" $\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}$ must satisfy

$$\begin{aligned} \eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} &= 1 \\ \left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 - \left(\frac{\mathrm{d}\vec{x}}{\mathrm{d}\tau}\right)^2 &= 1, \end{aligned} \tag{1.46}$$

for every value of τ. So, when we choose boundary conditions at $\tau = \tau_0$, we should respect the above identity. Once we have chosen consistent boundary conditions at $\tau = \tau_0$, we can rest assured that (1.46) will be satisfied for all values of τ. Indeed,

$$\begin{aligned} m\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\right) &= 2\eta_{\mu\nu}\left(m\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2}\right)\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \\ &= 2\eta_{\mu\nu}f^\mu\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \\ &\equiv K(x). \end{aligned} \tag{1.47}$$

Notice that $K(x)$ is invariant under Lorentz transformations. Indeed,

$$\begin{aligned} K'(x) &= 2\eta_{\mu\nu}f'^\mu\frac{\mathrm{d}x'^\nu}{\mathrm{d}\tau} \\ &= 2\eta_{\mu\nu}\Lambda^\mu{}_\rho f^\rho\Lambda^\nu{}_\sigma\frac{\mathrm{d}x^\sigma}{\mathrm{d}\tau} \\ &= 2\underbrace{\eta_{\mu\nu}\Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma}_{=\eta_{\rho\sigma}}f^\rho\frac{\mathrm{d}x^\sigma}{\mathrm{d}\tau} \\ &= 2\eta_{\rho\sigma}f^\rho\frac{\mathrm{d}x^\sigma}{\mathrm{d}\tau} = K(x). \end{aligned} \tag{1.48}$$

Let us compute it in the rest frame; by the above result, it is then the same in all inertial frames:

$$f^\mu\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\eta_{\mu\nu}\Big|_{\text{rest}} = \underbrace{f^0}_{=0}\frac{\mathrm{d}x^0}{\mathrm{d}\tau} - f^i\underbrace{\frac{\mathrm{d}x^i}{\mathrm{d}\tau}}_{=0} = 0. \tag{1.49}$$

Thus,

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\right) = 0, \tag{1.50}$$

and $\eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}$ is independent of proper time.

1.3 Relativistic energy and momentum

Let us now define

$$p^\mu = m\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}, \tag{1.51}$$

which seems like an intuitive generalisation of $\vec{p} = m\frac{\mathrm{d}\vec{x}}{\mathrm{d}t}$ from classical mechanics. Under a Lorentz transformation,

$$p'^\mu = \Lambda^\mu{}_\nu p^\nu. \tag{1.52}$$

In the rest frame,

$$p^\mu_{\text{rest}} = m\frac{\mathrm{d}x^\mu}{\mathrm{d}t} \Rightarrow \begin{cases} p^0 = m, \\ p^i = 0. \end{cases} \tag{1.53}$$

In the frame where the particle has a velocity $\vec{v}$, we have

$$\begin{aligned} p^\mu &= \Lambda^\mu{}_\nu(\vec{v}) p^\nu_{\text{rest}} \\ \rightarrow &\boxed{p^0 = m\gamma} \\ \rightarrow &\boxed{p^i = m\gamma v_i} \end{aligned} \tag{1.54}$$

Let us analyse the time component,

$$p^0 = m\gamma = \frac{m}{\sqrt{1-\vec{v}^2}} = m + \overbrace{\frac{1}{2}m\vec{v}^2}^{\text{kinetic energy of a free particle}} + \mathcal{O}(\vec{v}^4). \tag{1.55}$$

We will identify p^0 with the energy of the particle. We have

$$\begin{aligned} E = p^0 = m\gamma, \\ \text{and } \vec{p} = m\gamma\vec{v}. \end{aligned} \tag{1.56}$$

We also have

$$E^2 - \vec{p}^2 = m^2\gamma^2(1-\vec{v}^2) = m^2, \tag{1.57}$$

or

$$\boxed{E = \sqrt{m^2 + \vec{p}^2}} \tag{1.58}$$

Finally, let us recall that

$$f^\mu = m\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} = \frac{\mathrm{d}p^\mu}{\mathrm{d}\tau}. \tag{1.59}$$

Remark. For us the mass m is always a constant physical parameter. We will not combine $m\gamma = \frac{m}{\sqrt{1-\vec{v}^2}}$ into a "relativistic mass" ~~$m(\vec{v}) = \frac{m}{\sqrt{1-\vec{v}^2}}$~~. While this can give us a nice qualitative picture to describe various phenomena with words, it can also cause confusion.

1.4 Vectors and tensors in special relativity

• **Contravariant vectors:**

Spacetime intervals transform as

$$\mathrm{d}x^\mu \rightarrow \mathrm{d}x'^\mu = \Lambda^\mu{}_\nu \mathrm{d}x^\nu. \tag{1.60}$$

Definition: Every four-component object which transforms in the same way as spacetime intervals,

$$V^\mu \to V'^\mu = \Lambda^\mu{}_\nu V^\nu, \tag{1.61}$$

is called a contravariant vector. Examples of contravariant vectors are the four-momentum and the relativistic force

$$p^\mu = m\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}, \quad f^\mu = \frac{\mathrm{d}p^\mu}{\mathrm{d}\tau}, \tag{1.62}$$

which transform as in (1.61).

• **Covariant vectors:**

The spacetime derivative, under a Lorentz transformation, changes as

$$\frac{\partial}{\partial x^\mu} \to \frac{\partial}{\partial x'^\mu} = \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial}{\partial x^\rho} = \Lambda_\mu{}^\rho \frac{\partial}{\partial x^\rho}, \tag{1.63}$$

where we recognised that $\frac{\partial x^\rho}{\partial x'^\mu} = \Lambda_\mu{}^\rho$, the inverse of a Lorentz transformation. **Remark.** Here $\Lambda_\mu{}^\rho \equiv \eta_{\mu\nu}\eta^{\rho\sigma}\Lambda^\nu{}_\sigma$ (indices moved with η, as used below). That this equals $(\Lambda^{-1})^\rho{}_\mu$ is not automatic: it follows from (1.19), $\Lambda^T\eta\Lambda = \eta \Rightarrow \Lambda^{-1} = \eta^{-1}\Lambda^T\eta$, and only then is the step $\Lambda^\mu{}_\rho\Lambda_\mu{}^\sigma = \delta^\sigma{}_\rho$ below justified (Weinberg §2.1). Indeed, we have

$$\begin{aligned} &\frac{\partial x'^\mu}{\partial x'^\nu} = \delta^\mu{}_\nu \;\to\; \frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x^\rho}{\partial x'^\nu} = \delta^\mu{}_\nu \\ \to\; &\Lambda^\mu{}_\rho\frac{\partial x^\rho}{\partial x'^\nu} = \delta^\mu{}_\nu \;\to\; \Lambda^\mu{}_\rho\Lambda_\mu{}^\sigma\frac{\partial x^\rho}{\partial x'^\nu} = \delta^\mu{}_\nu\Lambda_\mu{}^\sigma \\ \to\; &\delta^\sigma{}_\rho\frac{\partial x^\rho}{\partial x'^\nu} = \Lambda_\nu{}^\sigma \to \boxed{\frac{\partial x^\sigma}{\partial x'^\nu} = \Lambda_\nu{}^\sigma}. \end{aligned} \tag{1.64}$$

So, the spacetime derivative transforms with the inverse of a Lorentz transformation. Any vector which transforms with the inverse of a Lorentz transformation is called a covariant vector,

$$U_\mu \to U'_\mu = \Lambda_\mu{}^\nu U_\nu. \tag{1.65}$$

Duality of contravariant and covariant vectors

Consider a contravariant vector V^μ, which obeys (1.61). Multiplying with the metric, we construct the object

$$V_\mu \equiv \eta_{\mu\nu}V^\nu, \tag{1.66}$$

V_μ is a covariant vector. Indeed,

$$\begin{aligned} V'_\mu &= \eta_{\mu\nu}V'^\nu = \eta_{\mu\nu}\Lambda^\nu{}_\rho V^\rho \\ &= \Lambda^\nu{}_\rho\eta_{\mu\nu}\delta^\rho{}_\sigma V^\sigma \\ &= \left(\Lambda^\nu{}_\rho\eta_{\mu\nu}\eta^{\rho\kappa}\right)\left(\eta_{\kappa\sigma}V^\sigma\right) \\ &= \Lambda_\mu{}^\kappa V_\kappa. \end{aligned} \tag{1.67}$$

Analogously, we can construct a contravariant vector out of a covariant one, using the inverse metric. Since the metric is known, the dual of a vector carries the same information as the original vector.

• **Scalars:**

There are quantities which do not transform under Lorentz transformations. These quantities are called scalars. For example, the contraction

$$A \cdot B \equiv A^\mu B_\mu, \tag{1.68}$$

of a contravariant and a covariant vector is invariant (scalar) under Lorentz transformations. Indeed,

$$\begin{aligned} A^\mu B_\mu \to A'^\mu B'_\mu &= \left(\Lambda^\mu{}_\nu A^\nu\right)\left(\Lambda_\mu{}^\rho B_\rho\right) \\ &= \left(\Lambda^\mu{}_\nu \Lambda_\mu{}^\rho\right) A^\nu B_\rho = \delta^\rho{}_\nu A^\nu B_\rho \\ &= A^\rho B_\rho. \end{aligned} \tag{1.69}$$

• **Tensors:**

We can define objects with multiple indices, which we call tensors,

$$T^{\mu_1\mu_2\cdots}{}_{\nu_1\nu_2\cdots}. \tag{1.70}$$

Products of vectors are tensors, and every tensor transforms like such a product under a Lorentz transformation,

$$T'^{\mu_1\mu_2\cdots}{}_{\nu_1\nu_2\cdots} = \Lambda^{\mu_1}{}_{\alpha_1}\Lambda^{\mu_2}{}_{\alpha_2}\cdots\Lambda_{\nu_1}{}^{\beta_1}\Lambda_{\nu_2}{}^{\beta_2}\cdots T^{\alpha_1\alpha_2\cdots}{}_{\beta_1\beta_2\cdots}. \tag{1.71}$$

For example, $T^{\mu\nu} = A^\mu B^\nu$ transforms as

$$\begin{aligned} T'^{\mu\nu} = A'^\mu B'^\nu &= \left(\Lambda^\mu{}_\rho A^\rho\right)\left(\Lambda^\nu{}_\sigma B^\sigma\right) \\ &= \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma A^\rho B^\sigma \\ &= \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma T^{\rho\sigma}. \end{aligned} \tag{1.72}$$

However, not all tensors are products of vectors.

Metric $\eta_{\mu\nu}$: This is a covariant tensor, as can be seen from the definition of the Lorentz transformation,

$$\eta'_{\rho\sigma} = \Lambda_\rho{}^\mu \Lambda_\sigma{}^\nu \eta_{\mu\nu} = \eta_{\rho\sigma}. \tag{1.73}$$

Kronecker delta: This is a tensor with one contravariant and one covariant index. Indeed,

$$\begin{aligned} \delta^\mu{}_\nu \to \delta'^\mu{}_\nu &= \Lambda^\mu{}_\rho \Lambda_\nu{}^\sigma \delta^\rho{}_\sigma \\ &= \underbrace{\Lambda^\mu{}_\rho \Lambda_\nu{}^\rho}_{\Lambda\cdot\Lambda^{-1}} = \underbrace{\delta^\mu{}_\nu}_{\mathbb{1}}. \end{aligned} \tag{1.74}$$

Levi-Civita tensor: This is defined as a fully antisymmetric tensor with

$$\varepsilon^{\alpha\beta\gamma\delta} = \begin{cases} +1, & \alpha\beta\gamma\delta \text{ even permutation of } 0123, \\ -1, & \alpha\beta\gamma\delta \text{ odd permutation of } 0123, \\ 0, & \text{otherwise.} \end{cases} \tag{1.75}$$

Let us show that the Levi-Civita symbol is a tensor under proper Lorentz transformations. We transform it as a tensor,

$$\varepsilon^{\alpha\beta\gamma\delta} \to \varepsilon'^{\alpha\beta\gamma\delta} = \Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu \Lambda^\gamma{}_\rho \Lambda^\delta{}_\sigma \varepsilon^{\mu\nu\rho\sigma}. \tag{1.76}$$

The right-hand side is fully antisymmetric in the indices $\alpha\beta\gamma\delta$. Let us calculate it explicitly for $\alpha\beta\gamma\delta = 0123$,

$$\Lambda^0{}_\mu \Lambda^1{}_\nu \Lambda^2{}_\rho \Lambda^3{}_\sigma \varepsilon^{\mu\nu\rho\sigma}, \tag{1.77}$$

and for $\alpha\beta\gamma\delta = 1023$

$$\Lambda^1{}_\nu \Lambda^0{}_\mu \Lambda^2{}_\rho \Lambda^3{}_\sigma \varepsilon^{\nu\mu\rho\sigma} = -\Lambda^0{}_\mu \Lambda^1{}_\nu \Lambda^2{}_\rho \Lambda^3{}_\sigma \varepsilon^{\mu\nu\rho\sigma}, \tag{1.78}$$

where we relabelled indices. For $\alpha\beta\gamma\delta = 0023$ we have

$$\begin{aligned} \Lambda^0{}_\mu \Lambda^0{}_\nu \Lambda^2{}_\rho \Lambda^3{}_\sigma \varepsilon^{\mu\nu\rho\sigma} &= \Lambda^0{}_\mu \Lambda^0{}_\nu \Lambda^2{}_\rho \Lambda^3{}_\sigma (-\varepsilon^{\nu\mu\rho\sigma}) \\ &= -\Lambda^0{}_\mu \Lambda^0{}_\nu \Lambda^2{}_\rho \Lambda^3{}_\sigma \varepsilon^{\mu\nu\rho\sigma} \\ \rightarrow \quad &= 0. \end{aligned} \tag{1.79}$$

Thus,

$$\varepsilon^{\alpha\beta\gamma\delta} \rightarrow \varepsilon'^{\alpha\beta\gamma\delta} = \Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu \Lambda^\gamma{}_\rho \Lambda^\delta{}_\sigma \varepsilon^{\mu\nu\rho\sigma} = \text{const} \cdot \varepsilon^{\alpha\beta\gamma\delta}, \tag{1.80}$$

which is also fully antisymmetric. The constant can be determined simply. For $\alpha\beta\gamma\delta = 0123$ we have

$$\begin{aligned} \Lambda^0{}_\mu \Lambda^1{}_\nu \Lambda^2{}_\rho \Lambda^3{}_\sigma \varepsilon^{\mu\nu\rho\sigma} &= \text{const} \cdot \varepsilon^{0123} \\ \rightarrow \quad \det \Lambda &= \text{const} \cdot (+1). \end{aligned} \tag{1.81}$$

Since we only consider proper Lorentz transformations, $\det \Lambda = 1$ and we have

$$\varepsilon'^{\alpha\beta\gamma\delta} = \Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu \Lambda^\gamma{}_\rho \Lambda^\delta{}_\sigma \varepsilon^{\mu\nu\rho\sigma} = \varepsilon^{\alpha\beta\gamma\delta}. \tag{1.82}$$

Notice that for the general coordinate transformations needed in general relativity, the determinant of the transformation can differ from 1, and the Levi-Civita symbol is no longer a tensor; it is rather a tensor density (Chapter 3).

Finally, we can construct new tensors from an existing one by lowering and raising indices,

$$\begin{aligned} T^\mu{}_\nu &= \eta_{\nu\rho} T^{\mu\rho}, \\ T_{\mu\nu} &= \eta_{\mu\rho} \eta_{\nu\sigma} T^{\rho\sigma}, \end{aligned} \tag{1.83}$$

and we can contract indices as

$$T^\mu{}_\mu = \eta_{\mu\nu} T^{\mu\nu}, \quad \text{etc.} \tag{1.84}$$

Here $T^\mu{}_\mu$ is a scalar, as shown above for the contraction of two vectors.

1.5 Currents and densities

Consider a set $\{n\}$ of particles with 4-velocities $\frac{\mathrm{d}x_n^\mu}{\mathrm{d}\tau}$. Suppose that each particle is endowed with some intrinsic conserved quantity such as an electric charge q_n. We define the relativistic charge and current density

$$J^\mu = \sum_n q_n \frac{\mathrm{d}x_n^\mu}{\mathrm{d}t} \delta^{(3)}(\vec{x} - \vec{x}_n(t)). \tag{1.85}$$

The δ-function localises the charges at the positions $\vec{x}_n(t)$ of the particles. For $\mu = 0$,

$$\begin{aligned} J^0 &= \sum_n q_n \frac{\mathrm{d}x_n^0}{\mathrm{d}t}\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \\ &= \sum_n q_n \frac{\mathrm{d}t}{\mathrm{d}t}\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \\ &= \sum_n q_n\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)) = \rho \qquad \text{(charge density)}. \end{aligned} \tag{1.86}$$

For $\mu = i = 1, 2, 3$ we have

$$\begin{aligned} J^i &= \sum_n q_n \frac{\mathrm{d}x_n^i}{\mathrm{d}t}\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \\ &= \sum_n q_n v_n^i\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \qquad \text{(current density)}. \end{aligned} \tag{1.87}$$

The divergence of the spatial current is

$$\begin{aligned} \frac{\partial}{\partial x^i} J^i &= \sum_n q_n \frac{\mathrm{d}x_n^i}{\mathrm{d}t} \frac{\partial}{\partial x^i} [\delta^{(3)}(\vec{x} - \vec{x}_n(t))] \\ &= \sum_n q_n \frac{\mathrm{d}x_n^i}{\mathrm{d}t} \left[-\frac{\partial}{\partial x_n^i} \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \right] \\ &= -\sum_n q_n \underbrace{\left[\frac{\mathrm{d}x_n^i}{\mathrm{d}t} \frac{\partial}{\partial x_n^i} \right]}_{\frac{\mathrm{d}}{\mathrm{d}t}} \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \\ &= -\frac{\mathrm{d}}{\mathrm{d}t} \left(\sum_n q_n\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \right) = -\frac{\mathrm{d}J^0}{\mathrm{d}t} \\ \Rightarrow & \boxed{\frac{\mathrm{d}J^0}{\mathrm{d}t} + \vec{\nabla} \cdot \vec{J} = 0} \end{aligned} \tag{1.88}$$

which is the continuity equation. **Remark.** $J^0(\vec{x}, t)$ is a field, so $\frac{\mathrm{d}}{\mathrm{d}t}$ acting on it (and on $\delta^{(3)}(\vec{x} - \vec{x}_n(t))$) means $\frac{\partial}{\partial t}$ at fixed $\vec{x}$; the continuity equation reads $\partial_t J^0 + \vec{\nabla} \cdot \vec{J} = 0$ (Weinberg eq. 2.6.4). The same applies to the energy-momentum tensor below. It can be rewritten as

$$\boxed{\partial_\mu J^\mu = 0} \tag{1.89}$$

The continuity equation implies the conservation of charge. Indeed,

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t} \int \mathrm{d}^3x\, J^0(\vec{x}) = \int \mathrm{d}^3x \frac{\mathrm{d}}{\mathrm{d}t} J^0(\vec{x}) = -\int \mathrm{d}^3x \vec{\nabla} \cdot \vec{J}(\vec{x}) \underset{\text{Gauss}}{=} 0. \tag{1.90}$$

The current density is a 4-vector. To see this, we write

$$\begin{aligned} J^\mu &= \sum_n q_n \frac{\mathrm{d}x_n^\mu}{\mathrm{d}t}\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \\ &= \sum_n q_n \frac{\mathrm{d}x_n^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}\tau}{\mathrm{d}t}\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \\ &= \sum_n q_n \int \mathrm{d}t' \frac{\mathrm{d}\tau}{\mathrm{d}t'}\frac{\mathrm{d}x_n^\mu}{\mathrm{d}\tau}\, \delta^{(3)}(\vec{x} - \vec{x}_n(\tau))\delta(t - t'(\tau)) \\ &= \sum_n q_n \int \mathrm{d}\tau \frac{\mathrm{d}x_n^\mu}{\mathrm{d}\tau}\, \delta^{(4)}(x^\mu - x_n^\mu(\tau)). \end{aligned} \tag{1.91}$$

Under a Lorentz transformation,

$$\begin{aligned} J'^\mu &= \sum_n q_n \int \mathrm{d}\tau \frac{\mathrm{d}x_n'^\mu}{\mathrm{d}\tau}\, \delta^{(4)}(x'^\mu - x_n'^\mu(\tau)) \\ &= \sum_n q_n \int \mathrm{d}\tau\, \Lambda^\mu{}_\nu \frac{\mathrm{d}x_n^\nu}{\mathrm{d}\tau}\, \delta^{(4)}\left(\Lambda^\mu{}_\rho(x^\rho - x_n^\rho(\tau))\right) \\ &= \Lambda^\mu{}_\nu \sum_n q_n \int \mathrm{d}\tau \frac{\mathrm{d}x_n^\nu}{\mathrm{d}\tau} \underbrace{\frac{\delta^{(4)}(x^\mu - x_n^\mu(\tau))}{|\det \Lambda|}}_{=1} \\ &= \Lambda^\mu{}_\nu J^\nu. \end{aligned} \tag{1.92}$$

Note that for general coordinate transformations the determinant of the Jacobian is not one, and current densities transform differently from vectors.

1.6 Energy-momentum tensor

We can now extend the notion of a current density to describe the density and flow of energy and momentum. Recall the definition of the current density

$$J^\mu = \sum_n q_n \frac{\mathrm{d}x_n^\mu}{\mathrm{d}t}\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)),$$

and replace the charge q_n by the momentum components of the particle. This defines a tensor which is called the energy-momentum tensor,

$$T^{\mu\nu} = \sum_n P_n^\mu \frac{\mathrm{d}x_n^\nu}{\mathrm{d}t}\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)). \tag{1.93}$$

Despite its appearance, this is a symmetric tensor. Indeed,

$$\begin{aligned} T^{\mu\nu} &= \sum_n P_n^\mu \frac{\mathrm{d}x_n^\nu}{\mathrm{d}t}\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \\ &= \sum_n P_n^\mu \frac{m_n \frac{\mathrm{d}x_n^\nu}{\mathrm{d}\tau_n}}{m_n \frac{\mathrm{d}t}{\mathrm{d}\tau_n}}\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \\ &= \sum_n \frac{P_n^\mu P_n^\nu}{E_n}\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)), \end{aligned} \tag{1.94}$$

where we have used that

$$P_n^\mu = (E_n, \vec{P}_n) = (m_n\gamma_n, m_n\gamma_n\vec{v}_n), \quad \mathrm{d}t = \gamma_n \mathrm{d}\tau_n. \tag{1.95}$$

We can also write the energy-momentum tensor as an integral of the form

$$T^{\mu\nu} = \int \mathrm{d}\tau \sum_n P_n^\mu \frac{\mathrm{d}x_n^\nu}{\mathrm{d}\tau}\, \delta^{(4)}(x - x_n(\tau)). \tag{1.96}$$

In this form we see explicitly that this is a second-rank tensor. Let us now calculate the divergence of the energy-momentum tensor,

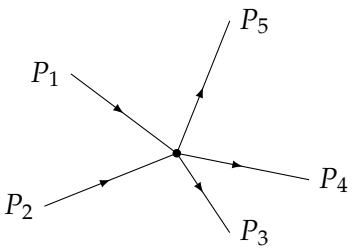
$$\begin{aligned}
\frac{\partial}{\partial x^i} T^{\mu i} &= \sum_n P_n^\mu \frac{\mathrm{d}x_n^i}{\mathrm{d}t} \frac{\partial}{\partial x^i} \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \\
&= \sum_n P_n^\mu \frac{\mathrm{d}x_n^i}{\mathrm{d}t} \left(-\frac{\partial}{\partial x_n^i}\right) \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \\
&= -\sum_n P_n^\mu \frac{\mathrm{d}}{\mathrm{d}t} \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \\
&= -\frac{\mathrm{d}}{\mathrm{d}t} \left(\sum_n P_n^\mu\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \right) + \sum_n \frac{\mathrm{d}P_n^\mu}{\mathrm{d}t}\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \\
&= -\frac{\mathrm{d}}{\mathrm{d}t} T^{\mu 0} + \sum_n \frac{\mathrm{d}P_n^\mu}{\mathrm{d}t}\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \\
\Rightarrow \frac{\mathrm{d}}{\mathrm{d}t} T^{\mu 0} + \frac{\partial}{\partial x^i} T^{\mu i} &= \sum_n \frac{\mathrm{d}P_n^\mu}{\mathrm{d}t}\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \\
\Rightarrow \boxed{\partial_\nu T^{\mu\nu} = G^\mu}
\end{aligned} \tag{1.97}$$

where

$$\begin{aligned}
G^\mu &= \sum_n \frac{\mathrm{d}P_n^\mu}{\mathrm{d}t}\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \\
&= \sum_n \frac{\mathrm{d}P_n^\mu}{\mathrm{d}\tau} \frac{\mathrm{d}\tau}{\mathrm{d}t}\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)) \\
&= \sum_n f_n^\mu \frac{\mathrm{d}\tau}{\mathrm{d}t}\, \delta^{(3)}(\vec{x} - \vec{x}_n(t)),
\end{aligned} \tag{1.98}$$

is the force density. Consider now a system of particles which interact only through collisions, with the 4-momentum conserved in each collision (Fig. 1.5).

e.g.



$$P_1 + P_2 = P_3 + P_4 + P_5.$$

Figure 1.5. A collision in which the total 4-momentum is conserved.

Then

$$\begin{aligned} G^\mu &= \sum_n \frac{\mathrm{d}P_n^\mu}{\mathrm{d}t}\,\delta^{(3)}(\vec{x}-\vec{x}_n(t)) \\ &= \sum_{\text{collisions}} \delta^{(3)}(\vec{x}-\vec{x}_{\text{collisions}}(t))\,\frac{\mathrm{d}}{\mathrm{d}t}\underbrace{\sum_{n\text{ in collisions}} P_n^\mu}_{\text{conserved}} = 0. \end{aligned} \tag{1.99}$$

In this case,

$$\begin{aligned} &\partial_\nu T^{\mu\nu} = 0 \\ &\Rightarrow P^\mu \equiv \int \mathrm{d}^3x\, T^{\mu 0} = \int \mathrm{d}^3x \sum_n P_n^\mu\,\delta^{(3)}(\vec{x}-\vec{x}_n(t)) \\ &\qquad = \sum_n P_n^\mu = \text{ conserved.} \end{aligned} \tag{1.100}$$

1.7 Spin

Consider a system with a conserved energy-momentum tensor, $\partial_\nu T^{\mu\nu} = 0$. We can construct an angular momentum tensor which is also conserved,

$$M^{\alpha\mu\nu} \equiv x^\mu T^{\nu\alpha} - x^\nu T^{\mu\alpha}. \tag{1.101}$$

Then

$$\begin{aligned} \frac{\partial}{\partial x^\alpha} M^{\alpha\mu\nu} &= \frac{\partial}{\partial x^\alpha}\left(x^\mu T^{\nu\alpha} - x^\nu T^{\mu\alpha}\right) \\ &= \delta^\mu{}_\alpha T^{\nu\alpha} - \delta^\nu{}_\alpha T^{\mu\alpha} + x^\mu \cancelto{0}{\frac{\partial}{\partial x^\alpha} T^{\nu\alpha}} - x^\nu \cancelto{0}{\frac{\partial}{\partial x^\alpha} T^{\mu\alpha}} \\ &= T^{\nu\mu} - T^{\mu\nu} \underset{\text{symmetry}}{=} 0. \end{aligned} \tag{1.102}$$

The total angular momentum

$$J^{\mu\nu} = \int \mathrm{d}^3x\, \left(x^\mu T^{\nu 0} - x^\nu T^{\mu 0}\right), \tag{1.103}$$

is then conserved. Notice that under translations $x^\mu \longrightarrow x^\mu + a^\mu = x'^\mu$,

$$J^{\mu\nu} \longrightarrow J^{\mu\nu} + (a^\mu P^\nu - a^\nu P^\mu). \tag{1.104}$$

But there is a part of the angular momentum which is intrinsic, i.e. independent of the choice of the origin of the coordinate system. This part is called spin. It can be obtained from

$$S_\alpha \equiv \frac{1}{2}\varepsilon_{\alpha\beta\gamma\delta} J^{\beta\gamma} V^\delta\,, \qquad \text{with } V^\delta \equiv \frac{P^\delta}{\sqrt{P^2}}, \tag{1.105}$$

the 4-velocity of the system. Under translations $x^\mu \to x^\mu + a^\mu$, we then find

$$\begin{aligned} S_\mu \to S'_\mu &= S_\mu + \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}(a^\nu P^\rho - a^\rho P^\nu)V^\sigma \\ &= S_\mu + \frac{1}{\sqrt{P^2}} \underbrace{\varepsilon_{\mu\nu\rho\sigma}}_{\text{antisymm.}} a^\nu \underbrace{P^\rho P^\sigma}_{\text{symm.}} \\ &= S_\mu + 0. \end{aligned} \tag{1.106}$$

So, the spin vector is invariant under translations (unlike orbital angular momentum).

In the centre-of-mass frame of the system,

$$\sum_n \vec{P}_n = 0 \;\Rightarrow\; P^\mu = \sqrt{P^2}(1,\vec{0}) \quad \text{and} \quad V^\mu = \frac{P^\mu}{\sqrt{P^2}} = (1,\vec{0}). \tag{1.107}$$

Then

$$\begin{aligned} S_\mu &= \frac{1}{2}\varepsilon_{\mu\alpha\beta\gamma}J^{\alpha\beta}V^\gamma = \frac{1}{2}\varepsilon_{\mu\alpha\beta0}J^{\alpha\beta}1 \\ \Rightarrow\; S_0 &= \frac{1}{2}\varepsilon_{0\alpha\beta0}J^{\alpha\beta} = 0, \\ S_1 &= \frac{1}{2}\varepsilon_{1\alpha\beta0}J^{\alpha\beta} = \frac{1}{2}\varepsilon_{1230}J^{23} + \frac{1}{2}\varepsilon_{1320}J^{32} \\ &= \varepsilon_{1230}\frac{J^{23}-J^{32}}{2} = -\varepsilon_{0123}\frac{J^{23}-J^{32}}{2}. \end{aligned} \tag{1.108}$$

Recall that

$$\begin{aligned} J^{23} &= \int \mathrm{d}^3x\,(x^2T^{30} - x^3T^{20}) = -J^{32} \\ \Rightarrow\; S_1 &= (-\varepsilon_{0123})J^{23}. \end{aligned} \tag{1.109}$$

Also,

$$\begin{aligned} \varepsilon_{0123} &= \eta_{0\mu}\eta_{1\nu}\eta_{2\rho}\eta_{3\sigma}\varepsilon^{\mu\nu\rho\sigma} \\ &= \eta_{00}\eta_{11}\eta_{22}\eta_{33}\varepsilon^{0123} \\ &= (+1)(-1)(-1)(-1)(+1) = -1, \end{aligned} \tag{1.110}$$

so that

$$\boxed{S_1 = J^{23}} \tag{1.111}$$

Similarly, we find that $\boxed{S_2 = J^{31}}$ and $\boxed{S_3 = J^{12}}$. **Remark.** These are covariant components. Since $J^{23} = \int \mathrm{d}^3x\,(x^2T^{30} - x^3T^{20})$ is the x-component of the ordinary angular momentum, with the signature $(+,-,-,-)$ used here the contravariant spin is $S^i = -S_i = -(J^{23}, J^{31}, J^{12})$, i.e. minus the angular momentum. Weinberg (§2.9) gets $S = (J^{23}, J^{31}, J^{12})$ without this sign because he uses $\eta = \mathrm{diag}(-1,1,1,1)$, where $S^i = S_i$. With the present signature one may define S_α with an extra minus sign (as for the Pauli–Lubanski vector). The spin vector has three components (rather than four) in every inertial frame. Indeed, we find that

$$S_\mu V^\mu = \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}J^{\nu\rho}V^\sigma V^\mu = \frac{1}{2} \underbrace{\varepsilon_{\mu\nu\rho\sigma}}_{\text{antisymm.}} J^{\nu\rho} \underbrace{V^\sigma V^\mu}_{\text{symm.}} = 0. \tag{1.112}$$

Therefore, the spin 4-vector always satisfies the constraint $S_\mu V^\mu = 0$, which reduces its independent components from four to three.

Remark. In the above derivations we have often used that the contraction of a symmetric (S) and an antisymmetric (A) tensor vanishes. Indeed,

$$\begin{aligned} S_{\mu\nu}A^{\mu\nu} &= S_{\mu\nu}\frac{A^{\mu\nu} - A^{\nu\mu}}{2} \\ &= \frac{1}{2}S_{\mu\nu}A^{\mu\nu} - \frac{1}{2}S_{\mu\nu}A^{\nu\mu} \\ &= \frac{1}{2}S_{\mu\nu}A^{\mu\nu} - \frac{1}{2}S_{\nu\mu}A^{\mu\nu} \\ &= \frac{1}{2}\underbrace{\left(S_{\mu\nu} - S_{\nu\mu}\right)}_{=0}A^{\mu\nu} = 0. \end{aligned} \tag{1.113}$$

2

The principle of equivalence

Let us return to Newtonian mechanics and write the equations of motion for a particle in a homogeneous gravitational field. The particle is additionally subject to other forces $\vec{F}$. We have, for an inertial observer,

$$m_I \ddot{\vec{x}} = m_G \vec{g} + \vec{F}. \tag{2.1}$$

In the above, we have distinguished between the mass coefficient of the acceleration, the so-called inertial mass m_I, and the gravitational mass m_G, which enters the gravitational force on the right-hand side. We now perform a coordinate transformation

$$\vec{x} = \vec{x}' + \frac{1}{2}\vec{g}t^2 \quad (t = t'). \tag{2.2}$$

In this new, accelerated frame we have

$$m_I \ddot{\vec{x}}' = (m_G - m_I)\vec{g} + \vec{F}'. \tag{2.3}$$

In classical mechanics we postulate that the gravitational and inertial masses are equal (or equivalent),

$$m_I = m_G = m. \tag{2.4}$$

Then, the equations of motion in the accelerated frame take the form

$$m \ddot{\vec{x}}' = \vec{F}'. \tag{2.5}$$

In the above, the force of gravity has been eliminated. We have thus found that a gravitational force present in one coordinate system can be traded for an acceleration.

A condition for the above was that the gravitational field is uniform. This is of course unrealistic at large scales. However, if we restrict ourselves to very small regions of space and short times, we can consider the gravitational field to be sufficiently uniform. This sounds like an unwanted approximation, but it is not. We can make our spacetime intervals infinitesimally small, and compensate for this by erecting infinitely many local, freely falling inertial frames, one for each infinitesimal region of spacetime. **Remark.** The cancellation is exact only at a single spacetime point (to first order in the distance from it): in a locally inertial frame one can achieve $g_{\mu\nu} = \eta_{\mu\nu}$ and $\partial_\lambda g_{\mu\nu} = 0$ at the point, but not the vanishing of the second derivatives. The residual tidal accelerations are encoded in the Riemann tensor and are what distinguishes a genuine gravitational field from an accelerated frame (Weinberg Sec. 3.1 and Ch. 6).

Einstein took this observation and used it as a physical principle upon which he constructed the theory of general relativity. Concretely, the principle states:

Locally (i.e. in small spacetime regions) we can erect coordinate systems in which the laws of physics take the form which holds in inertial frames of reference. This form is the one we expect in special relativity.

Let us apply this principle to a single particle in an arbitrary gravitational field. The particle is subjected only to the gravitational force and no other forces act on it. Einstein's principle of equivalence tells us that in small spacetime regions we can find coordinate transformations

$$\underbrace{x^\alpha}_{\text{lab-frame}} \to \underbrace{\xi^\alpha}_{\substack{\text{locally}\\ \text{inertial}\\ \text{frame}}}, \tag{2.6}$$

that cancel the effect of gravity. One needs a different transformation of this kind in different patches of spacetime. However, such transformations must exist, according to the equivalence principle, for all patches.

Let us focus on one such local patch. In the new reference frame (ξ^α coordinates) the laws of physics take their special-relativistic form. In particular, proper time intervals are given by

$$\mathrm{d}\tau^2 = \eta_{\mu\nu}\mathrm{d}\xi^\mu \mathrm{d}\xi^\nu, \tag{2.7}$$

and the equations of motion for the particle are those of a free particle, that is,

$$\frac{\mathrm{d}^2\xi^\mu}{\mathrm{d}\tau^2} = 0. \tag{2.8}$$

The metric $\eta_{\mu\nu}$ in (2.7) is the Minkowski metric $\eta_{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)$.

What is the form of (2.6) and (2.7) in the lab frame, where gravity is not eliminated? We have

$$\begin{aligned} \mathrm{d}\tau^2 &= \eta_{\mu\nu}\mathrm{d}\xi^\mu\mathrm{d}\xi^\nu = \eta_{\mu\nu}\frac{\partial\xi^\mu}{\partial x^\alpha}\frac{\partial\xi^\nu}{\partial x^\beta}\mathrm{d}x^\alpha\mathrm{d}x^\beta \\ &\to \boxed{\mathrm{d}\tau^2 = g_{\alpha\beta}(x)\,\mathrm{d}x^\alpha\mathrm{d}x^\beta} \end{aligned} \tag{2.9}$$

and the metric is not Minkowskian anymore:

$$\boxed{g_{\alpha\beta} \equiv \eta_{\mu\nu}\frac{\partial\xi^\mu}{\partial x^\alpha}\frac{\partial\xi^\nu}{\partial x^\beta}} \tag{2.10}$$

We will soon see that the metric encodes the gravitational potential. The equations of

motion of the particle in terms of the lab-frame variables are

$$\begin{aligned} 0 &= \frac{\mathrm{d}^2\xi^\alpha}{\mathrm{d}\tau^2} = \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\mathrm{d}\xi^\alpha}{\mathrm{d}\tau} = \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\right) \\ &= \frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} + \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial\xi^\alpha}{\partial x^\mu} \\ &= \frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} + \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\partial^2\xi^\alpha}{\partial x^\mu\partial x^\nu} \\ \rightarrow\ 0 &= \frac{\partial x^\gamma}{\partial\xi^\alpha}\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} + \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\partial x^\gamma}{\partial\xi^\alpha}\frac{\partial^2\xi^\alpha}{\partial x^\mu\partial x^\nu} \\ \Rightarrow\ 0 &= \delta^\gamma{}_\mu\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} + \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\partial x^\gamma}{\partial\xi^\alpha}\frac{\partial^2\xi^\alpha}{\partial x^\mu\partial x^\nu}, \end{aligned} \tag{2.11}$$

$$\Rightarrow \boxed{\frac{\mathrm{d}^2x^\gamma}{\mathrm{d}\tau^2} + \Gamma^\gamma{}_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} = 0} \tag{2.12}$$

where

$$\boxed{\Gamma^\gamma{}_{\mu\nu} \equiv \frac{\partial x^\gamma}{\partial\xi^\alpha}\frac{\partial^2\xi^\alpha}{\partial x^\mu\partial x^\nu}} \tag{2.13}$$

is called the affine connection.

Physically, it plays the role of the Newtonian gravitational force per unit mass. This is best seen in the rest frame of the particle, where

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \left(\frac{\mathrm{d}t}{\mathrm{d}\tau}, \overbrace{\frac{\mathrm{d}\vec{x}}{\mathrm{d}\tau}}^{0}\right) = \left(\frac{\mathrm{d}t}{\mathrm{d}\tau}, \vec{0}\right). \tag{2.14}$$

Then (2.12) gives

$$\begin{aligned} \frac{\mathrm{d}^2x^\alpha}{\mathrm{d}\tau^2} + \Gamma^\alpha{}_{00}\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 &= 0 \\ \frac{\mathrm{d}^2x^\alpha}{\mathrm{d}t^2}\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 + \Gamma^\alpha{}_{00}\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 &= 0 \\ \rightarrow\ \frac{\mathrm{d}^2x^\alpha}{\mathrm{d}t^2} &= -\Gamma^\alpha{}_{00}. \end{aligned} \tag{2.15}$$

For $\alpha = 1, 2, 3$, this is the analogue of the Newtonian equation

$$\frac{\mathrm{d}^2\vec{x}}{\mathrm{d}t^2} = \vec{F}_{\mathrm{grav}}/m. \tag{2.16}$$

Before we proceed to determine the affine connection, we should make a remark about the metric of (2.10). If the coordinate transformation $\xi^\alpha \rightarrow x^\alpha$ can be inverted, which is what one expects physically, then the metric has an inverse,

$$g^{\mu\nu} \equiv \eta^{\rho\sigma}\frac{\partial x^\mu}{\partial\xi^\rho}\frac{\partial x^\nu}{\partial\xi^\sigma}. \tag{2.17}$$

Indeed,

$$\begin{aligned} g^{\mu\rho} g_{\rho\nu} &= \left(\eta^{\alpha\beta} \frac{\partial x^\mu}{\partial \xi^\alpha} \frac{\partial x^\rho}{\partial \xi^\beta}\right) \left(\eta_{\gamma\delta} \frac{\partial \xi^\gamma}{\partial x^\rho} \frac{\partial \xi^\delta}{\partial x^\nu}\right) \\ &= \eta^{\alpha\beta} \eta_{\gamma\delta} \frac{\partial x^\mu}{\partial \xi^\alpha} \frac{\partial \xi^\delta}{\partial x^\nu} \underbrace{\frac{\partial x^\rho}{\partial \xi^\beta} \frac{\partial \xi^\gamma}{\partial x^\rho}}_{=\frac{\partial \xi^\gamma}{\partial \xi^\beta} = \delta^\gamma{}_\beta} \\ &= \cdots = \delta^\mu{}_\nu \,, \end{aligned} \tag{2.18}$$

where the dots stand for the remaining contractions: in order, the metric $\eta_{\gamma\delta}$ with the Kronecker delta, the two Minkowski metrics with each other, the resulting Kronecker delta with the ξ-derivatives, and the remaining derivatives with each other.

Let us now return to the affine connection $\Gamma^\alpha{}_{\mu\nu}$. It can be derived from the metric as

$$\Gamma^\alpha{}_{\mu\nu} = \frac{1}{2} g^{\alpha\rho} \left\{ \frac{\partial g_{\rho\nu}}{\partial x^\mu} + \frac{\partial g_{\rho\mu}}{\partial x^\nu} - \frac{\partial g_{\mu\nu}}{\partial x^\rho} \right\}. \tag{2.19}$$

Proof:

$$\begin{aligned} \frac{\partial g_{\alpha\beta}}{\partial x^\gamma} &= \frac{\partial}{\partial x^\gamma} \left[\eta_{\mu\nu} \frac{\partial \xi^\mu}{\partial x^\alpha} \frac{\partial \xi^\nu}{\partial x^\beta} \right] \\ &= \langle \alpha\gamma | \beta \rangle + \langle \beta\gamma | \alpha \rangle \,, \end{aligned} \tag{2.20}$$

with

$$\langle \alpha\beta | \gamma \rangle \equiv \frac{\partial^2 \xi^\mu}{\partial x^\alpha \partial x^\beta} \frac{\partial \xi^\nu}{\partial x^\gamma} \eta_{\mu\nu}. \tag{2.21}$$

Then

$$\begin{aligned} \frac{\partial g_{\rho\mu}}{\partial x^\nu} + \frac{\partial g_{\rho\nu}}{\partial x^\mu} - \frac{\partial g_{\mu\nu}}{\partial x^\rho} &= \cancel{\langle \rho\nu | \mu \rangle} + \langle \mu\nu | \rho \rangle + \cancel{\langle \rho\mu | \nu \rangle} + \langle \mu\nu | \rho \rangle - \cancel{\langle \rho\mu | \nu \rangle} - \cancel{\langle \rho\nu | \mu \rangle} \\ &= 2 \langle \mu\nu | \rho \rangle \,. \end{aligned} \tag{2.22}$$

Thus,

$$\begin{aligned} \frac{1}{2} g^{\alpha\rho} \left\{ \frac{\partial g_{\rho\nu}}{\partial x^\mu} + \frac{\partial g_{\rho\mu}}{\partial x^\nu} - \frac{\partial g_{\mu\nu}}{\partial x^\rho} \right\} &= \frac{1}{2} 2 \langle \mu\nu | \rho \rangle \, g^{\alpha\rho} \\ &= g^{\alpha\rho} \frac{\partial^2 \xi^\beta}{\partial x^\mu \partial x^\nu} \frac{\partial \xi^\gamma}{\partial x^\rho} \eta_{\beta\gamma} \\ &= \eta^{\varepsilon\delta} \frac{\partial x^\alpha}{\partial \xi^\varepsilon} \frac{\partial x^\rho}{\partial \xi^\delta} \frac{\partial^2 \xi^\beta}{\partial x^\mu \partial x^\nu} \frac{\partial \xi^\gamma}{\partial x^\rho} \eta_{\beta\gamma} \\ &= \eta^{\varepsilon\delta} \frac{\partial x^\alpha}{\partial \xi^\varepsilon} \frac{\partial^2 \xi^\beta}{\partial x^\mu \partial x^\nu} \delta^\gamma{}_\delta \eta_{\beta\gamma} \\ &= \frac{\partial x^\alpha}{\partial \xi^\beta} \frac{\partial^2 \xi^\beta}{\partial x^\mu \partial x^\nu} = \Gamma^\alpha{}_{\mu\nu}. \end{aligned} \tag{2.23}$$

In summary, we have shown that the affine connection can be computed from the metric via

$$\Gamma^\alpha{}_{\mu\nu} = \frac{1}{2} g^{\alpha\rho} \left\{ \frac{\partial g_{\rho\nu}}{\partial x^\mu} + \frac{\partial g_{\rho\mu}}{\partial x^\nu} - \frac{\partial g_{\mu\nu}}{\partial x^\rho} \right\}. \tag{2.19}$$

The right-hand side of the above equation has a wider mathematical meaning. It is known as the Christoffel symbol of the second kind and is also denoted by $\left\{ {\alpha \atop \mu\nu} \right\}$. Notice that the affine connection is symmetric under the exchange $\mu \leftrightarrow \nu$.

2.1 Newtonian limit

We shall now apply the equations of motion derived above in the limit of

- a slowly moving particle
- inside a static and weak gravitational field.

As the particle moves slowly, we have

$$\frac{\mathrm{d}x^i}{\mathrm{d}\tau} = \frac{\mathrm{d}x^i}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}\tau} = \underbrace{v_i}_{\text{velocity}}\frac{\mathrm{d}t}{\mathrm{d}\tau} \ll \frac{\mathrm{d}t}{\mathrm{d}\tau}. \tag{2.24}$$

Then, the equations of motion simplify to

$$\begin{aligned} &\frac{\mathrm{d}^2x^\alpha}{\mathrm{d}\tau^2} + \Gamma^\alpha{}_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} = 0 \\ \rightarrow\quad &\frac{\mathrm{d}^2x^\alpha}{\mathrm{d}t^2}\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 + \Gamma^\alpha{}_{00}\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 \approx 0 \\ \rightarrow\quad &\boxed{\frac{\mathrm{d}^2x^\alpha}{\mathrm{d}t^2} \approx -\Gamma^\alpha{}_{00}} \end{aligned} \tag{2.25}$$

Remark. Strictly, $\frac{\mathrm{d}^2x^i}{\mathrm{d}\tau^2} = \frac{\mathrm{d}^2x^i}{\mathrm{d}t^2}\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 + \frac{\mathrm{d}x^i}{\mathrm{d}t}\frac{\mathrm{d}^2t}{\mathrm{d}\tau^2}$, and the second step above for $\alpha = 0$ would read $0 = -\Gamma^0_{00}$. A cleaner order (Weinberg Sec. 3.4) is to keep $\frac{\mathrm{d}^2x^\alpha}{\mathrm{d}\tau^2} + \Gamma^\alpha_{00}(\frac{\mathrm{d}t}{\mathrm{d}\tau})^2 = 0$, show below that $\Gamma^0_{00} = 0$ so $\frac{\mathrm{d}t}{\mathrm{d}\tau}$ is constant, and only then divide by $(\frac{\mathrm{d}t}{\mathrm{d}\tau})^2$ in the spatial equations. The relevant components of the affine connection are

$$\Gamma^\alpha{}_{00} = \frac{1}{2}g^{\alpha\rho}\left\{\frac{\partial g_{0\rho}}{\partial x^0} + \frac{\partial g_{0\rho}}{\partial x^0} - \frac{\partial g_{00}}{\partial x^\rho}\right\}. \tag{2.26}$$

As the field is static, we can neglect the time derivatives of the metric and thus,

$$\Gamma^\alpha{}_{00} \approx -\frac{1}{2}g^{\alpha\rho}\frac{\partial g_{00}}{\partial x^\rho}. \tag{2.27}$$

Finally, we use the approximation that the gravitational field is weak (e.g. far away from very massive bodies). We can then write the metric as the Minkowski metric plus a small perturbation,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}\,, \quad |h_{\mu\nu}| \text{ small}. \tag{2.28}$$

Then

$$\begin{aligned} \Gamma^\alpha{}_{00} &\approx -\frac{1}{2}\left(\eta^{\alpha\rho} - h^{\alpha\rho}\right)\frac{\partial}{\partial x^\rho}(\eta_{00} + h_{00}) \\ &\approx -\frac{1}{2}\eta^{\alpha\rho}\frac{\partial h_{00}}{\partial x^\rho}. \end{aligned} \tag{2.29}$$

The components are

$$\Gamma^0{}_{00} \approx -\frac{1}{2}\eta^{0\rho}\frac{\partial h_{00}}{\partial x^\rho} = -\frac{1}{2}\underbrace{\frac{\partial h_{00}}{\partial t}}_{0,\ \text{static}} = 0, \tag{2.30}$$

$$\Gamma^j{}_{00} \approx -\frac{1}{2}\eta^{j\rho}\frac{\partial h_{00}}{\partial x^\rho} = \frac{1}{2}\frac{\partial h_{00}}{\partial x^j}. \tag{2.31}$$

Then, the equations of motion become

$$\begin{aligned} &\frac{\mathrm{d}^2 t}{\mathrm{d}\tau^2} + \cancelto{0}{\Gamma^0_{00}} \left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 = 0 \\ \rightarrow &\boxed{\frac{\mathrm{d}^2 t}{\mathrm{d}\tau^2} = 0} \end{aligned} \tag{2.32}$$

and

$$\frac{\mathrm{d}^2 x^j}{\mathrm{d}t^2} = -\Gamma^j{}_{00} = -\frac{1}{2}\frac{\partial h_{00}}{\partial x^j}, \tag{2.33}$$

or, in a vectorial form,

$$\boxed{\frac{\mathrm{d}^2 \vec{x}}{\mathrm{d}t^2} = -\vec{\nabla}\left(\frac{1}{2}h_{00}\right)} \tag{2.34}$$

We can compare this result with the Newtonian theory, which gives

$$\frac{\mathrm{d}^2 \vec{x}}{\mathrm{d}t^2} = -\vec{\nabla}\phi, \tag{2.35}$$

with $\phi(\vec{x})$ being the gravitational potential. We naturally identify

$$h_{00} = 2\phi + \text{const.} \tag{2.36}$$

At infinity, we expect the gravitational potential to vanish and the metric to be Minkowskian,

$$\overbrace{\lim_{|\vec{x}|\to\infty} h_{00}(\vec{x})}^{0} = 2\overbrace{\lim_{|\vec{x}|\to\infty} \phi(\vec{x})}^{0} + \text{const}, \tag{2.37}$$

so that the constant vanishes and

$$h_{00} = 2\phi(\vec{x}). \tag{2.38}$$

The 00 component of the metric is thus

$$g_{00} = 1 + 2\phi(\vec{x}). \tag{2.39}$$

As we know, the metric defines the geometrical properties of spacetime. These properties are altered by gravity. In what follows, we will explore some of the phenomena which emerge due to gravitational effects on spacetime.

Let us get an estimate of how much the metric changes with respect to the Minkowski metric on the surface of the Earth. The potential at the surface is given by

$$\begin{aligned} \phi = -\frac{GM}{R} &= -\left(\frac{GM}{R^2}\right)R \\ &= -\underbrace{g}_{9{,}81\,\mathrm{m/s^2}}\overbrace{R}^{6371\,\mathrm{km}} \\ &\overset{\text{natural units}}{=} -\frac{gR}{c^2} \approx -\frac{10\cdot 6\cdot 10^6}{(3\cdot 10^8)^2} \\ &\approx -0.7\cdot 10^{-9}. \end{aligned} \tag{2.40}$$

So, at the surface of the Earth, the deviation of the spacetime geometry from Minkowski space is tiny. This makes direct experimental tests of general relativity difficult, but not impossible, and by now the theory has been tested in various ways.

2.2 Clocks and light in a gravitational field

Consider a clock which is at rest in our lab frame, inside a non-zero gravitational field. We would like to compare its ticking rate with the ticking rate of the same clock at rest in the absence of the gravitational field:

$$\Delta t\Big|_{\text{gravity \& at rest}} \quad \text{vs} \quad \Delta t\Big|_{\text{no gravity \& at rest}}$$

To perform this comparison we will use the equivalence principle. First, we will perform a coordinate transformation (not a Lorentz transformation) which cancels gravity in the spacetime region of the clock. Second, we will perform a Lorentz transformation which will bring us to an equivalent inertial frame with gravity cancelled out and the clock at rest. The procedure is sketched in Fig. 2.1.

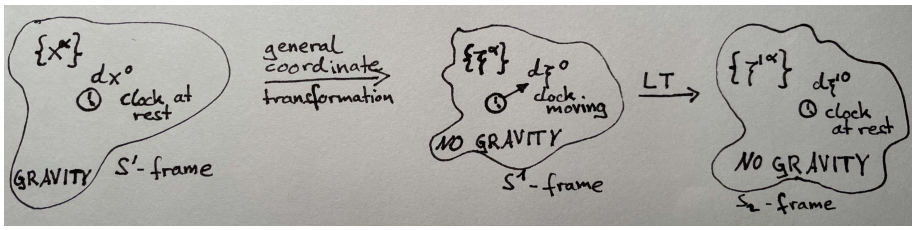


Figure 2.1. Left: the lab frame (coordinates x^α), with gravity and the clock at rest. Middle: after a general coordinate transformation, a locally inertial frame (ξ^α) without gravity, in which the clock moves. Right: after a Lorentz transformation, the inertial frame S_2 (ξ'^α), in which the clock is at rest.

In the S_2 frame, the time interval between two ticks of the clock is a proper time interval,

$$\mathrm{d}\xi'^0 = \mathrm{d}\tau, \tag{2.41}$$

because the clock is at rest as viewed in an inertial frame.

Moreover, the proper time interval is the same for all inertial observers, and thus

$$\mathrm{d}\tau = \mathrm{d}\xi'^0 = \sqrt{\eta_{\mu\nu}\mathrm{d}\xi^\mu \mathrm{d}\xi^\nu}. \tag{2.42}$$

It remains to perform the general coordinate transformation, which expresses $\mathrm{d}\tau$ (the ticking interval at rest in the absence of gravitation) in terms of the spacetime interval which separates two successive ticks in our lab frame. We get

$$\begin{aligned} \mathrm{d}\tau &= \sqrt{\eta_{\mu\nu}\mathrm{d}\xi^\mu \mathrm{d}\xi^\nu} = \sqrt{\underbrace{\eta_{\mu\nu}\frac{\partial\xi^\mu}{\partial x^\rho}\frac{\partial\xi^\nu}{\partial x^\sigma}}_{=g_{\rho\sigma}}\mathrm{d}x^\rho \mathrm{d}x^\sigma} \\ &= \sqrt{g_{\rho\sigma}(x)\,\mathrm{d}x^\rho \mathrm{d}x^\sigma}. \end{aligned} \tag{2.43}$$

In the lab frame the clock is at rest: $\mathrm{d}x^1 = \mathrm{d}x^2 = \mathrm{d}x^3 = 0$, $\mathrm{d}x^0 \equiv \Delta t \neq 0$. Thus,

$$\mathrm{d}\tau = \sqrt{g_{00}(x)\,\mathrm{d}x^0 \mathrm{d}x^0} = \sqrt{g_{00}(x)}\Delta t, \tag{2.44}$$

$$\Rightarrow \Delta t = \frac{\mathrm{d}\tau}{\sqrt{g_{00}(x)}}. \tag{2.45}$$

Here Δt is the time elapsed between ticks of the clock at rest inside the gravitational field and $d\tau$ is the time elapsed between ticks of the clock at rest outside the gravitational field.

Equation (2.45) allows us to compare the rates of clocks placed at rest at various points within the gravitational field. This is more practical than taking our clock to a place with no gravity. For two points x_A and x_B we have

$$\Delta t_A = \frac{d\tau}{\sqrt{g_{00}(x_A)}}, \quad \Delta t_B = \frac{d\tau}{\sqrt{g_{00}(x_B)}}, \tag{2.46}$$

$$\Rightarrow \frac{\Delta t_A}{\Delta t_B} = \sqrt{\frac{g_{00}(x_B)}{g_{00}(x_A)}}. \tag{2.47}$$

For the ticking frequencies $\nu \equiv (\Delta t)^{-1}$ of the clocks, this gives

$$\frac{\nu_B}{\nu_A} = \sqrt{\frac{g_{00}(x_B)}{g_{00}(x_A)}}, \tag{2.48}$$

$$\Rightarrow \nu_B = \nu_A \sqrt{\frac{g_{00}(x_B)}{g_{00}(x_A)}}. \tag{2.49}$$

Let us now use the weak-field approximation. We have found the expression $g_{00}(x) \approx 1 + 2\phi(x)$, Eq. (2.39), for the 00 component of the metric in terms of the gravitational potential. Thus, in a weak field $\nu_B \approx \nu_A \left(1 + 2\phi(x_B)\right)^{1/2} \left(1 + 2\phi(x_A)\right)^{-1/2}$, or, after a Taylor expansion,

$$\boxed{\nu_B \approx \nu_A (1 + \phi(x_B) - \phi(x_A))} \tag{2.50}$$

This equation applies not only to traditional clocks, but to the frequency of any physical process. For example, ν_A and ν_B could refer to the frequency of light produced with identical mechanisms (same atomic transitions) at points A and B in spacetime. **Remark.** This step needs one more ingredient: in a static field the light signal from B arrives at A with the same coordinate-time period Δt_B it had at emission, because successive wave crests take the same coordinate time to travel from B to A ($g_{\mu\nu}$ is t-independent). Only then is the observed ratio of received to local frequency at A equal to $\sqrt{g_{00}(x_B)/g_{00}(x_A)}$ (Weinberg Sec. 3.5). Let us choose A to be a point at the surface of the Earth and B a point at the surface of the Sun. Then

$$\underset{\text{sun}}{\nu_\odot} \approx \underset{\text{earth}}{\nu_\oplus}(1 + \phi_\odot - \phi_\oplus). \tag{2.51}$$

We have already estimated the gravitational potential at the surface of the Earth to be about $\phi_\oplus \approx -10^{-9}$. The Sun has a mass of about $3 \cdot 10^5$ times the mass of the Earth and a radius of about 10^2 times the radius of the Earth. A very rough back-of-the-envelope estimate then gives

$$\phi_\odot \approx \frac{3 \cdot 10^5}{10^2} \phi_\oplus \approx -2 \cdot 10^{-6}. \tag{2.52}$$

We conclude that

$$\nu_\odot \approx \nu_\oplus (1 - 2 \cdot 10^{-6}), \quad \Rightarrow \frac{\nu_\odot - \nu_\oplus}{\nu_\oplus} \approx -2 \cdot 10^{-6}. \tag{2.53}$$

For a photon in the optical spectrum, this corresponds to a redshift (a negative shift) of about two parts in a million. This is really tiny and difficult to discern, but there have been measurements which are consistent with such a shift.

2.3 Geodesics and the motion of a particle in a gravitational field

We have derived the equations of motion of a particle in the gravitational field by means of the principle of equivalence,

$$\begin{aligned} \frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}\tau^2} + \Gamma^\alpha{}_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} = 0, \\ \mathrm{d}\tau^2 = g_{\mu\nu}(x) \mathrm{d}x^\mu \mathrm{d}x^\nu, \end{aligned} \tag{2.54}$$

and

$$\Gamma^\alpha{}_{\mu\nu} = \frac{1}{2} g^{\alpha\rho} \left\{ \frac{\partial g_{\rho\nu}}{\partial x^\mu} + \frac{\partial g_{\rho\mu}}{\partial x^\nu} - \frac{\partial g_{\mu\nu}}{\partial x^\rho} \right\}. \tag{2.55}$$

We will now show that the equations of motion describe a geodesic in spacetime. Namely, the trajectory makes the proper time interval stationary (for timelike curves, locally maximal) between any two points:

$$\begin{aligned} \delta \int_{\tau_A}^{\tau_B} \mathrm{d}\tau &= \delta \int_{\tau_A}^{\tau_B} \sqrt{g_{\mu\nu}(x) \mathrm{d}x^\mu \mathrm{d}x^\nu} \\ &= \delta \int_{p_A}^{p_B} \mathrm{d}p \sqrt{g_{\mu\nu}(x) \frac{\mathrm{d}x^\mu}{\mathrm{d}p} \frac{\mathrm{d}x^\nu}{\mathrm{d}p}} \quad (p\text{: parameter describing the trajectory}) \\ &= \int_{p_A}^{p_B} \frac{\mathrm{d}p}{2} \underbrace{\left(g_{\mu\nu}(x) \frac{\mathrm{d}x^\mu}{\mathrm{d}p} \frac{\mathrm{d}x^\nu}{\mathrm{d}p} \right)^{-1/2}}_{(\frac{\mathrm{d}\tau}{\mathrm{d}p})^2} \left(\frac{\partial g_{\mu\nu}}{\partial x^\alpha} \delta x^\alpha \frac{\mathrm{d}x^\mu}{\mathrm{d}p} \frac{\mathrm{d}x^\nu}{\mathrm{d}p} + 2 g_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}p} \frac{\mathrm{d}\delta x^\nu}{\mathrm{d}p} \right) \\ &= \int_{p_A}^{p_B} \frac{\mathrm{d}p}{2} \left(\frac{\mathrm{d}p}{\mathrm{d}\tau} \right) \left[\frac{\partial g_{\mu\nu}}{\partial x^\alpha} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \delta x^\alpha + 2 g_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \frac{\mathrm{d}\delta x^\nu}{\mathrm{d}\tau} \right] \left(\frac{\mathrm{d}\tau}{\mathrm{d}p} \right)^2 \\ &= \int_{\tau_A}^{\tau_B} \mathrm{d}\tau \left[\frac{1}{2} \frac{\partial g_{\mu\nu}}{\partial x^\alpha} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \delta x^\alpha - \frac{\mathrm{d}}{\mathrm{d}\tau} \left(g_{\mu\alpha} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \right) \delta x^\alpha \right] \\ &= \int_{\tau_A}^{\tau_B} \mathrm{d}\tau \left[\frac{1}{2} \frac{\partial g_{\mu\nu}}{\partial x^\alpha} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} - \frac{\partial g_{\mu\alpha}}{\partial x^\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} - g_{\mu\alpha} \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} \right] \delta x^\alpha \\ &= \int_{\tau_A}^{\tau_B} \mathrm{d}\tau \left[\left(\frac{1}{2} \frac{\partial g_{\mu\nu}}{\partial x^\alpha} - \frac{1}{2} \frac{\partial g_{\mu\alpha}}{\partial x^\nu} - \frac{1}{2} \frac{\partial g_{\nu\alpha}}{\partial x^\mu} \right) \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} - g_{\mu\alpha} \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} \right] \delta x^\alpha \\ &= - \int_{\tau_A}^{\tau_B} \mathrm{d}\tau \, g_{\alpha\mu} \left[\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu{}_{\nu\beta} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} \right] \delta x^\alpha. \end{aligned} \tag{2.56}$$

The bracket above vanishes by the geodesic equation, leading to

$$\boxed{\delta \int_{\tau_A}^{\tau_B} \mathrm{d}\tau = 0} \tag{2.57}$$

Alternatively, instead of using the equivalence principle, we could have used the variational principle to derive our equations of motion.

Let us now make a remark about how we can set the boundary conditions for solving the equations of motion of a particle moving in a gravitational field with a known metric $g_{\mu\nu}$.

The system of differential equations that we must solve is

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu{}_{\rho\sigma} \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} \frac{\mathrm{d}x^\sigma}{\mathrm{d}\tau} = 0, \tag{2.58}$$

$$\text{and} \qquad g_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} = 1. \tag{2.59}$$

To satisfy (2.59), we only need to make sure, by choosing appropriate boundary conditions, that it holds at an initial proper time τ_0. It then remains satisfied at all proper times. Indeed,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\tau}\left(g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\right) &= 2g_{\mu\nu}\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} + \left(\frac{\mathrm{d}}{\mathrm{d}\tau}g_{\mu\nu}\right)\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \\
&= 2g_{\mu\nu}\left(-\Gamma^\mu{}_{\rho\sigma}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau}\frac{\mathrm{d}x^\sigma}{\mathrm{d}\tau}\right)\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} + \frac{\partial g_{\rho\nu}}{\partial x^\sigma}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau}\frac{\mathrm{d}x^\sigma}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \\
&= \left[-\underbrace{g_{\mu\nu}g^{\mu\delta}}_{\delta^\delta{}_\nu}\left(\frac{\partial g_{\sigma\delta}}{\partial x^\rho} + \frac{\partial g_{\rho\delta}}{\partial x^\sigma} - \frac{\partial g_{\rho\sigma}}{\partial x^\delta}\right) + \frac{\partial g_{\rho\nu}}{\partial x^\sigma}\right]\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau}\frac{\mathrm{d}x^\sigma}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \\
&= \left[-\frac{\partial g_{\sigma\nu}}{\partial x^\rho} - \cancel{\frac{\partial g_{\rho\nu}}{\partial x^\sigma}} + \frac{\partial g_{\rho\sigma}}{\partial x^\nu} + \cancel{\frac{\partial g_{\rho\nu}}{\partial x^\sigma}}\right]\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau}\frac{\mathrm{d}x^\sigma}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \\
&= \left[-\frac{\partial g_{\sigma\nu}}{\partial x^\rho} + \frac{\partial g_{\sigma\nu}}{\partial x^\rho}\right]\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau}\frac{\mathrm{d}x^\sigma}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} = 0.
\end{aligned} \tag{2.60}$$

So, as long as (2.58) is satisfied and the initial conditions are chosen to also satisfy (2.59) at τ_0, (2.59) remains satisfied. We can thus focus our attention on solving (2.58).

3

Tensors and tensor densities in general relativity

With the principle of equivalence, we have been able to deduce simple laws for the motion of a particle inside a gravitational field from the laws of special relativity in the absence of gravity. We could apply the same technique in order to derive more complicated laws of physics. For example, we could obtain in the same way the laws that govern the motion of a particle in both a gravitational and an electromagnetic field from Maxwell's theory, which is valid only in the absence of gravity.

However, there is an easier recipe to derive laws of physics which are consistent with the principle of equivalence. This is the method of general covariance. According to this method, we seek equations that satisfy two conditions:

(i) They are valid in all coordinate systems, i.e. they are covariant under general coordinate transformations (maintaining the same form).

(ii) They are valid in the absence of gravity.

Indeed, if conditions (i) and (ii) are satisfied, then the equations hold in all local inertial frames that can be erected at each spacetime point, as requested by the principle of equivalence.

Equations expressing physical laws contain spacetime coordinates, their derivatives and fields, such as the electromagnetic field. These mathematical objects have or must be endowed with precise transformation properties under general coordinate transformations. In what follows, we will classify their transformation properties, introducing scalars, vectors, tensors and densities. With this classification at hand, we will then be able to write equations which transform covariantly under general coordinate transformations.

3.1 Scalars, vectors, tensors

We will now examine how various physical quantities may transform under a general coordinate transformation,

$$x^\mu \longrightarrow x'^\mu = x'^\mu(x). \tag{3.1}$$

The simplest objects are scalars, which remain the same in all frames.

• **Scalars:**

$$\phi(x) \to \phi'(x') = \phi(x). \tag{3.2}$$

There are 4-component objects which are called contravariant vectors. Under general coordinate transformations they transform as

• **Contravariant vectors:**

$$V^\mu(x) \to V'^\mu(x') = V^\rho(x)\frac{\partial x'^\mu}{\partial x^\rho}. \tag{3.3}$$

For example, infinitesimal spacetime intervals are contravariant vectors. Indeed,

$$\mathrm{d}x'^\mu = \frac{\partial x'^\mu}{\partial x^\rho}\mathrm{d}x^\rho. \tag{3.4}$$

Other noteworthy 4-component objects are covariant vectors, which transform as

• **Covariant vectors:**

$$U_\mu \to U'_\mu(x') = \frac{\partial x^\rho}{\partial x'^\mu}U_\rho(x). \tag{3.5}$$

An example of a covariant vector is the derivative of a scalar quantity. Indeed,

$$\frac{\partial\phi(x)}{\partial x^\mu} \to \frac{\partial\phi'(x')}{\partial x'^\mu} = \frac{\partial\phi(x)}{\partial x'^\mu} = \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial\phi(x)}{\partial x^\rho}. \tag{3.6}$$

We now move on to noteworthy objects with 16 components, the second-rank tensors. These can be:

• **Contravariant second-rank tensors:** $A^{\mu\nu}(x)$ which transform as

$$A^{\mu\nu}(x) \to A'^{\mu\nu}(x') = \frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x'^\nu}{\partial x^\sigma}A^{\rho\sigma}(x). \tag{3.7}$$

• **Covariant second-rank tensors:** $B_{\mu\nu}(x)$ which transform as

$$B_{\mu\nu}(x) \to B'_{\mu\nu}(x') = \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x^\sigma}{\partial x'^\nu}B_{\rho\sigma}(x). \tag{3.8}$$

• **Mixed second-rank tensors:** $C^\mu{}_\nu(x)$ which transform as

$$C^\mu{}_\nu(x) \to C'^\mu{}_\nu(x') = \frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x^\sigma}{\partial x'^\nu}C^\rho{}_\sigma(x). \tag{3.9}$$

An example of a contravariant second-rank tensor is the inverse of the metric. In terms of the coordinates $\{\xi\}$ of the locally inertial frames postulated by the principle of equivalence and the lab-frame coordinates $\{x\}$, it reads

$$g^{\mu\nu}(x) = \eta^{\rho\sigma}\frac{\partial x^\mu}{\partial\xi^\rho}\frac{\partial x^\nu}{\partial\xi^\sigma}. \tag{3.10}$$

In a new coordinate system $\{x'\}$, we have

$$\begin{aligned} g'^{\mu\nu}(x') &= \eta^{\rho\sigma}\frac{\partial x'^{\mu}}{\partial \xi^{\rho}}\frac{\partial x'^{\nu}}{\partial \xi^{\sigma}} = \eta^{\rho\sigma}\left(\frac{\partial x'^{\mu}}{\partial x^{\alpha}}\frac{\partial x^{\alpha}}{\partial \xi^{\rho}}\right)\left(\frac{\partial x'^{\nu}}{\partial x^{\beta}}\frac{\partial x^{\beta}}{\partial \xi^{\sigma}}\right) \\ &= \left(\eta^{\rho\sigma}\frac{\partial x^{\alpha}}{\partial \xi^{\rho}}\frac{\partial x^{\beta}}{\partial \xi^{\sigma}}\right)\frac{\partial x'^{\mu}}{\partial x^{\alpha}}\frac{\partial x'^{\nu}}{\partial x^{\beta}} \\ &= g^{\alpha\beta}(x)\frac{\partial x'^{\mu}}{\partial x^{\alpha}}\frac{\partial x'^{\nu}}{\partial x^{\beta}}. \end{aligned} \tag{3.11}$$

An example of a covariant second-rank tensor is the metric,

$$g_{\mu\nu}(x) = \eta_{\alpha\beta}\frac{\partial \xi^{\alpha}}{\partial x^{\mu}}\frac{\partial \xi^{\beta}}{\partial x^{\nu}}. \tag{3.12}$$

In a new coordinate system, we have

$$g'_{\mu\nu}(x') = \eta_{\alpha\beta}\frac{\partial \xi^{\alpha}}{\partial x'^{\mu}}\frac{\partial \xi^{\beta}}{\partial x'^{\nu}} = \cdots = g_{\rho\sigma}(x)\frac{\partial x^{\rho}}{\partial x'^{\mu}}\frac{\partial x^{\sigma}}{\partial x'^{\nu}}, \tag{3.13}$$

following the same steps as for the inverse metric. This is the transformation of a covariant second-rank tensor, as anticipated.

An example of a mixed second-rank tensor is the Kronecker delta symbol. We can represent the Kronecker delta as

$$\delta^{\mu}{}_{\nu} \equiv \frac{\partial x^{\mu}}{\partial x^{\nu}}. \tag{3.14}$$

In a new coordinate system,

$$\delta'^{\mu}{}_{\nu} = \frac{\partial x'^{\mu}}{\partial x'^{\nu}} = \frac{\partial x'^{\mu}}{\partial x^{\rho}}\frac{\partial x^{\rho}}{\partial x'^{\nu}} = \delta^{\rho}{}_{\sigma}\frac{\partial x'^{\mu}}{\partial x^{\rho}}\frac{\partial x^{\sigma}}{\partial x'^{\nu}}. \tag{3.15}$$

Of course, numerically we obtain $\delta'^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu}$. However, this numerical equality should not hide the transformation property, which is that of a mixed second-rank tensor (rather than that of a scalar).

Similarly, we can define tensors of arbitrary rank

$$T^{\alpha_1\alpha_2\cdots}{}_{\beta_1\beta_2\cdots}(x), \tag{3.16}$$

which transform as

$$T'^{\alpha_1\alpha_2\cdots}{}_{\beta_1\beta_2\cdots}(x') = \frac{\partial x'^{\alpha_1}}{\partial x^{\rho_1}}\frac{\partial x'^{\alpha_2}}{\partial x^{\rho_2}}\cdots\frac{\partial x^{\sigma_1}}{\partial x'^{\beta_1}}\frac{\partial x^{\sigma_2}}{\partial x'^{\beta_2}}\cdots T^{\rho_1\rho_2\cdots}{}_{\sigma_1\sigma_2\cdots}(x). \tag{3.17}$$

The rank of a tensor equals the number of its free (not summed over) indices. Conventionally, we write the indices α which induce a contravariant transformation factor $\frac{\partial x'^{\alpha}}{\partial x^{\rho}}$ in an upper position. Similarly, we write the indices β which induce a covariant transformation factor $\frac{\partial x^{\sigma}}{\partial x'^{\beta}}$ at a lower position.

Setting an upper and a lower index equal and summing over their values is called a contraction. A contraction gives a tensor of lower rank. For example,

$$\begin{aligned} T^{\mu\nu}{}_{\mu} \to T'^{\mu\nu}{}_{\mu} &= \frac{\partial x'^{\mu}}{\partial x^{\rho}}\frac{\partial x'^{\nu}}{\partial x^{\sigma}}\frac{\partial x^{\alpha}}{\partial x'^{\mu}}T^{\rho\sigma}{}_{\alpha} \\ &= \frac{\partial x'^{\nu}}{\partial x^{\sigma}}\left(\frac{\partial x'^{\mu}}{\partial x^{\rho}}\frac{\partial x^{\alpha}}{\partial x'^{\mu}}\right)T^{\rho\sigma}{}_{\alpha} = \frac{\partial x'^{\nu}}{\partial x^{\sigma}}\frac{\partial x^{\alpha}}{\partial x^{\rho}}T^{\rho\sigma}{}_{\alpha} \\ &= \frac{\partial x'^{\nu}}{\partial x^{\sigma}}\delta^{\alpha}{}_{\rho}T^{\rho\sigma}{}_{\alpha} = \frac{\partial x'^{\nu}}{\partial x^{\sigma}}T^{\rho\sigma}{}_{\rho}. \end{aligned} \tag{3.18}$$

This is the transformation of a rank-1 tensor (a vector) rather than a rank-3 tensor.

Lowering and raising indices

Let us multiply a contravariant vector V^μ with the metric, $g_{\rho\sigma}V^\mu$, and contract the index μ of the vector with one of the indices of the metric. The result is a covariant vector,

$$V_\sigma = g_{\mu\sigma}V^\mu, \tag{3.19}$$

since

$$\begin{aligned} V'_\sigma = g'_{\mu\sigma}V'^\mu &= \left(g_{\rho\alpha}\frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x^\alpha}{\partial x'^\sigma}\right)\left(\frac{\partial x'^\mu}{\partial x^\delta}V^\delta\right) \\ &= g_{\rho\alpha}V^\delta \underbrace{\frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x'^\mu}{\partial x^\delta}}_{=\delta^\rho{}_\delta}\frac{\partial x^\alpha}{\partial x'^\sigma} = g_{\rho\alpha}V^\rho\frac{\partial x^\alpha}{\partial x'^\sigma} \\ &= V_\alpha\frac{\partial x^\alpha}{\partial x'^\sigma}. \end{aligned} \tag{3.20}$$

Similarly, we can contract a covariant vector with the inverse of the metric to construct its contravariant dual,

$$U^\mu \equiv g^{\mu\nu}U_\nu, \tag{3.21}$$

which transforms as

$$\begin{aligned} U'^\mu = g'^{\mu\nu}U'_\nu &= g^{\rho\sigma}\frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x'^\nu}{\partial x^\sigma}\frac{\partial x^\gamma}{\partial x'^\nu}U_\gamma \\ &= g^{\rho\sigma}U_\sigma\frac{\partial x'^\mu}{\partial x^\rho} = U^\rho\frac{\partial x'^\mu}{\partial x^\rho}. \end{aligned} \tag{3.22}$$

Finally, we remark that we can use the metric tensor or its inverse to lower or raise indices in arbitrary tensors. This modifies their transformation properties under general coordinate transformations. but it does not alter their physical content. The 2^N versions of a rank-N tensor obtained in this way are called duals of each other:

$$T^{\alpha_1}{}_{\alpha_2}{}^{\alpha_3\alpha_4\cdots}{}_{\ldots}, \; T^{\alpha_1\alpha_2}{}_{\alpha_3}{}^{\alpha_4\cdots}{}_{\ldots}, \cdots, \; T_{\alpha_1\alpha_2\alpha_3\alpha_4\cdots} \tag{3.23}$$

3.2 Tensor densities

Consider (minus) the determinant of the metric,

$$g \equiv -\det\left(g_{\mu\nu}\right). \tag{3.24}$$

Under a general coordinate transformation, it transforms as

$$\begin{aligned} g \to g' = -\det\left(g'_{\mu\nu}\right) &= -\det\left(\frac{\partial x^\rho}{\partial x'^\mu}g_{\rho\sigma}\frac{\partial x^\sigma}{\partial x'^\nu}\right) \\ &= -\det\left(g_{\rho\sigma}\right)\det\left(\frac{\partial x^\rho}{\partial x'^\mu}\right)^2 = g\left|\frac{\partial x}{\partial x'}\right|^2 \\ &= g\left|\frac{\partial x'}{\partial x}\right|^{-2}, \end{aligned} \tag{3.25}$$

where $\left|\frac{\partial x'}{\partial x}\right|$ is the Jacobian of the general coordinate transformation. Hence, the determinant of the metric is not a scalar. A quantity which transforms as a scalar up to a power of the Jacobian is called a scalar density. Generically, a scalar density $\underline{O}$ transforms as

$$\underline{O} \to \underline{O}' = \underline{O} \left|\frac{\partial x'}{\partial x}\right|^w , \tag{3.26}$$

where the power w is called the weight of the scalar density. **Remark.** With this convention g has weight -2, $\sqrt{g}$ weight -1, d^4x weight $+1$, the Levi-Civita symbol $\varepsilon^{\mu\nu\rho\sigma}$ (numerically the same in all frames) weight -1 and $\varepsilon_{\mu\nu\rho\sigma}$ weight $+1$, as in Weinberg Sec. 4.4. Textbooks differ in the sign convention for the weight, so the numbers should be compared with care. Similarly, there are objects which transform as tensors up to a power of the Jacobian,

$$\underline{O}^{\mu_1\cdots}{}_{\nu_1\cdots} \to \underline{O}'^{\mu_1\cdots}{}_{\nu_1\cdots} = \frac{\partial x'^{\mu_1}}{\partial x^{\alpha_1}} \cdots \frac{\partial x^{\beta_1}}{\partial x'^{\nu_1}} \cdots \left|\frac{\partial x'}{\partial x}\right|^w \underline{O}^{\alpha_1\cdots}{}_{\beta_1\cdots}. \tag{3.27}$$

The above defines a tensor density of weight w. Notice that we can turn a tensor density into an ordinary tensor by multiplying it with an appropriate power of the determinant of the metric. Concretely,

$$\begin{aligned} g^{w/2}\underline{O}^{\mu_1\cdots}{}_{\nu_1\cdots} \to & g'^{w/2}\underline{O}'^{\mu_1\cdots}{}_{\nu_1\cdots} \\ &= g^{w/2}\left(\left|\frac{\partial x'}{\partial x}\right|^{-2}\right)^{w/2} \underline{O}^{\alpha_1\cdots}{}_{\beta_1\cdots} \left|\frac{\partial x'}{\partial x}\right|^w \left(\frac{\partial x'^{\mu_1}}{\partial x^{\alpha_1}}\right)\cdots\left(\frac{\partial x^{\beta_1}}{\partial x'^{\nu_1}}\right)\cdots \\ &= g^{w/2}\underline{O}^{\alpha_1\cdots}{}_{\beta_1\cdots} \left(\frac{\partial x'^{\mu_1}}{\partial x^{\alpha_1}}\right)\cdots\left(\frac{\partial x^{\beta_1}}{\partial x'^{\nu_1}}\right)\cdots . \end{aligned} \tag{3.28}$$

We observe that the 4-dimensional volume element is a scalar density of weight 1. Indeed,

$$\mathrm{d}^4x \equiv \mathrm{d}x^0\mathrm{d}x^1\mathrm{d}x^2\mathrm{d}x^3 \to \boxed{\mathrm{d}^4x' = \left|\frac{\partial x'}{\partial x}\right| \mathrm{d}^4x}. \tag{3.29}$$

The combination $\sqrt{g}\mathrm{d}^4x$ is invariant under general coordinate transformations. Indeed,

$$\sqrt{g}\mathrm{d}^4x \to \sqrt{g'}\mathrm{d}^4x' = \sqrt{g\left|\frac{\partial x'}{\partial x}\right|^{-2}}\mathrm{d}^4x \left|\frac{\partial x'}{\partial x}\right| = \sqrt{g}\mathrm{d}^4x. \tag{3.30}$$

The fully antisymmetric Levi-Civita symbol is also a tensor density. Consider the quantity

$$\varepsilon^{\alpha\beta\gamma\delta}\frac{\partial x'^{\mu}}{\partial x^{\alpha}}\frac{\partial x'^{\nu}}{\partial x^{\beta}}\frac{\partial x'^{\rho}}{\partial x^{\gamma}}\frac{\partial x'^{\sigma}}{\partial x^{\delta}} = \mathrm{const}\cdot\varepsilon^{\mu\nu\rho\sigma}. \tag{3.31}$$

For $\mu\nu\rho\sigma = 0,1,2,3$ we obtain

$$\mathrm{const} = \varepsilon^{\alpha\beta\gamma\delta}\frac{\partial x'^{0}}{\partial x^{\alpha}}\frac{\partial x'^{1}}{\partial x^{\beta}}\frac{\partial x'^{2}}{\partial x^{\gamma}}\frac{\partial x'^{3}}{\partial x^{\delta}} = \det\left(\frac{\partial x'}{\partial x}\right) = \left|\frac{\partial x'}{\partial x}\right|. \tag{3.32}$$

Thus,

$$\varepsilon^{\alpha\beta\gamma\delta}\frac{\partial x'^{\mu}}{\partial x^{\alpha}}\frac{\partial x'^{\nu}}{\partial x^{\beta}}\frac{\partial x'^{\rho}}{\partial x^{\gamma}}\frac{\partial x'^{\sigma}}{\partial x^{\delta}} = \left|\frac{\partial x'}{\partial x}\right|\varepsilon^{\mu\nu\rho\sigma}. \tag{3.33}$$

From the above we conclude that the Levi-Civita symbol is a tensor density.

Exercise: Prove that $\varepsilon^{\mu\nu\rho\sigma}/\sqrt{g}$ and $\sqrt{g}\,\varepsilon_{\mu\nu\rho\sigma}$ are tensors.

3.3 Transformation of the affine connection

The affine connection is given by

$$\Gamma^{\alpha}{}_{\mu\nu} = \frac{\partial x^{\alpha}}{\partial \xi^{\rho}} \frac{\partial^2 \xi^{\rho}}{\partial x^{\mu} \partial x^{\nu}}. \tag{3.34}$$

Under a general coordinate transformation $x \to x'$, we have

$$\begin{aligned}
\Gamma'^{\alpha}{}_{\mu\nu} &= \frac{\partial x'^{\alpha}}{\partial \xi^{\rho}} \frac{\partial^2 \xi^{\rho}}{\partial x'^{\mu} \partial x'^{\nu}} = \frac{\partial x'^{\alpha}}{\partial \xi^{\rho}} \left[\frac{\partial}{\partial x'^{\mu}} \frac{\partial \xi^{\rho}}{\partial x'^{\nu}} \right] \\
&= \frac{\partial x'^{\alpha}}{\partial \xi^{\rho}} \left[\frac{\partial}{\partial x'^{\mu}} \left(\frac{\partial x^{\beta}}{\partial x'^{\nu}} \frac{\partial \xi^{\rho}}{\partial x^{\beta}} \right) \right] \\
&= \frac{\partial x'^{\alpha}}{\partial \xi^{\rho}} \frac{\partial x^{\beta}}{\partial x'^{\nu}} \frac{\partial}{\partial x'^{\mu}} \frac{\partial \xi^{\rho}}{\partial x^{\beta}} + \frac{\partial x'^{\alpha}}{\partial \xi^{\rho}} \frac{\partial \xi^{\rho}}{\partial x^{\beta}} \frac{\partial^2 x^{\beta}}{\partial x'^{\mu} \partial x'^{\nu}} \\
&= \underbrace{\frac{\partial x^{\delta}}{\partial \xi^{\rho}} \frac{\partial^2 \xi^{\rho}}{\partial x^{\gamma} \partial x^{\beta}}}_{=\Gamma^{\delta}{}_{\gamma\beta}} \underbrace{\frac{\partial x'^{\alpha}}{\partial x^{\delta}} \frac{\partial x^{\beta}}{\partial x'^{\nu}} \frac{\partial x^{\gamma}}{\partial x'^{\mu}}}_{\substack{\text{tensor-like part} \\ \text{of transformation}}} + \underbrace{\frac{\partial x'^{\alpha}}{\partial x^{\beta}} \frac{\partial^2 x^{\beta}}{\partial x'^{\mu} \partial x'^{\nu}}}_{\substack{\text{non-tensorial part} \\ \text{of transformation}}}.
\end{aligned} \tag{3.35}$$

So, the affine connection is not a tensor. Under a general coordinate transformation it changes as

$$\Gamma^{\alpha}{}_{\mu\nu} \to \Gamma'^{\alpha}{}_{\mu\nu} = \frac{\partial x'^{\alpha}}{\partial x^{\rho}} \frac{\partial x^{\beta}}{\partial x'^{\mu}} \frac{\partial x^{\gamma}}{\partial x'^{\nu}} \Gamma^{\rho}{}_{\beta\gamma} + \frac{\partial x'^{\alpha}}{\partial x^{\beta}} \frac{\partial^2 x^{\beta}}{\partial x'^{\mu} \partial x'^{\nu}}. \tag{3.36}$$

This non-tensorial transformation of the affine connection is needed to make the geodesic equation for a particle in a gravitational field generally covariant. Indeed, one can show that the combination

$$\frac{\mathrm{d}^2 x^{\mu}}{\mathrm{d}\tau^2} + \Gamma^{\mu}{}_{\alpha\beta} \frac{\mathrm{d}x^{\alpha}}{\mathrm{d}\tau} \frac{\mathrm{d}x^{\beta}}{\mathrm{d}\tau}, \tag{3.37}$$

transforms as a vector (Exercise).

It is easier to prove the above if we derive an equivalent formula for the transformation of the affine connection. We have

$$\begin{aligned}
\frac{\partial x'^{\alpha}}{\partial x'^{\mu}} = \delta^{\alpha}{}_{\mu} &= \frac{\partial x'^{\alpha}}{\partial x^{\beta}} \frac{\partial x^{\beta}}{\partial x'^{\mu}} \\
\to \frac{\partial}{\partial x'^{\nu}} \left(\frac{\partial x'^{\alpha}}{\partial x^{\beta}} \right) \frac{\partial x^{\beta}}{\partial x'^{\mu}} &= - \frac{\partial x'^{\alpha}}{\partial x^{\beta}} \frac{\partial^2 x^{\beta}}{\partial x'^{\mu} \partial x'^{\nu}} \\
\Rightarrow - \frac{\partial^2 x'^{\alpha}}{\partial x^{\beta} \partial x^{\gamma}} \frac{\partial x^{\beta}}{\partial x'^{\mu}} \frac{\partial x^{\gamma}}{\partial x'^{\nu}} &= \frac{\partial x'^{\alpha}}{\partial x^{\beta}} \frac{\partial^2 x^{\beta}}{\partial x'^{\mu} \partial x'^{\nu}},
\end{aligned} \tag{3.38}$$

and thus

$$\Gamma^{\alpha}{}_{\mu\nu} \to \Gamma'^{\alpha}{}_{\mu\nu} = \frac{\partial x'^{\alpha}}{\partial x^{\rho}} \frac{\partial x^{\beta}}{\partial x'^{\mu}} \frac{\partial x^{\gamma}}{\partial x'^{\nu}} \Gamma^{\rho}{}_{\beta\gamma} - \frac{\partial x^{\beta}}{\partial x'^{\mu}} \frac{\partial x^{\gamma}}{\partial x'^{\nu}} \frac{\partial^2 x'^{\alpha}}{\partial x^{\beta} \partial x^{\gamma}}. \tag{3.39}$$

Let us now contract the connection with a vector:

$$\begin{aligned}
\Gamma^{\alpha}{}_{\mu\nu} V^{\nu} &\to \frac{\partial x'^{\alpha}}{\partial x^{\rho}} \frac{\partial x^{\sigma}}{\partial x'^{\mu}} \Gamma^{\rho}{}_{\sigma\nu} V^{\nu} - \frac{\partial x^{\beta}}{\partial x'^{\mu}} \frac{\partial x^{\gamma}}{\partial x'^{\nu}} \frac{\partial^2 x'^{\alpha}}{\partial x^{\beta} \partial x^{\gamma}} \frac{\partial x'^{\nu}}{\partial x^{\delta}} V^{\delta} \\
&= \frac{\partial x'^{\alpha}}{\partial x^{\rho}} \frac{\partial x^{\sigma}}{\partial x'^{\mu}} \Gamma^{\rho}{}_{\sigma\nu} V^{\nu} - \frac{\partial x^{\beta}}{\partial x'^{\mu}} \frac{\partial^2 x'^{\alpha}}{\partial x^{\beta} \partial x^{\gamma}} V^{\gamma}.
\end{aligned} \tag{3.40}$$

3.4 Covariant derivative

Let us now look at the derivative of a contravariant vector,

$$\begin{aligned}\frac{\partial V^\alpha}{\partial x^\mu} \to \frac{\partial V'^\alpha}{\partial x'^\mu} &= \frac{\partial}{\partial x'^\mu}\left(\frac{\partial x'^\alpha}{\partial x^\gamma} V^\gamma\right) \\ &= \underbrace{\frac{\partial x^\beta}{\partial x'^\mu}\frac{\partial^2 x'^\alpha}{\partial x^\beta \partial x^\gamma} V^\gamma}_{\text{non-tensorial}} + \frac{\partial x'^\alpha}{\partial x^\gamma}\frac{\partial x^\beta}{\partial x'^\mu}\frac{\partial V^\gamma}{\partial x^\beta}.\end{aligned} \tag{3.41}$$

We then define

$$V^\alpha{}_{;\mu} \equiv \frac{\partial V^\alpha}{\partial x^\mu} + \Gamma^\alpha{}_{\mu\nu} V^\nu, \tag{3.42}$$

which transforms as a tensor,

$$\boxed{V'^\alpha{}_{;\mu} = \frac{\partial x'^\alpha}{\partial x^\beta}\frac{\partial x^\gamma}{\partial x'^\mu} V^\beta{}_{;\gamma}}. \tag{3.43}$$

This is the covariant derivative of a contravariant vector. For a covariant vector, we can find an analogous definition. We start from

$$\begin{aligned}\frac{\partial V_\alpha}{\partial x^\mu} \to \frac{\partial V'_\alpha}{\partial x'^\mu} &= \frac{\partial}{\partial x'^\mu}\left(\frac{\partial x^\beta}{\partial x'^\alpha} V_\beta\right) \\ &= \frac{\partial x^\beta}{\partial x'^\alpha}\frac{\partial x^\gamma}{\partial x'^\mu}\frac{\partial V_\beta}{\partial x^\gamma} + \frac{\partial^2 x^\beta}{\partial x'^\alpha \partial x'^\mu} V_\beta.\end{aligned} \tag{3.44}$$

The transformation of the contraction of the vector with the affine connection is

$$\Gamma^\beta{}_{\alpha\mu} V_\beta \to \Gamma'^\beta{}_{\alpha\mu} V'_\beta = \frac{\partial x^\gamma}{\partial x'^\alpha}\frac{\partial x^\sigma}{\partial x'^\mu}\Gamma^\delta{}_{\gamma\sigma} V_\delta + \frac{\partial x'^\beta}{\partial x^\delta}\frac{\partial^2 x^\delta}{\partial x'^\alpha \partial x'^\mu}\frac{\partial x^\rho}{\partial x'^\beta} V_\rho. \tag{3.45}$$

We then define

$$V_{\alpha;\mu} \equiv \frac{\partial V_\alpha}{\partial x^\mu} - \Gamma^\delta{}_{\alpha\mu} V_\delta, \tag{3.46}$$

which transforms as a tensor.

For a tensor, we can analogously construct a covariant derivative

$$T^\alpha{}_{\beta\gamma;\mu} = \frac{\partial}{\partial x^\mu} T^\alpha{}_{\beta\gamma} - \Gamma^\rho{}_{\beta\mu} T^\alpha{}_{\rho\gamma} - \Gamma^\rho{}_{\gamma\mu} T^\alpha{}_{\beta\rho} + \Gamma^\alpha{}_{\mu\rho} T^\rho{}_{\beta\gamma}. \tag{3.47}$$

For a contracted tensor we have

$$\begin{aligned}T^{\mu\lambda}{}_{\lambda;\sigma} &= \frac{\partial}{\partial x^\sigma} T^{\mu\lambda}{}_{\lambda} - \cancel{\Gamma^\alpha{}_{\lambda\sigma} T^{\mu\lambda}{}_{\alpha}} + \Gamma^\mu{}_{\sigma\alpha} T^{\alpha\lambda}{}_{\lambda} + \cancel{\Gamma^\lambda{}_{\alpha\sigma} T^{\mu\alpha}{}_{\lambda}} \\ &= \frac{\partial}{\partial x^\sigma} T^{\mu\lambda}{}_{\lambda} + \Gamma^\mu{}_{\sigma\alpha} T^{\alpha\lambda}{}_{\lambda}.\end{aligned} \tag{3.48}$$

Let us also note that

$$g_{\mu\nu;\rho} = g^{\mu\nu}{}_{;\rho} = \delta^\mu{}_{\nu;\rho} = 0. \tag{3.49}$$

These equations hold in locally inertial frames, where $g_{\mu\nu} = \eta_{\mu\nu}$ and the affine connection vanishes. Being tensor equations, they then hold in arbitrary gravitational fields. (Exercise: Check this explicitly!)

3.4.1 Covariant curl

For a covariant vector, the connection terms cancel in the antisymmetrised derivative,

$$\begin{aligned} V_{\mu;\nu} - V_{\nu;\mu} &= \frac{\partial V_\mu}{\partial x^\nu} - \cancel{\Gamma^\alpha_{\mu\nu} V_\alpha} - \frac{\partial V_\nu}{\partial x^\mu} + \cancel{\Gamma^\alpha_{\nu\mu} V_\alpha} \\ &= \frac{\partial V_\mu}{\partial x^\nu} - \frac{\partial V_\nu}{\partial x^\mu}. \end{aligned} \tag{3.50}$$

3.4.2 Covariant divergence

Consider a contravariant vector V^μ. Its covariant divergence is

$$V^\mu{}_{;\mu} = \frac{\partial V^\mu}{\partial x^\mu} + \Gamma^\mu{}_{\mu\rho} V^\rho. \tag{3.51}$$

The contracted affine connection is

$$\Gamma^\mu{}_{\mu\rho} = \frac{1}{2} g^{\mu\alpha} \left(\frac{\partial g_{\mu\alpha}}{\partial x^\rho} + \cancel{\frac{\partial g_{\rho\alpha}}{\partial x^\mu}} - \cancel{\frac{\partial g_{\mu\rho}}{\partial x^\alpha}} \right) = \frac{1}{2} g^{\mu\alpha} \frac{\partial g_{\mu\alpha}}{\partial x^\rho}. \tag{3.52}$$

This can be simplified further. With some linear algebra one can show that

$$\boxed{\Gamma^\mu{}_{\mu\rho} = \frac{1}{2} \frac{\partial}{\partial x^\rho} \log\left(-\det g_{\mu\nu}\right)} \tag{3.53}$$

Proof: Let us start with a matrix $M(x)$ and assume a variation $x \to x + \delta x$,

$$\begin{aligned} \delta \log\det M(x) &= \log(\det M(x+\delta x)) - \log\det M(x) \\ &= \log \frac{\det(M(x+\delta x))}{\det M(x)} = \log\det\left(M^{-1}(x) M(x+\delta x)\right) \\ &= \mathrm{Tr}\left(\log\left(M^{-1}(x) M(x+\delta x)\right)\right) \\ &= \mathrm{Tr}\left(\log\left(M^{-1}(x)(M(x) + \frac{\partial M(x)}{\partial x}\delta x + \cdots)\right)\right) \\ &= \mathrm{Tr}\left(\log\left(1 + M^{-1}(x)\frac{\partial M(x)}{\partial x}\delta x\right)\right) \\ &= \mathrm{Tr}\left(M^{-1}(x)\frac{\partial M(x)}{\partial x}\right)\delta x. \end{aligned} \tag{3.54}$$

The left-hand side is also

$$\begin{aligned} \delta \log\det(M(x)) &= \frac{\partial}{\partial x} \log\det(M(x))\delta x \\ &= \frac{1}{\det M(x)} \frac{\partial}{\partial x}(\det M(x))\delta x = \frac{\partial}{\partial x} \log(\det M(x))\, \delta x. \end{aligned} \tag{3.55}$$

Thus,

$$\det\left(M^{-1}(x)\right) \frac{\partial}{\partial x} \det(M(x)) = \mathrm{Tr}\left(M^{-1}(x) \frac{\partial}{\partial x} M(x)\right). \tag{3.56}$$

For the metric matrix $g_{\mu\nu}(x)$ the inverse is $g^{\mu\nu}(x)$. Taking M to be the metric, we have

$$\begin{aligned} \det g^{\mu\nu} \frac{\partial}{\partial x^\rho} \det\left(g_{\mu\nu}\right) &= g^{\mu\alpha} \frac{\partial}{\partial x^\rho} g_{\alpha\mu} \\ \Rightarrow \Gamma^\mu{}_{\mu\rho} = \frac{1}{2} g^{\mu\alpha} \frac{\partial}{\partial x^\rho} g_{\alpha\mu} &= \frac{1}{2 \det g_{\mu\nu}} \frac{\partial}{\partial x^\rho} \det g_{\mu\nu} = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^\rho} \sqrt{g}, \end{aligned} \tag{3.57}$$

where $g \equiv -\det g_{\mu\nu}$. Thus, the divergence becomes

$$\begin{aligned} V^\mu{}_{;\mu} &= \frac{\partial V^\mu}{\partial x^\mu} + \Gamma^\mu{}_{\mu\rho} V^\rho \\ &= \frac{\partial V^\mu}{\partial x^\mu} + \frac{1}{\sqrt{g}} \left(\frac{\partial}{\partial x^\rho} \sqrt{g} \right) V^\rho \\ \Rightarrow \boxed{V^\mu{}_{;\mu} &= \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^\mu} \left(\sqrt{g} V^\mu \right)} \end{aligned} \tag{3.58}$$

We can now deduce a covariant form of Gauss's theorem. If V^μ vanishes at infinity, we have

$$\begin{aligned} &\int \mathrm{d}^4 x \partial_\mu (\sqrt{g} V^\mu) = 0 \\ \Rightarrow &\boxed{\int \underbrace{\mathrm{d}^4 x \sqrt{g}}_{\text{invariant}} \underbrace{V^\mu{}_{;\mu}}_{\text{cov. deriv.}} = 0} \end{aligned} \tag{3.59}$$

Exercise: How does $V^\mu{}_{;\mu}$ transform under general coordinate transformations?

3.4.3 Covariant differentiation along a curve

Often we are interested in properties of particles moving along a trajectory in spacetime, which are then defined only along this curve. Therefore, we need a covariant derivative for tensors defined only on a curve, as opposed to everywhere in space.

For example, consider a quantity $\Pi^{\mu\nu\rho}(x(\tau))$ which is defined on a trajectory $x(\tau)$ through spacetime (Fig. 3.1). How would we compute its derivative?
In classical mechanics we worked with $\vec{x}(t)$, $\frac{\vec{x}(t+\delta t)-\vec{x}(t)}{\delta t}$, or in special relativity with $\vec{x}(\tau)$, $t(\tau)$. An appropriate derivative for a vector $V^\mu(x^\alpha(\tau))$ living on the trajectory was then

$$\frac{\mathrm{d}V^\mu(\tau)}{\mathrm{d}\tau}. \tag{3.60}$$

Problem: this is not covariant!

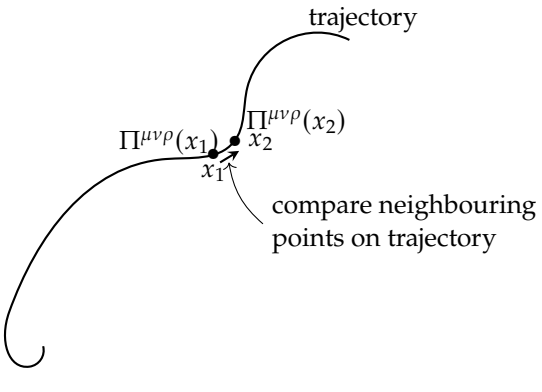


Figure 3.1. A tensor defined only along a trajectory. Its derivative compares neighbouring points on the curve.

We thus introduce the covariant derivative along a curve

$$\frac{\mathrm{D}A^\mu}{\mathrm{D}\tau} \equiv \frac{\mathrm{d}A^\mu}{\mathrm{d}\tau} + \Gamma^\mu{}_{\nu\lambda} \frac{\mathrm{d}x^\lambda}{\mathrm{d}\tau} A^\nu(\tau). \tag{3.61}$$

Similarly, we have

$$\frac{\mathrm{D}A_\mu}{\mathrm{D}\tau} \equiv \frac{\mathrm{d}A_\mu}{\mathrm{d}\tau} - \Gamma^\lambda{}_{\mu\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}A_\lambda, \tag{3.62}$$

for a covariant vector. The generalisation to tensors with several upper and lower indices is analogous to the ordinary covariant derivative.

Exercise: Check that $\frac{\mathrm{d}A^\mu}{\mathrm{d}\tau}$ and $\frac{\mathrm{d}A_\mu}{\mathrm{d}\tau}$ are not vectors, but that $\frac{\mathrm{D}A^\mu}{\mathrm{D}\tau}$ is.

4

Writing laws of physics consistently in general relativity

In Chapter 3 we studied the mathematical properties of tensors in general relativity. Tensors allow us to generalise the laws of special relativity to include gravity, by casting the corresponding equations in a form which is covariant under general coordinate transformations. We will show how this can be achieved in particle mechanics and in electrodynamics.

4.1 Particle mechanics

4.1.1 Acceleration

In the absence of gravity, for an inertial observer with coordinates $\{\xi^\mu\}$, the four-velocity of a particle is

$$V^\mu = \frac{\mathrm{d}\xi^\mu}{\mathrm{d}\tau}. \tag{4.1}$$

For a free particle, the equations of motion read

$$\frac{\mathrm{d}V^\mu}{\mathrm{d}\tau} = 0. \tag{4.2}$$

In the presence of gravity, we work in a lab frame with coordinates $\{x^\alpha\}$. There, we can also define a four-velocity, which is a vector under general coordinate transformations,

$$V^\mu \equiv \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}. \tag{4.3}$$

Equation (4.2) is not covariant under general coordinate transformations. We can make it covariant by replacing the derivative with a covariant derivative. An appropriate equation reads

$$\begin{aligned} & \frac{\mathrm{D}V^\mu}{\mathrm{D}\tau} = 0 \\ \Leftrightarrow\ & \frac{\mathrm{d}V^\mu}{\mathrm{d}\tau} + \Gamma^\mu{}_{\rho\sigma} V^\rho \frac{\mathrm{d}x^\sigma}{\mathrm{d}\tau} = 0 \\ \Leftrightarrow\ & \boxed{\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu{}_{\rho\sigma} \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} \frac{\mathrm{d}x^\sigma}{\mathrm{d}\tau} = 0} \end{aligned} \tag{4.4}$$

which is the equation we found earlier by a more direct use of the principle of equivalence.

It is easy to verify that the above equation $\left(\frac{DV^\mu}{D\tau} = 0\right)$ satisfies the principle of equivalence. Indeed, if we perform a coordinate transformation to a locally inertial frame, in which gravity is cancelled at a given point, it reduces to (4.2) ($g_{\mu\nu} \to \eta_{\mu\nu}$, $\Gamma^\mu{}_{\rho\sigma} \to 0$). The existence of such a frame is guaranteed by the principle of equivalence, while the covariance of (4.4) under general coordinate transformations guarantees that the equation keeps its form under this transformation. **Remark.** Mathematically, for any metric one can reach $g_{\mu\nu} = \eta_{\mu\nu}$ and $\Gamma^\mu_{\rho\sigma} = 0$ at a chosen point (e.g. with Riemann normal coordinates). However, $\Gamma \to 0$ holds only at that point, not in a finite region (Weinberg Sec. 4.1).

4.1.2 Spin

In special relativity the spin of a free particle is conserved for an inertial observer,

$$\frac{dS_\mu}{d\tau} = 0. \tag{4.5}$$

In the rest frame of the particle, the spin four-vector takes the form $S_\mu = (0, \vec{S})$. Thus $S_0 = 0$, a condition which can be written covariantly, and hence in any inertial frame, as

$$S_\mu V^\mu = 0, \tag{4.6}$$

which in the rest frame ($d\tau = dt$) gives

$$S_0 \frac{dt}{d\tau} + S_i \frac{\overset{0}{\cancel{dx^i}}}{d\tau} = 0 \quad \Rightarrow \quad S_0 = 0. \tag{4.7}$$

In the absence of gravity, the magnitude of the spin is also conserved,

$$\begin{aligned} \frac{dS_\mu}{d\tau} = 0 \;&\to\; S^\mu \frac{dS_\mu}{d\tau} = 0 \\ \to\; \frac{d}{d\tau}(S_\mu S^\mu) &= 0. \end{aligned} \tag{4.8}$$

In the presence of gravity, we can generalise (4.5) as

$$\begin{aligned} &\frac{DS_\mu}{D\tau} = 0 \\ \Leftrightarrow\; &\frac{dS_\mu}{d\tau} - \underbrace{\Gamma^\rho{}_{\mu\nu}}_{\text{effects of gravity}} S_\rho V^\nu = 0. \end{aligned} \tag{4.9}$$

For the magnitude of the spin we have

$$\frac{D}{D\tau}(S_\alpha S^\alpha) = 0. \tag{4.10}$$

As $S_\alpha S^\alpha$ is a scalar, the covariant derivative reduces to an ordinary derivative. So, the above simplifies to $\frac{d}{d\tau}(S_\alpha S^\alpha) = 0$. Thus, we conclude that gravity changes the direction of the spin of a falling particle, but not its magnitude (spin precession).

4.2 Electrodynamics

Let us now look at Maxwell's equations.

(A) In the absence of gravity, for an inertial observer, we have

$$\vec{\nabla}\cdot\vec{E} = \rho, \tag{4.11}$$

$$\vec{\nabla}\times\vec{B} = \frac{\partial\vec{E}}{\partial t} + \vec{j}, \tag{4.12}$$

$$\vec{\nabla}\cdot\vec{B} = 0, \tag{4.13}$$

$$\vec{\nabla}\times\vec{E} = -\frac{\partial\vec{B}}{\partial t}. \tag{4.14}$$

The $\vec{E}$, $\vec{B}$ fields make up the components of the field strength tensor

$$F^{\mu\nu} : \begin{cases} E_i = F^{0i}, \\ B_k = \frac{1}{2}\varepsilon_{ijk}F^{ij} \text{ (e.g. } B_2 = F^{31},\ B_3 = F^{12}), \end{cases} \tag{4.15}$$

with $F^{\mu\nu} = -F^{\nu\mu}$. Also $J^\mu \equiv (\rho, \vec{j})$. Substituting (4.15) into (4.11), we have

$$\begin{aligned} \frac{\partial}{\partial x^i}E_i = J^0 \ &\Rightarrow\ \frac{\partial}{\partial x^i}F^{0i} = J^0 \\ \Rightarrow\ \frac{\partial \cancelto{0}{F^{00}}}{\partial x^0} + \frac{\partial}{\partial x^i}F^{0i} = J^0 \ &\Rightarrow\ \boxed{\frac{\partial F^{0\mu}}{\partial x^\mu} = J^0}. \end{aligned} \tag{4.16}$$

Similarly, (4.12) becomes

$$\begin{aligned} \varepsilon_{ijk}\frac{\partial}{\partial x^j}\left(\frac{1}{2}\varepsilon_{klm}F^{lm}\right) &= +\frac{\partial F^{0i}}{\partial x^0} + J^i \\ \Rightarrow \begin{vmatrix}\delta_{il} & \delta_{im}\\ \delta_{jl} & \delta_{jm}\end{vmatrix}\frac{\partial}{\partial x^j}\frac{F^{lm}}{2} &= +\frac{\partial F^{0i}}{\partial x^0} + J^i \\ \Rightarrow\ \frac{\partial}{\partial x^j}\left(\frac{F^{ij}-F^{ji}}{2}\right) &= -\frac{\partial F^{i0}}{\partial x^0} + J^i \\ \Rightarrow\ \frac{\partial F^{i0}}{\partial x^0} + \frac{\partial F^{ij}}{\partial x^j} = J^i \ &\Rightarrow\ \boxed{\frac{\partial F^{i\mu}}{\partial x^\mu} = J^i}. \end{aligned} \tag{4.17}$$

Putting the last two results together, we have

$$\boxed{\frac{\partial F^{\mu\nu}}{\partial x^\mu} = -J^\nu} \tag{4.18}$$

Equation (4.13) becomes

$$\begin{aligned} 0 = \vec{\nabla}\cdot\vec{B} = \frac{\partial}{\partial x^i}B_i &= \frac{1}{2}\frac{\partial}{\partial x^i}\varepsilon_{ijk}F^{jk} \\ 0 &= \varepsilon_{ijk}\frac{\partial}{\partial x^i}F^{jk} \\ \Rightarrow\ 0 &= \frac{\partial}{\partial x^1}F^{23} + \frac{\partial}{\partial x^2}F^{31} + \frac{\partial}{\partial x^3}F^{12}. \end{aligned} \tag{4.19}$$

Equation (4.14) gives an analogous result. Collectively, (4.13) and (4.14) become

$$\boxed{\frac{\partial}{\partial x^\mu}F_{\nu\rho} + \frac{\partial}{\partial x^\nu}F_{\rho\mu} + \frac{\partial}{\partial x^\rho}F_{\mu\nu} = 0} \tag{4.20}$$

(B) We will now attempt to generalise (4.18) and (4.20) in the presence of gravity. Our method is to replace derivatives by covariant derivatives, and the Minkowski metric by the metric $g_{\mu\nu}$. The Maxwell equations thus become

$$F^{\mu\nu}{}_{;\mu} = -J^\nu, \tag{4.21}$$

$$F_{\mu\nu;\rho} + F_{\nu\rho;\mu} + F_{\rho\mu;\nu} = 0. \tag{4.22}$$

The second equation simplifies because the field strength tensor is antisymmetric. Writing out the covariant derivatives, we have

$$\begin{aligned} &\frac{\partial}{\partial x^\rho}F_{\mu\nu} + \frac{\partial}{\partial x^\mu}F_{\nu\rho} + \frac{\partial}{\partial x^\nu}F_{\rho\mu} \\ &\quad - \Gamma^\alpha{}_{\nu\rho}F_{\mu\alpha} - \Gamma^\alpha{}_{\mu\rho}F_{\alpha\nu} - \Gamma^\alpha{}_{\mu\rho}F_{\nu\alpha} - \Gamma^\alpha{}_{\nu\mu}F_{\alpha\rho} - \Gamma^\alpha{}_{\rho\nu}F_{\alpha\mu} - \Gamma^\alpha{}_{\mu\nu}F_{\rho\alpha} = 0 \\ \Rightarrow &\boxed{\frac{\partial}{\partial x^\rho}F_{\mu\nu} + \frac{\partial}{\partial x^\mu}F_{\nu\rho} + \frac{\partial}{\partial x^\nu}F_{\rho\mu} = 0} \end{aligned} \tag{4.23}$$

The divergence of the field strength tensor is

$$\begin{aligned} F^{\mu\nu}{}_{;\mu} &= \frac{\partial F^{\mu\nu}}{\partial x^\mu} + \Gamma^\mu{}_{\mu\rho}F^{\rho\nu} + \underbrace{\Gamma^\nu{}_{\mu\rho}\cancel{F^{\mu\rho}}}_{\text{sym.}\times\text{antisym.}} 0 \\ &= \frac{\partial F^{\mu\nu}}{\partial x^\mu} + \Gamma^\mu{}_{\mu\rho}F^{\rho\nu}. \end{aligned} \tag{4.24}$$

Then, analogously to what we did for the divergence of a contravariant vector, we have

$$F^{\mu\nu}{}_{;\mu} = \frac{1}{\sqrt{g}}\frac{\partial}{\partial x^\mu}\left(\sqrt{g}F^{\mu\nu}\right), \tag{4.25}$$

and the remaining Maxwell equations become

$$\begin{aligned} F^{\mu\nu}{}_{;\mu} &= -J^\nu \\ \rightarrow \frac{1}{\sqrt{g}}\frac{\partial}{\partial x^\mu}\left(\sqrt{g}F^{\mu\nu}\right) &= -J^\nu \\ \rightarrow \boxed{\frac{\partial}{\partial x^\mu}\left(\sqrt{g}F^{\mu\nu}\right) = -\sqrt{g}J^\nu} \end{aligned} \tag{4.26}$$

Acting with an ordinary derivative on the above, we find

$$\begin{aligned} \underbrace{\frac{\partial^2}{\partial x^\nu \partial x^\mu}}_{\text{symm.}}\underbrace{\left(\sqrt{g}F^{\mu\nu}\right)}_{\text{antisymm.}} &= -\frac{\partial}{\partial x^\nu}\left(\sqrt{g}J^\nu\right) \\ \rightarrow 0 &= \frac{\partial}{\partial x^\nu}(\sqrt{g}J^\nu) \\ \rightarrow 0 &= \frac{1}{\sqrt{g}}\frac{\partial}{\partial x^\nu}(\sqrt{g}J^\nu) \\ \rightarrow &\boxed{J^\nu{}_{;\nu} = 0} \end{aligned} \tag{4.27}$$

which is the charge conservation equation in general relativity.

Exercise: Show that the above equation leads to the conservation of the electric charge.

Let us summarise what we have achieved so far. We have developed a covariant formalism which allows us to write equations that retain their form in all coordinate systems. These equations include the effects of gravitation, provided that in the absence of gravity they reduce to valid equations of special relativity. Our final recipe is simple. We take an equation of special relativity and perform the replacements

$$\eta_{\mu\nu} \to g_{\mu\nu}(x), \tag{4.28}$$

$$\frac{\partial}{\partial x^\mu} \to \nabla_\mu \qquad \text{(covariant derivative)}, \tag{4.29}$$

to obtain the corresponding physical law in the presence of gravity. **Remark.** This minimal-coupling recipe is not unique when the special-relativistic equation contains second derivatives: covariant derivatives do not commute ($[\nabla_\mu, \nabla_\nu]$ gives curvature), and terms involving the curvature tensor that vanish in flat space (e.g. $\xi R\phi^2$ for a scalar field) cannot be excluded by the principle of equivalence. Weinberg (Sec. 4.1) argues they are suppressed only when the scale of the system is small compared with the scale of the gravitational field; Wald Sec. 4.2 discusses the ambiguity. The only missing ingredient needed to use these equations is the metric $g_{\mu\nu}(x)$. In what follows we will pave the way towards the equations of general relativity which determine, in principle, the metric. These are the so-called Einstein field equations.

5

Einstein field equations

5.1 Riemann curvature tensor

We have physically associated the metric with the gravitational potential in the Newtonian limit. Recall that

$$g_{00}(x) \approx 1 + 2\phi(x). \tag{5.1}$$

In Newtonian mechanics the gravitational potential satisfies the Poisson equation

$$\boxed{\nabla^2\phi(x) = 4\pi G\rho(x)} \tag{5.2}$$

We are now seeking a generalisation of the above equation which is also valid at high velocities and in strong fields. So far, we have been successful by imposing a symmetry on our equations, namely covariance under general coordinate transformations. We will use general covariance to generalise (5.2) as well. We will seek an equation in tensor form which collapses to (5.2) for weak gravitational fields.

We will start by generalising the left-hand side of (5.2). We will seek tensors which contain second derivatives of the metric, as suggested by the left-hand side of the equation. We will see that the possibilities are very constrained.

A tensor with second derivatives of the metric emerges in the commutator

$$A_{\mu;\nu;\rho} - A_{\mu;\rho;\nu}, \tag{5.3}$$

where A_μ is a vector. We have

$$\begin{aligned}
A_{\mu;\nu;\rho} = (A_{\mu;\nu})_{;\rho} &= \frac{\partial}{\partial x^\rho}A_{\mu;\nu} - \Gamma^\alpha{}_{\mu\rho}A_{\alpha;\nu} - \Gamma^\alpha{}_{\nu\rho}A_{\mu;\alpha} \\
&= \frac{\partial}{\partial x^\rho}\left(\frac{\partial A_\mu}{\partial x^\nu} - \Gamma^\beta{}_{\mu\nu}A_\beta\right) - \Gamma^\alpha{}_{\mu\rho}\left(\frac{\partial A_\alpha}{\partial x^\nu} - \Gamma^\beta{}_{\alpha\nu}A_\beta\right) - \Gamma^\alpha{}_{\nu\rho}\left(\frac{\partial A_\mu}{\partial x^\alpha} - \Gamma^\beta{}_{\mu\alpha}A_\beta\right) \\
&= \frac{\partial^2 A_\mu}{\partial x^\nu \partial x^\rho} - \Gamma^\alpha{}_{\mu\nu}\frac{\partial A_\alpha}{\partial x^\rho} - \Gamma^\alpha{}_{\mu\rho}\frac{\partial A_\alpha}{\partial x^\nu} - \Gamma^\alpha{}_{\nu\rho}\frac{\partial A_\mu}{\partial x^\alpha} + \Gamma^\alpha{}_{\nu\rho}\Gamma^\beta{}_{\mu\alpha}A_\beta \\
&\quad - \left(\frac{\partial \Gamma^\beta{}_{\mu\nu}}{\partial x^\rho} - \Gamma^\alpha{}_{\mu\rho}\Gamma^\beta{}_{\alpha\nu}\right)A_\beta.
\end{aligned} \tag{5.4}$$

The first line is symmetric under the exchange $\nu \leftrightarrow \rho$. Taking the commutator, we get

$$A_{\mu;\nu;\rho} - A_{\mu;\rho;\nu} = -R^{\beta}{}_{\mu\nu\rho} A_{\beta}, \tag{5.5}$$

with

$$R^{\beta}{}_{\mu\nu\rho} \equiv \frac{\partial \Gamma^{\beta}{}_{\mu\nu}}{\partial x^{\rho}} - \frac{\partial \Gamma^{\beta}{}_{\mu\rho}}{\partial x^{\nu}} - \Gamma^{\alpha}{}_{\mu\rho}\Gamma^{\beta}{}_{\alpha\nu} + \Gamma^{\alpha}{}_{\mu\nu}\Gamma^{\beta}{}_{\alpha\rho}, \tag{5.6}$$

the so-called Riemann curvature tensor. It is indeed a tensor, as the left-hand side of (5.5) is a tensor and A_{β} on the right-hand side is a vector. Similarly, we can find that

$$V^{\mu}{}_{;\nu;\rho} - V^{\mu}{}_{;\rho;\nu} = R^{\mu}{}_{\beta\nu\rho} V^{\beta}, \tag{5.7}$$

and for tensors (as an example)

$$T^{\lambda}{}_{\mu;\nu;\kappa} - T^{\lambda}{}_{\mu;\kappa;\nu} = R^{\lambda}{}_{\sigma\nu\kappa} T^{\sigma}{}_{\mu} - R^{\sigma}{}_{\mu\nu\kappa} T^{\lambda}{}_{\sigma} \quad \text{etc.} \tag{5.8}$$

The curvature tensor is very special. We will prove that it is unique, i.e. it is the only tensor which can be constructed from the metric and its first and second derivatives and which is linear in the second derivatives of the metric (and thus contains $\nabla^2\phi$)! The uniqueness of the Riemann tensor will be essential for generalising the Poisson equation of Newtonian mechanics to the field equations of Einstein.

Proof: Focus on one point X and erect a locally inertial frame. The affine connection vanishes at X, but its derivatives do not,

$$\Gamma^{\alpha}{}_{\mu\nu} = 0\,, \quad \text{but} \quad \frac{\partial \Gamma^{\mu}{}_{\nu\rho}}{\partial x^{\sigma}} \neq 0 \ \text{ in general.} \tag{5.9}$$

Now consider coordinate transformations $x' \to x$ which leave the affine connection zero at the point X; they form a subset of all coordinate transformations. For a generic coordinate transformation, the affine connection transforms as

$$\Gamma^{\mu}{}_{\nu\rho} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}} \frac{\partial x'^{\beta}}{\partial x^{\nu}} \frac{\partial x'^{\gamma}}{\partial x^{\rho}} \Gamma'^{\alpha}{}_{\beta\gamma} + \frac{\partial x^{\mu}}{\partial x'^{\beta}} \frac{\partial^2 x'^{\beta}}{\partial x^{\nu} \partial x^{\rho}}, \tag{5.10}$$

and at the point X, where $\Gamma' = \Gamma = 0$, we have

$$\left(\frac{\partial^2 x'^{\beta}(x)}{\partial x^{\nu} \partial x^{\rho}} \right)_{x=X} = 0. \tag{5.11}$$

The derivative of the affine connection, under these specialised transformations, transforms as

$$\frac{\partial \Gamma^{\mu}{}_{\nu\rho}}{\partial x^{\sigma}} = \underbrace{\frac{\partial x^{\mu}}{\partial x'^{\alpha}} \frac{\partial x'^{\beta}}{\partial x^{\nu}} \frac{\partial x'^{\gamma}}{\partial x^{\rho}} \frac{\partial x'^{\delta}}{\partial x^{\sigma}} \frac{\partial \Gamma'^{\alpha}{}_{\beta\gamma}}{\partial x'^{\delta}}}_{\text{tensor-like}} - \underbrace{\frac{\partial^3 x^{\mu}}{\partial x'^{\alpha} \partial x'^{\beta} \partial x'^{\gamma}} \frac{\partial x'^{\alpha}}{\partial x^{\nu}} \frac{\partial x'^{\beta}}{\partial x^{\rho}} \frac{\partial x'^{\gamma}}{\partial x^{\sigma}}}_{\text{non-tensor-like}} \Bigg|_{x=X}. \tag{5.12}$$

So, the derivative of the affine connection is clearly not a tensor. Can we construct an object which is linear in the second derivative of the metric and transforms as a tensor (at least, for our special coordinate transformations above)? As the affine connection vanishes, the only objects available to us are its derivatives. We can write

$$T^{\mu}{}_{\nu\rho\sigma} = a_1 \frac{\partial \Gamma^{\mu}{}_{\nu\rho}}{\partial x^{\sigma}} + a_2 \frac{\partial \Gamma^{\mu}{}_{\rho\sigma}}{\partial x^{\nu}} + a_3 \frac{\partial \Gamma^{\mu}{}_{\sigma\nu}}{\partial x^{\rho}}, \tag{5.13}$$

which transforms as

$$T^{\mu}{}_{\nu\rho\sigma} = \frac{\partial x^{\mu}}{\partial x'^{\alpha}}\frac{\partial x'^{\beta}}{\partial x^{\nu}}\frac{\partial x'^{\gamma}}{\partial x^{\rho}}\frac{\partial x'^{\delta}}{\partial x^{\sigma}}T'^{\alpha}{}_{\beta\gamma\delta} - (a_1+a_2+a_3)\frac{\partial^3 x^{\mu}}{\partial x'^{\alpha}\partial x'^{\beta}\partial x'^{\gamma}}\frac{\partial x'^{\alpha}}{\partial x^{\nu}}\frac{\partial x'^{\beta}}{\partial x^{\rho}}\frac{\partial x'^{\gamma}}{\partial x^{\sigma}}\bigg|_{x=X}. \quad (5.14)$$

For the non-tensorial part to vanish we must have $a_3 = -a_1 - a_2$. So, the general form is

$$\begin{aligned} T^{\mu}{}_{\nu\rho\sigma}\Big|_{x=X} &= a_1\left(\frac{\partial\Gamma^{\mu}{}_{\nu\rho}}{\partial x^{\sigma}} - \frac{\partial\Gamma^{\mu}{}_{\nu\sigma}}{\partial x^{\rho}}\right)_{x=X} + a_2\left(\frac{\partial\Gamma^{\mu}{}_{\sigma\rho}}{\partial x^{\nu}} - \frac{\partial\Gamma^{\mu}{}_{\sigma\nu}}{\partial x^{\rho}}\right)_{x=X} \\ &= a_1 R^{\mu}{}_{\nu\rho\sigma}\Big|_{x=X} + a_2 R^{\mu}{}_{\sigma\rho\nu}\Big|_{x=X}. \end{aligned} \quad (5.15)$$

So, at the point X in our locally inertial reference frame, the only object which we can construct to behave as a tensor under our special coordinate transformations is the Riemann tensor (or a superposition of Riemann tensors). Suppose there exists a second tensor $B^{\mu}{}_{\nu\rho\sigma}$ which is also linear in the second derivatives of the metric. At $x = X$, in the coordinates of our locally inertial frame, we must have

$$B^{\mu}{}_{\nu\rho\sigma} = a_1 R^{\mu}{}_{\nu\rho\sigma} + a_2 R^{\mu}{}_{\sigma\rho\nu}. \quad (5.16)$$

But this is an identity between tensors, so it is valid in all coordinate systems. This concludes the proof of the uniqueness of the Riemann tensor, as the only tensor linearly dependent on the second derivative of the metric. **Remark.** For other index structures, B may also contain Riemann tensors with permuted indices contracted with $g_{\mu\nu}$ or $g^{\mu\nu}$, plus terms built from $g_{\mu\nu}$ alone (Weinberg, Sec. 6.2).

From the Riemann tensor we can form contractions. The following ones are very important in general relativity:

- Ricci tensor: $R_{\mu\nu} \equiv R^{\alpha}{}_{\mu\alpha\nu}$,
- Curvature scalar: $R \equiv g^{\mu\nu}R_{\mu\nu}$.

These will enter Einstein's field equations, as we will see later on.

5.1.1 Parallel transport and the Riemann tensor

Consider a particle with spin $S_{\mu}(\tau)$ following the trajectory $x^{\mu}(\tau)$. We have already seen that the equation of motion for the spin is

$$\frac{DS_{\mu}}{D\tau} = 0 \;\Rightarrow\; \boxed{\frac{dS_{\mu}}{d\tau} = \Gamma^{\alpha}{}_{\mu\rho}\frac{dx^{\rho}}{d\tau}S_{\alpha}} \quad (5.17)$$

We say that the spin vector is parallel transported along the trajectory if its equation of motion is given by the above. The reason is that, viewed from a locally inertial frame, the vector remains the same ($\frac{dS_{\mu}}{d\tau} = 0$) as it is transported along the trajectory.

Let us now think of a closed trajectory (e.g. a gyroscope in orbit around the Earth; see Fig. 5.1).

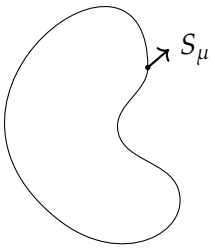


Figure 5.1. A vector transported around a closed loop.

The question we would like to investigate is by how much the spin vector (or any other parallel-transported vector) changes after a full orbit, when it returns to the same point. To start, we will consider small closed curves. The affine connection can be expanded as

$$\Gamma^{\mu}{}_{\nu\rho}(x) = \Gamma^{\mu}{}_{\nu\rho}(X) + (x^\alpha - X^\alpha)\frac{\partial \Gamma^{\mu}{}_{\nu\rho}}{\partial x^\alpha}(X) + \cdots , \tag{5.18}$$

where $X^\alpha \equiv x^\alpha(\tau_0)$ is a point on the curve. The differential equation for parallel transport gives

$$\begin{aligned} \frac{S_\mu(\tau) - S_\mu(\tau_0)}{\tau - \tau_0} &\simeq \Gamma^{\alpha}{}_{\mu\rho}\frac{x^\rho(\tau) - x^\rho(\tau_0)}{\tau - \tau_0} S_\alpha(\tau_0) \\ S_\mu(\tau) &= S_\mu(\tau_0) + \Gamma^{\alpha}{}_{\mu\rho}(X)(x^\rho(\tau) - X^\rho)S_\alpha(\tau_0) + \cdots \end{aligned} \tag{5.19}$$

This is only a first approximation, which we can improve by one order. Let us return to the original differential equation,

$$\begin{aligned} \frac{\mathrm{d}S_\mu}{\mathrm{d}\tau} &= \Gamma^{\alpha}{}_{\mu\rho} S_\alpha \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} \\ \int_{\tau_0}^{\tau} \mathrm{d}S_\mu &= \int_{\tau_0}^{\tau} \mathrm{d}\tau\, \Gamma^{\alpha}{}_{\mu\rho} S_\alpha \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} \\ \Rightarrow\ S_\mu(\tau) &= S_\mu(\tau_0) + \int_{\tau_0}^{\tau} \mathrm{d}\tau\, \Gamma^{\alpha}{}_{\mu\rho} S_\alpha \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} \\ &\approx S_\mu(\tau_0) + \int_{\tau_0}^{\tau} \mathrm{d}\tau \left[\Gamma^{\alpha}{}_{\mu\rho}(X) + \frac{\partial \Gamma^{\alpha}{}_{\mu\rho}}{\partial x^\beta}(X)(x^\beta - X^\beta) + \cdots\right] \\ &\qquad \cdot \left[S_\alpha(\tau_0) + \Gamma^{\gamma}{}_{\alpha\delta}(X)(x^\delta - X^\delta)S_\gamma(x) + \cdots\right] \cdot \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} \\ &\approx S_\mu(\tau_0) + \Gamma^{\alpha}{}_{\mu\rho}(X)S_\alpha(\tau_0)\int_{\tau_0}^{\tau} \mathrm{d}\tau\, \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} \\ &\qquad + \left[\frac{\partial \Gamma^{\alpha}{}_{\mu\rho}}{\partial x^\beta}(X)S_\alpha(\tau_0) + \Gamma^{\omega}{}_{\mu\rho}(X)\Gamma^{\alpha}{}_{\omega\beta}(X)S_\alpha(\tau_0)\right]\int_{\tau_0}^{\tau} \mathrm{d}\tau\, (x^\beta - X^\beta)\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau}. \end{aligned} \tag{5.20}$$

For a closed loop,

$$\oint \mathrm{d}\tau\, \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} = 0. \tag{5.21}$$

Thus, the change in the vector becomes

$$\Delta S_\mu = \left[\frac{\partial \Gamma^{\alpha}{}_{\mu\rho}}{\partial x^\beta}(X) + \Gamma^{\omega}{}_{\mu\rho}\Gamma^{\alpha}{}_{\omega\beta}\right] S_\alpha(X) \oint \mathrm{d}\tau\, \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} x^\beta. \tag{5.22}$$

Notice that the term in brackets resembles the Riemann tensor and that the integral is antisymmetric in $\rho \leftrightarrow \beta$. Indeed,

$$\begin{aligned} \oint \mathrm{d}\tau \, \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} x^\beta &= \oint \mathrm{d}\tau \left[\frac{\mathrm{d}}{\mathrm{d}\tau}(x^\rho x^\beta) - x^\rho \frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} \right] \\ &= - \oint \mathrm{d}\tau \, x^\rho \frac{\mathrm{d}x^\beta}{\mathrm{d}\tau}. \end{aligned} \tag{5.23}$$

We can then write, using antisymmetry,

$$\Delta S_\mu = \frac{1}{2} R^\alpha{}_{\mu\rho\beta}(X) S_\alpha(X) \oint \mathrm{d}\tau \, \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} x^\beta. \tag{5.24}$$

We conclude that if $R^\alpha{}_{\mu\nu\rho} = 0$, then $\Delta S_\mu = 0$: there is no change in the vector.

We have worked with small closed curves. If the Riemann tensor vanishes in a large region, we can split a large loop in this region into small cells. As in the proof of Stokes' theorem, the contributions from the common boundaries of neighbouring cells cancel, leaving only the contribution of the external boundary. Thus, the vector is unchanged after a full orbit around any closed loop in this region.

Consider now the transition from a point $x(\tau_0)$ to $x(\tau_1)$ via two different paths, A and B (Fig. 5.2).

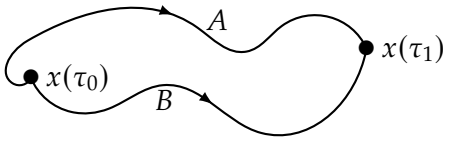


Figure 5.2. Two paths connecting the points $x(\tau_0)$ and $x(\tau_1)$.

If the Riemann tensor vanishes in the region between the two paths, the change of the vector is the same along both, $\Delta S_A = \Delta S_B$, since path A followed by path B reversed is a closed loop. **Remark.** In curved spacetime, parallel transport is path dependent. The region between the paths must also be simply connected (cf. Weinberg Sec. 6.3, Wald Sec. 3.2).

In Minkowski spacetime, where $g_{\mu\nu}(x) = \eta_{\mu\nu}$ and $\Gamma^\mu{}_{\nu\rho} = 0$ everywhere, the Riemann curvature tensor vanishes. The converse is also true. One can show that

(i) if $R^\alpha{}_{\mu\nu\rho} = 0$ and

(ii) the metric $g_{\mu\nu}(x)$ at a point X has three negative eigenvalues and one positive eigenvalue (as $\eta_{\mu\nu}$),

then there exists a coordinate system ξ^α in which the metric is the Minkowski metric everywhere, $g_{\mu\nu}(\xi) = \eta_{\mu\nu}$ $\forall \xi$.

Proof: The metric matrix is symmetric. We can always bring it to diagonal form with an appropriate orthogonal matrix O, such that

$$O \mathbf{g} O^T = D, \tag{5.25}$$

where **g** is the metric matrix and D is diagonal. In components,

$$O^{\alpha}{}_{\mu} g^{\mu\nu} O^{\beta}{}_{\nu} = D^{\alpha\beta} = \begin{cases} D^{\alpha}, & \alpha = \beta, \\ 0, & \alpha \neq \beta. \end{cases} \tag{5.26}$$

We have one positive and three negative eigenvalues. Let us then choose D^0 positive and D^1, D^2, D^3 negative. We can then define normalised matrices d through

$$\begin{aligned} O^{0}{}_{\mu} &= \sqrt{+D^0}\, d^{0}{}_{\mu}, \\ O^{i}{}_{\mu} &= \sqrt{-D^i}\, d^{i}{}_{\mu}. \end{aligned} \tag{5.27}$$

Then the diagonalisation equation takes the form

$$\boxed{g^{\mu\nu}(X) d^{\alpha}{}_{\mu} d^{\beta}{}_{\nu} = \eta^{\alpha\beta}} \tag{5.28}$$

Let us now define a quantity $D^{\alpha}{}_{\mu}(x)$ at an arbitrary point x through the equation

$$\frac{\partial D^{\alpha}{}_{\mu}(x)}{\partial x^{\nu}} = \Gamma^{\lambda}{}_{\mu\nu}(x) D^{\alpha}{}_{\lambda}(x), \tag{5.29}$$

with the boundary condition $D^{\alpha}{}_{\mu}(X) = d^{\alpha}{}_{\mu}$. Let us now consider the derivative

$$(g^{\mu\nu} D^{\alpha}{}_{\mu} D^{\beta}{}_{\nu})_{;\lambda} = 2 g^{\mu\nu} \underbrace{D^{\alpha}{}_{\mu;\lambda}}_{=0} D^{\beta}{}_{\nu} = 0. \tag{5.30}$$

But $g^{\mu\nu} D^{\alpha}{}_{\mu} D^{\beta}{}_{\nu}$ is a scalar. So, its covariant derivative equals its ordinary derivative. Thus, we have shown that

$$\begin{aligned} & \frac{\partial}{\partial x^{\lambda}} \left(g^{\mu\nu} D^{\alpha}{}_{\mu} D^{\beta}{}_{\nu} \right) = 0 \\ \Rightarrow \; & g^{\mu\nu}(x) D^{\alpha}{}_{\mu}(x) D^{\beta}{}_{\nu}(x) = g^{\mu\nu}(X) D^{\alpha}{}_{\mu}(X) D^{\beta}{}_{\nu}(X) \\ & \qquad = g^{\mu\nu}(X) d^{\alpha}{}_{\mu} d^{\beta}{}_{\nu} \\ & \qquad = \eta^{\alpha\beta}. \end{aligned} \tag{5.31}$$

Thus, there is a coordinate transformation encoded in $D^{\alpha}{}_{\mu}(x)$ which transforms the metric into the Minkowski metric. The appropriate coordinate transformation ($x \to \xi$) is

$$\frac{\partial \xi^{\alpha}}{\partial x^{\mu}} = D^{\alpha}{}_{\mu}. \tag{5.32}$$

Then

$$g^{\mu\nu}(x) \frac{\partial \xi^{\alpha}}{\partial x^{\mu}} \frac{\partial \xi^{\beta}}{\partial x^{\nu}} = \eta^{\alpha\beta}, \tag{5.33}$$

which is the transformation law of the (inverse) metric tensor. In the above derivation we have made a crucial assumption, namely that a solution of

$$\frac{\partial D^{\alpha}{}_{\mu}}{\partial x^{\nu}} = \Gamma^{\lambda}{}_{\mu\nu} D^{\alpha}{}_{\lambda}, \tag{5.34}$$

does exist at every point in spacetime. This is guaranteed if $R^{\alpha}{}_{\mu\nu\rho} = 0$. To see this, consider an equation of the same form,

$$\frac{\partial S_\mu(x)}{\partial x^\nu} = \Gamma^{\lambda}{}_{\mu\nu} S_\lambda(x), \tag{5.35}$$

for a general vector field $S_\mu(x)$ defined everywhere in spacetime. Along an arbitrary curve $x^\mu(\tau)$, we multiply with $\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}$ and obtain

$$\begin{aligned} \frac{\partial S_\mu(x)}{\partial x^\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} &= \Gamma^{\lambda}{}_{\mu\nu}(x)\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} S_\lambda(x) \\ \Leftrightarrow \quad \frac{\mathrm{d}S_\mu}{\mathrm{d}\tau} &= \Gamma^{\lambda}{}_{\mu\nu}(x)\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} S_\lambda(x(\tau)). \end{aligned} \tag{5.36}$$

This ordinary differential equation along the curve can always be solved uniquely. The field $S_\mu(x)$ exists if the result is independent of the curve chosen to connect X and x, i.e. if $\Delta S_\mu = 0$ around every closed loop. By the previous subsection, this holds when $R^{\alpha}{}_{\mu\nu\rho} = 0$ (in a simply connected region). **Remark.** Equivalently, the integrability condition $\partial_\rho(\Gamma^{\lambda}{}_{\mu\nu} S_\lambda) = \partial_\nu(\Gamma^{\lambda}{}_{\mu\rho} S_\lambda)$ reduces to $R^{\lambda}{}_{\mu\nu\rho} S_\lambda = 0$. In addition, $\partial\xi^\alpha/\partial x^\mu = D^{\alpha}{}_{\mu}$ is integrable because $\partial_\nu D^{\alpha}{}_{\mu} = \Gamma^{\lambda}{}_{\mu\nu} D^{\alpha}{}_{\lambda}$ is symmetric in $\mu\nu$ (Weinberg Sec. 6.3).

5.1.2 Algebraic properties of the Riemann tensor

The Riemann tensor is given by

$$R^{\mu}{}_{\nu\rho\sigma} = \frac{\partial \Gamma^{\mu}{}_{\nu\rho}}{\partial x^\sigma} + \Gamma^{\alpha}{}_{\nu\rho}\Gamma^{\mu}{}_{\alpha\sigma} - \frac{\partial \Gamma^{\mu}{}_{\nu\sigma}}{\partial x^\rho} - \Gamma^{\alpha}{}_{\nu\sigma}\Gamma^{\mu}{}_{\alpha\rho}, \tag{5.37}$$

where the affine connection is given by

$$\Gamma^{\mu}{}_{\nu\rho} = g^{\mu\alpha}\Gamma_{\alpha\nu\rho} \text{ and } \Gamma_{\alpha\nu\rho} = \frac{1}{2}\left[\frac{\partial g_{\alpha\nu}}{\partial x^\rho} + \frac{\partial g_{\alpha\rho}}{\partial x^\nu} - \frac{\partial g_{\nu\rho}}{\partial x^\alpha}\right]. \tag{5.38}$$

Let us lower the first index,

$$R_{\mu\nu\rho\sigma} = g_{\mu\alpha} R^{\alpha}{}_{\nu\rho\sigma}. \tag{5.39}$$

After some straightforward algebra, using that

$$0 = \frac{\partial \delta^{\mu}{}_{\beta}}{\partial x^\sigma} = \frac{\partial}{\partial x^\sigma}(g^{\mu\alpha} g_{\alpha\beta}) \rightarrow g_{\alpha\beta}\frac{\partial}{\partial x^\sigma} g^{\mu\alpha} = -g^{\mu\alpha}\frac{\partial}{\partial x^\sigma} g_{\alpha\beta}, \tag{5.40}$$

we find

$$R_{\mu\nu\rho\sigma} = \frac{\partial \Gamma_{\mu\nu\rho}}{\partial x^\sigma} - \frac{\partial \Gamma_{\mu\nu\sigma}}{\partial x^\rho} - \Gamma^{\beta}{}_{\nu\rho}\Gamma_{\beta\mu\sigma} + \Gamma^{\beta}{}_{\nu\sigma}\Gamma_{\beta\mu\rho}, \tag{5.41}$$

or, more explicitly,

$$R_{\mu\nu\rho\sigma} = \frac{1}{2}\left[\frac{\partial^2 g_{\mu\rho}}{\partial x^\nu \partial x^\sigma} + \frac{\partial^2 g_{\nu\sigma}}{\partial x^\mu \partial x^\rho} - \frac{\partial^2 g_{\nu\rho}}{\partial x^\mu \partial x^\sigma} - \frac{\partial^2 g_{\mu\sigma}}{\partial x^\nu \partial x^\rho}\right] - g_{\alpha\beta}\Gamma^{\alpha}{}_{\mu\sigma}\Gamma^{\beta}{}_{\nu\rho} + g_{\alpha\beta}\Gamma^{\alpha}{}_{\mu\rho}\Gamma^{\beta}{}_{\nu\sigma}. \tag{5.42}$$

We then find the following properties:

Symmetry

$$R_{\mu\nu\rho\sigma} = R_{\rho\sigma\mu\nu}. \tag{5.43}$$

Antisymmetry

$$R_{\mu\nu\rho\sigma} = -R_{\mu\nu\sigma\rho}, \tag{5.44}$$
$$R_{\mu\nu\rho\sigma} = -R_{\nu\mu\rho\sigma}. \tag{5.45}$$

From the above it also follows that

$$R_{\mu\nu\rho\sigma} = R_{\nu\mu\sigma\rho}. \tag{5.46}$$

Cyclicity

$$R^{\mu}{}_{\nu\rho\sigma} + R^{\mu}{}_{\sigma\nu\rho} + R^{\mu}{}_{\rho\sigma\nu} = 0. \tag{5.47}$$

5.1.3 The Ricci tensor and Ricci scalar

We have already proven that the Riemann tensor and its linear combinations are the only possible tensors where second derivatives of the metric enter linearly. One such combination is the Ricci tensor,

$$R_{\mu\nu} \equiv R^{\alpha}{}_{\mu\alpha\nu} = g^{\alpha\beta} R_{\alpha\mu\beta\nu}. \tag{5.48}$$

The Ricci tensor is the only second-rank tensor that can be formed in this way, because of the antisymmetry of the Riemann tensor. Indeed,

$$\underbrace{g^{\alpha\beta}}_{\text{symm.}} \underbrace{R_{\alpha\beta\gamma\delta}}_{\text{antisymm.}} = 0, \quad \text{etc.} \tag{5.49}$$

We can also form a scalar, the Ricci or curvature scalar,

$$R \equiv g^{\mu\nu} R_{\mu\nu} = R_{\alpha\beta\gamma\delta} g^{\alpha\gamma} g^{\beta\delta}. \tag{5.50}$$

No other scalar possibilities exist due to the antisymmetry and cyclicity of the Riemann tensor. Notice also that

$$\begin{aligned} 0 &= \varepsilon^{\mu\nu\rho\sigma} \overbrace{(R_{\mu\nu\rho\sigma} + R_{\mu\rho\sigma\nu} + R_{\mu\sigma\nu\rho})}^{=0} \\ &= R_{\mu\nu\rho\sigma}(\varepsilon^{\mu\nu\rho\sigma} + \varepsilon^{\mu\rho\sigma\nu} + \varepsilon^{\mu\sigma\nu\rho}) \\ &= R_{\mu\nu\rho\sigma}(\varepsilon^{\mu\nu\rho\sigma} + (-1)(-1)\varepsilon^{\mu\nu\rho\sigma} + (-1)(-1)\varepsilon^{\mu\nu\rho\sigma}) \\ \rightarrow 0 &= \varepsilon^{\mu\nu\rho\sigma} R_{\mu\nu\rho\sigma}. \end{aligned} \tag{5.51}$$

So, while we could potentially contract the Riemann tensor with a Levi-Civita symbol to form a scalar, this contraction yields zero.

5.1.4 The Bianchi identities

One can prove the following Bianchi identity,

$$R_{\mu\nu\rho\sigma;\tau} + R_{\mu\nu\tau\rho;\sigma} + R_{\mu\nu\sigma\tau;\rho} = 0. \tag{5.52}$$

We now multiply with $g^{\mu\rho}$. Since $g^{\mu\rho}{}_{;\beta} = 0$, we arrive at

$$\begin{aligned} R^{\rho}{}_{\nu\rho\sigma;\tau} + R^{\rho}{}_{\nu\tau\rho;\sigma} + R^{\rho}{}_{\nu\sigma\tau;\rho} &= 0 \\ R^{\rho}{}_{\nu\rho\sigma;\tau} - R^{\rho}{}_{\nu\rho\tau;\sigma} + R^{\rho}{}_{\nu\sigma\tau;\rho} &= 0 \\ \Rightarrow \boxed{R_{\nu\sigma;\tau} - R_{\nu\tau;\sigma} + R^{\rho}{}_{\nu\sigma\tau;\rho} = 0} & \end{aligned} \tag{5.53}$$

Multiplying with $g^{\nu\sigma}$, we now get

$$\begin{aligned} R_{;\tau} - R^{\sigma}{}_{\tau;\sigma} + g^{\rho\beta} g^{\nu\sigma} R_{\beta\nu\sigma\tau;\rho} &= 0 \\ R_{;\tau} - R^{\sigma}{}_{\tau;\sigma} - g^{\rho\beta} g^{\nu\sigma} R_{\nu\beta\sigma\tau;\rho} &= 0 \\ R_{;\tau} - R^{\sigma}{}_{\tau;\sigma} - g^{\rho\beta} R_{\beta\tau;\rho} &= 0 \\ 2R^{\sigma}{}_{\tau;\sigma} - R_{;\tau} &= 0 \\ R^{\sigma}{}_{\tau;\sigma} - \frac{1}{2}\left(\delta^{\sigma}{}_{\tau} R\right)_{;\sigma} &= 0 \\ \Rightarrow \boxed{\left(R^{\sigma}{}_{\tau} - \frac{1}{2}\delta^{\sigma}{}_{\tau} R\right)_{;\sigma} = 0} & \end{aligned} \tag{5.54}$$

or, equivalently (Exercise),

$$\boxed{\left(R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R\right)_{;\mu} = 0} \tag{5.55}$$

5.2 The energy-momentum tensor

We are looking for a relativistic generalisation of Poisson's equation

$$\nabla^2 \phi(x) = 4\pi G \rho(x), \tag{5.56}$$

of Newtonian mechanics. We have made progress by realising that the only tensor which contains the second derivatives of the metric (the analogue of the potential in general relativity) is the Riemann tensor. We now turn our attention to the right-hand side of the above equation, where we encounter the mass density,

$$\rho(\vec{x}, t) = \sum_n m_n \, \delta^{(3)}(\vec{x} - \vec{x}_n(t)), \tag{5.57}$$

where the sum ranges over all particles and the δ-function localises them on their trajectories. The mass density is one of the components of the energy-momentum tensor, which in special relativity reads

$$T^{\mu\nu}(\vec{x}, t) = \sum_n P_n^{\mu} \frac{\mathrm{d}x_n^{\nu}}{\mathrm{d}t} \, \delta^{(3)}(\vec{x} - \vec{x}_n(t)) = \sum_n m_n \int \mathrm{d}\tau \, \frac{\mathrm{d}x_n^{\mu}}{\mathrm{d}\tau} \frac{\mathrm{d}x_n^{\nu}}{\mathrm{d}\tau} \, \delta(x - x_n(\tau)). \tag{5.58}$$

In special relativity, the right-hand side of the above equation is manifestly a tensor since $\frac{\mathrm{d}x_n^\mu}{\mathrm{d}\tau}$ is a vector and the δ-function is a scalar. In general relativity the δ-function is not a scalar but a scalar density. Consider a scalar field $\phi(x)$. This can be written as

$$\begin{aligned} \phi(x) &= \int \mathrm{d}^4 y\, \delta(x-y)\phi(y) \\ \Rightarrow \underbrace{\phi(x)}_{\text{scalar}} &= \int \underbrace{\mathrm{d}^4 y \sqrt{g}}_{\text{scalar}} \underbrace{\frac{\delta(x-y)}{\sqrt{g}}}_{\text{scalar}} \underbrace{\phi(y)}_{\text{scalar}} . \end{aligned} \tag{5.59}$$

For the above to be an invariant equation under general coordinate transformations, we must have that $\frac{\delta(x-y)}{\sqrt{g}}$ rather than $\delta(x-y)$ is a scalar. We can now write down an energy-momentum tensor for matter in general relativity. It reads

$$\boxed{T^{\mu\nu}(x) = \frac{1}{\sqrt{g(x)}} \sum_n m_n \int \mathrm{d}\tau \, \frac{\mathrm{d}x_n^\mu}{\mathrm{d}\tau} \frac{\mathrm{d}x_n^\nu}{\mathrm{d}\tau} \, \delta(x - x_n(\tau))} \tag{5.60}$$

The energy-momentum tensor of special relativity satisfies the continuity equation

$$\frac{\partial T^{\mu\nu}}{\partial x^\mu} = G^\nu , \tag{5.61}$$

where G^ν on the right-hand side is the force density. Its generalisation in general relativity is

$$T^{\mu\nu}{}_{;\mu} = G^\nu , \tag{5.62}$$

where G^ν is an appropriate generalisation of the force density (excluding the gravitational force). The effect of gravitation in (5.62) resides in the covariant differentiation. Explicitly,

$$\begin{aligned} T^{\mu\nu}{}_{;\mu} &= \frac{\partial T^{\mu\nu}}{\partial x^\mu} + \Gamma^\mu{}_{\mu\rho} T^{\rho\nu} + \Gamma^\nu{}_{\mu\rho} T^{\mu\rho} \\ &= \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^\mu} \left(\sqrt{g} T^{\mu\nu} \right) + \Gamma^\nu{}_{\mu\rho} T^{\mu\rho} . \end{aligned} \tag{5.63}$$

Thus,

$$\begin{aligned} T^{\mu\nu}{}_{;\mu} &= G^\nu \\ \Rightarrow \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^\mu} (\sqrt{g} T^{\mu\nu}) &= G^\nu - \underbrace{\Gamma^\nu{}_{\mu\rho} T^{\mu\rho}}_{\text{gravitational "force density"}} . \end{aligned} \tag{5.64}$$

5.3 Einstein's field equations

We now have all the ingredients to build the relativistic generalisation of Newton's law $\nabla^2\phi = 4\pi G\rho$. As $\nabla^2\phi$ is contained inside the Riemann tensor and $\rho = T^{00}$ inside the energy-momentum tensor, we can make an ansatz of the form

$$G^{\mu\nu} = c_0 T^{\mu\nu} , \tag{5.65}$$

where $G^{\mu\nu}$ is formed, in general, from the Riemann tensor and the metric. As $T^{\mu\nu} = T^{\nu\mu}$, we expect $G^{\mu\nu}$ to be symmetric. Also, taking the covariant divergence of both sides, we have that

$$G^{\mu\nu}{}_{;\mu} = c_0 T^{\mu\nu}{}_{;\mu}. \tag{5.66}$$

In the absence of non-gravitational forces, we have $T^{\mu\nu}{}_{;\mu} = 0$ and thus we must also have

$$G^{\mu\nu}{}_{;\mu} = 0. \tag{5.67}$$

We have seen that the combination

$$R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R \tag{5.68}$$

has a vanishing covariant divergence. So, the most general symmetric second-rank tensor of this type with vanishing covariant divergence is

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R + \Lambda g^{\mu\nu}, \tag{5.69}$$

with $G^{\mu\nu}{}_{;\mu} = 0$. **Remark.** Two steps are implicit here. (i) $T^{\mu\nu}{}_{;\mu} = 0$ must hold for the total energy-momentum tensor (matter plus electromagnetic and all other non-gravitational fields), so the condition $G^{\mu\nu}{}_{;\mu} = 0$ does not rely on the absence of forces. (ii) The general symmetric tensor linear in the Riemann tensor is $aR^{\mu\nu} + bg^{\mu\nu}R + \Lambda g^{\mu\nu}$; its divergence is $(a/2 + b)R^{;\nu}$, which vanishes for all metrics only if $b = -a/2$, and a is absorbed into c_0 (Weinberg Sec. 7.1). Lovelock's theorem makes this the unique choice in four dimensions. We have arrived at the equation

$$\boxed{R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R + \Lambda g^{\mu\nu} = c_0 T^{\mu\nu}} \tag{5.70}$$

or, with lower indices,

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = c_0 T_{\mu\nu}. \tag{5.71}$$

Let us now compute the limit of weak, static gravitational fields and slow, non-relativistic particles. **Remark.** In the following derivation we drop the term $\Lambda g_{\mu\nu}$, i.e. we assume Λ to be negligible. Keeping it, the 00 equation becomes $\nabla^2\phi = 4\pi G\rho - \Lambda$, i.e. a repulsive potential $-\Lambda r^2/6$ for $\Lambda > 0$, negligible on solar-system scales (Carroll Sec. 4.2–4.3). We can decompose $g_{\mu\nu} \approx \eta_{\mu\nu} + h_{\mu\nu}$ with a small $h_{\mu\nu}$ and $\frac{\partial h_{\mu\nu}}{\partial x^0} \approx 0$ (static field approximation). Then the affine connection is

$$\Gamma^{\mu}{}_{\nu\rho} \approx \frac{1}{2}\eta^{\mu\alpha}\left[\frac{\partial h_{\alpha\nu}}{\partial x^\rho} + \frac{\partial h_{\alpha\rho}}{\partial x^\nu} - \frac{\partial h_{\nu\rho}}{\partial x^\alpha}\right] \sim O(h). \tag{5.72}$$

Terms of the form $\Gamma^{\alpha}{}_{\mu\nu}\Gamma^{\beta}{}_{\rho\sigma} \sim O(h^2)$ can be dropped in the linear approximation. The Ricci tensor is then

$$\begin{aligned} R_{\mu\nu} &\approx \frac{\partial \Gamma^{\alpha}{}_{\mu\alpha}}{\partial x^\nu} - \frac{\partial \Gamma^{\alpha}{}_{\mu\nu}}{\partial x^\alpha} \\ &\approx \frac{\eta^{\alpha\beta}}{2}\left[\frac{\partial}{\partial x^\nu}\left(\frac{\partial g_{\mu\beta}}{\partial x^\alpha} + \frac{\partial g_{\alpha\beta}}{\partial x^\mu} - \frac{\partial g_{\mu\alpha}}{\partial x^\beta}\right) - \frac{\partial}{\partial x^\alpha}\left(\frac{\partial g_{\mu\beta}}{\partial x^\nu} + \frac{\partial g_{\nu\beta}}{\partial x^\mu} - \frac{\partial g_{\mu\nu}}{\partial x^\beta}\right)\right] \\ \rightarrow\ R_{\mu\nu} &\approx \frac{\eta^{\alpha\beta}}{2}\left[\frac{\partial^2 h_{\alpha\beta}}{\partial x^\mu \partial x^\nu} - \frac{\partial^2 h_{\mu\alpha}}{\partial x^\beta \partial x^\nu} - \frac{\partial^2 h_{\nu\beta}}{\partial x^\alpha \partial x^\mu} + \frac{\partial^2 h_{\mu\nu}}{\partial x^\alpha \partial x^\beta}\right]. \end{aligned} \tag{5.73}$$

The R_{00} component will be of special interest. Ignoring $\frac{\partial h_{\mu\nu}}{\partial x^0}$ terms, within the static approximation, we have

$$R_{00} \approx \frac{\eta^{jk}}{2} \frac{\partial^2 h_{00}}{\partial x^j \partial x^k} = -\frac{1}{2}\nabla^2 h_{00}. \tag{5.74}$$

Then,

$$G_{00} \equiv R_{00} - \frac{1}{2} g_{00} R \approx R_{00} - \frac{1}{2}\eta_{00} R = R_{00} - \frac{R}{2}, \tag{5.75}$$

and thus

$$\boxed{R_{00} - \frac{R}{2} \approx c_0 T_{00} = c_0 \rho} \tag{5.76}$$

Let us now turn to the spatial components. Since $\frac{\mathrm{d}x^i}{\mathrm{d}t} \to 0$, we have $|T_{ij}| \ll T_{00}$. This then implies that

$$\begin{aligned} R_{ij} - \frac{\eta_{ij}}{2} R &\approx 0 \\ R_{ij} &\approx \frac{\eta_{ij}}{2} R \\ R_{kk} &\approx -\frac{3}{2} R. \end{aligned} \tag{5.77}$$

The curvature scalar can now be approximated as

$$\begin{aligned} R = g^{\mu\nu} R_{\mu\nu} = R_{00} - R_{kk} = R_{00} + \frac{3}{2} R \\ \Rightarrow \; R \approx -2 R_{00}. \end{aligned} \tag{5.78}$$

Thus, our ansatz for Einstein's equations becomes

$$\begin{aligned} R_{00} - \frac{1}{2} R &= c_0 \rho \\ R_{00} &= \frac{c_0}{2} \rho \\ \boxed{-\nabla^2 h_{00} = c_0 \rho} \end{aligned} \tag{5.79}$$

It now remains to identify the h_{00} metric perturbation with the Newtonian potential. We have already done this by comparing with the equation of motion in the Newtonian limit, where we found $h_{00} \approx 2\phi$. Substituting into the above equation, we obtain

$$\nabla^2 \phi = -\frac{c_0}{2} \rho. \tag{5.80}$$

Comparing with $\nabla^2 \phi = 4\pi G \rho$, we can identify

$$c_0 = -8\pi G. \tag{5.81}$$

Einstein's equations then read

$$\boxed{R_{\mu\nu} - \frac{g_{\mu\nu}}{2} R + \Lambda g_{\mu\nu} = -8\pi G T_{\mu\nu}} \tag{5.82}$$

Remark. The minus sign is a consequence of the Riemann convention used here (Weinberg's, eq. 6.1.5, with $R_{\mu\nu} = R^{\alpha}{}_{\mu\alpha\nu}$), which is minus that of MTW, Carroll and Wald; with theirs

(and their signature $(-,+,+,+)$) the same equation reads $R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = +8\pi G T_{\mu\nu}$, with the same Λ. Because $R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$ and $T_{\mu\nu}$ are unchanged under $g_{\mu\nu} \to -g_{\mu\nu}$, the sign of the $8\pi G$ term does not depend on the signature $(+,-,-,-)$; the term $\Lambda g_{\mu\nu}$ does change sign, so that in Weinberg's signature (same Riemann convention as here) the equation reads $R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R - \Lambda g_{\mu\nu} = -8\pi G T_{\mu\nu}$. With this sign $\Lambda > 0$ acts as a vacuum energy density $\rho_\Lambda = \Lambda/(8\pi G) > 0$; e.g. the vacuum solution $R_{\mu\nu} = \Lambda g_{\mu\nu}$ of the static form $B = 1/A = 1 - \Lambda r^2/3$ is de Sitter space, with a repulsive potential $\phi = -\Lambda r^2/6$.

The constant Λ is the so-called cosmological constant. Since it multiplies a term without derivatives, it must have units of inverse length squared,

$$\Lambda \sim \frac{1}{\text{Length}^2}. \tag{5.83}$$

Its introduction implies the existence of a physical length scale, which must be very large, as we do not observe significant deviations from Newtonian gravity at small length scales. However, at large scales the cosmological constant plays an important role, as it affects, for example, the expansion rate of the universe.

We can now find an alternative form of Einstein's equations. By contracting the μ and ν indices, we find

$$\begin{aligned} g^{\mu\nu}\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu}\right) &= -8\pi G T^{\mu}{}_{\mu} \\ R - \frac{4}{2}R + 4\Lambda &= -8\pi G T^{\mu}{}_{\mu} \\ 8\pi G T^{\mu}{}_{\mu} + 4\Lambda &= R. \end{aligned} \tag{5.84}$$

Then, substituting the above into Einstein's equations, we have

$$\begin{aligned} R_{\mu\nu} - \frac{g_{\mu\nu}}{2}\left(8\pi G T^{\alpha}{}_{\alpha} + 4\Lambda\right) + \Lambda g_{\mu\nu} &= -8\pi G T_{\mu\nu} \\ R_{\mu\nu} &= -8\pi G\left(T_{\mu\nu} - T^{\alpha}{}_{\alpha}\frac{g_{\mu\nu}}{2}\right) + \Lambda g_{\mu\nu}. \end{aligned} \tag{5.85}$$

Synopsis of the laws of general relativity

We are now at a point where we can write down the laws of gravity in general relativity and summarise our conventions:

- Metric:
$$\mathrm{d}\tau^2 = g_{\mu\nu}(x)\mathrm{d}x^\mu \mathrm{d}x^\nu. \tag{5.86}$$
- The motion of a particle in a gravitational field is determined by
$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} + \Gamma^{\mu}{}_{\nu\rho}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} = 0. \tag{5.87}$$
- The affine connection is given by
$$\Gamma^{\alpha}{}_{\mu\nu} = \frac{1}{2}g^{\alpha\beta}\left[\frac{\partial g_{\beta\mu}}{\partial x^\nu} + \frac{\partial g_{\beta\nu}}{\partial x^\mu} - \frac{\partial g_{\mu\nu}}{\partial x^\beta}\right]. \tag{5.88}$$

- The metric can be determined by solving Einstein's field equations

$$R_{\mu\nu} - \frac{g_{\mu\nu}}{2}R + \Lambda g_{\mu\nu} = -8\pi G T_{\mu\nu}. \tag{5.89}$$

- The curvature scalar is given by

$$R = g^{\mu\nu}R_{\mu\nu}. \tag{5.90}$$

- The Ricci tensor is given by

$$R_{\mu\nu} = R^{\alpha}{}_{\mu\alpha\nu}. \tag{5.91}$$

- The Riemann tensor is given by

$$R^{\alpha}{}_{\mu\nu\rho} = \frac{\partial \Gamma^{\alpha}{}_{\mu\nu}}{\partial x^{\rho}} + \Gamma^{\beta}{}_{\mu\nu}\Gamma^{\alpha}{}_{\beta\rho} - \frac{\partial \Gamma^{\alpha}{}_{\mu\rho}}{\partial x^{\nu}} - \Gamma^{\beta}{}_{\mu\rho}\Gamma^{\alpha}{}_{\beta\nu}. \tag{5.92}$$

- For $T_{\mu\nu} = 0$ and $\Lambda = 0$ one solution of Einstein's equations is the Minkowski metric,

$$g_{\mu\nu} = \eta_{\mu\nu}. \tag{5.93}$$

- In our conventions we take

$$\eta_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1). \tag{5.94}$$

Important note: The Minkowski metric is not the only solution of Einstein's equations. As we will see, there are also solutions describing gravitational waves propagating in empty space.

5.3.1 Coordinate conditions

The metric $g_{\mu\nu}(x)$ is a symmetric 4×4 matrix and contains ten independent elements. Einstein's equations

$$R_{\mu\nu} - \frac{g_{\mu\nu}}{2}R + \Lambda g_{\mu\nu} + 8\pi G T_{\mu\nu} = 0, \tag{5.95}$$

are also ten in number; however, four of them are not independent, due to the four Bianchi identities

$$\left(R^{\mu\nu} - \tfrac{1}{2}g^{\mu\nu}R + \Lambda g^{\mu\nu} + 8\pi G T^{\mu\nu}\right)_{;\mu} = 0. \tag{5.96}$$

This means that Einstein's equations suffice to determine only six of the ten components of $g_{\mu\nu}(x)$. The situation is analogous to that of the Maxwell equations in special relativity, which read

$$(\partial^2 g_{\mu\nu} - \partial_\mu \partial_\nu)A^{\nu} = 0. \tag{5.97}$$

If $A^{\nu}_{\mathrm{sol}}(x)$ is a solution of the above, then so is

$$A'^{\nu}_{\mathrm{sol}}(x) = A^{\nu}_{\mathrm{sol}}(x) + \partial^{\nu} f(x). \tag{5.98}$$

This is because the left-hand side of (5.97) satisfies a Bianchi-type identity, namely

$$\frac{\partial}{\partial x^{\mu}}\left[\partial^2 g^{\mu\nu} - \partial^{\mu}\partial^{\nu}\right]A_{\nu} = (\partial^2\partial^{\nu} - \partial^2\partial^{\nu})A_{\nu} = 0. \tag{5.99}$$

In electrodynamics we remove this ambiguity in the solutions by fixing the gauge with an additional gauge-fixing condition. For example, if we find a solution $A^\nu_{\text{sol}}(x)$, we can always construct a solution

$$A'^\nu_{\text{sol}}(x) = A^\nu_{\text{sol}}(x) + \frac{\partial f(x)}{\partial x_\nu}, \tag{5.100}$$

which satisfies the Lorenz gauge condition

$$\partial_\nu A'^\nu_{\text{sol}} = 0, \tag{5.101}$$

by using as $f(x)$ a solution of the differential equation

$$\partial^2 f(x) = -\partial_\nu A^\nu_{\text{sol}}(x). \tag{5.102}$$

Let us now return to general relativity. For Einstein's equations, too, we need to fix the gauge in order to make their solutions unique.

One possible way to do this is by imposing the so-called harmonic coordinate conditions, which demand

$$\Gamma^\lambda \equiv \Gamma^\lambda{}_{\mu\nu} g^{\mu\nu} = 0. \tag{5.103}$$

Let us work out the left-hand side of this equation,

$$\begin{aligned}
\Gamma^\lambda &= \frac{g^{\mu\nu} g^{\lambda\alpha}}{2} \left[\frac{\partial g_{\nu\alpha}}{\partial x^\mu} + \frac{\partial g_{\mu\alpha}}{\partial x^\nu} - \frac{\partial g_{\mu\nu}}{\partial x^\alpha} \right] \\
&= g^{\mu\nu} g^{\lambda\alpha} \frac{\partial g_{\mu\alpha}}{\partial x^\nu} - g^{\lambda\alpha} \left(\frac{1}{2} g^{\mu\nu} \frac{\partial g_{\mu\nu}}{\partial x^\alpha} \right) \\
&= g^{\mu\nu} \left(-\frac{\partial g^{\lambda\alpha}}{\partial x^\nu} g_{\mu\alpha} \right) - g^{\lambda\alpha} \frac{1}{\sqrt{g}} \frac{\partial \sqrt{g}}{\partial x^\alpha} \\
&= -\frac{\partial g^{\lambda\alpha}}{\partial x^\alpha} - \frac{1}{\sqrt{g}} g^{\lambda\alpha} \frac{\partial \sqrt{g}}{\partial x^\alpha} \\
&= -\frac{1}{\sqrt{g}} \frac{\partial}{\partial x^\alpha} \left(\sqrt{g} g^{\lambda\alpha} \right).
\end{aligned} \tag{5.104}$$

Thus, the harmonic coordinate condition becomes

$$\frac{\partial}{\partial x^\alpha} \left(\sqrt{g} g^{\alpha\lambda} \right) = 0. \tag{5.105}$$

(Compare with the Lorenz gauge condition $\frac{\partial A^\alpha}{\partial x^\alpha} = 0$ of electrodynamics.) We now have to show that it is possible to satisfy the harmonic coordinate condition. Suppose that we find a solution of Einstein's equations for the metric $g_{\mu\nu}(x)$ with $\Gamma^\lambda \equiv \Gamma^\lambda{}_{\mu\nu} g^{\mu\nu}$ which does not satisfy the harmonic coordinate condition $\Gamma^\lambda = 0$. All it takes to construct a solution which does satisfy the condition is a coordinate transformation. Indeed, under a transformation $x^\mu \to x'^\mu = x'^\mu(x)$, we have

$$\begin{aligned}
\Gamma'^\lambda &= g'^{\mu\nu} \Gamma'^\lambda{}_{\mu\nu} \\
&= \frac{\partial x'^\mu}{\partial x^\rho} \frac{\partial x'^\nu}{\partial x^\sigma} g^{\rho\sigma} \left[\frac{\partial x'^\lambda}{\partial x^\alpha} \frac{\partial x^\beta}{\partial x'^\mu} \frac{\partial x^\gamma}{\partial x'^\nu} \Gamma^\alpha{}_{\beta\gamma} - \frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x^\beta}{\partial x'^\nu} \frac{\partial^2 x'^\lambda}{\partial x^\alpha \partial x^\beta} \right] \\
&= \frac{\partial x'^\lambda}{\partial x^\alpha} \Gamma^\alpha - \frac{\partial^2 x'^\lambda}{\partial x^\rho \partial x^\sigma} g^{\rho\sigma}.
\end{aligned} \tag{5.106}$$

To get to $\Gamma'^{\lambda} = 0$, our coordinate transformation needs to satisfy the differential equation

$$\frac{\partial^2 x'^{\lambda}}{\partial x^{\rho} \partial x^{\sigma}} g^{\rho\sigma} - \frac{\partial x'^{\lambda}}{\partial x^{\alpha}} \Gamma^{\alpha} = 0. \tag{5.107}$$

(Compare with the equation $\partial^2 f = -\partial_{\mu} A^{\mu}_{\text{sol}}$ for fixing the Lorenz gauge in electrodynamics.)

The d'Alembert operator in general relativity

Consider a scalar $\phi(x)$. Its covariant derivative is

$$\phi_{;\mu} = \frac{\partial \phi}{\partial x^{\mu}}. \tag{5.108}$$

Taking one more covariant derivative we have

$$\phi_{;\mu;\nu} = \frac{\partial^2 \phi}{\partial x^{\mu} \partial x^{\nu}} - \Gamma^{\alpha}{}_{\mu\nu} \frac{\partial \phi}{\partial x^{\alpha}}. \tag{5.109}$$

Multiplying with $g^{\mu\nu}$ yields

$$g^{\mu\nu} \phi_{;\mu;\nu} = g^{\mu\nu} \frac{\partial^2 \phi}{\partial x^{\mu} \partial x^{\nu}} - g^{\mu\nu} \Gamma^{\alpha}{}_{\mu\nu} \frac{\partial \phi}{\partial x^{\alpha}}, \tag{5.110}$$

or, using the definition of Γ^{λ},

$$\boxed{\Box \phi \equiv \phi^{;\mu}{}_{;\mu} = g^{\mu\nu} \frac{\partial^2 \phi}{\partial x^{\mu} \partial x^{\nu}} - \Gamma^{\alpha} \frac{\partial \phi}{\partial x^{\alpha}}} \tag{5.111}$$

In the above we recognise a generalisation of the d'Alembert operator

$$\Box \equiv g^{\mu\nu} \frac{\partial^2}{\partial x^{\mu} \partial x^{\nu}} - g^{\mu\nu} \Gamma^{\alpha}{}_{\mu\nu} \frac{\partial}{\partial x^{\alpha}}. \tag{5.112}$$

↘ of GR

In terms of this operator, (5.107) reads

$$\Box x'^{\lambda}(x) = 0, \tag{5.113}$$

which justifies the name of harmonic coordinate conditions.

5.3.2 Energy, momentum and angular momentum

In special relativity we could define the energy and momentum of isolated systems from the conservation equation of the energy-momentum tensor $\partial_{\mu} T^{\mu\nu} = 0$. In general relativity the energy-momentum tensor on the right-hand side of Einstein's equations

$$R_{\mu\nu} - \frac{g_{\mu\nu}}{2} R = -8\pi G T_{\mu\nu}, \tag{5.114}$$

is not conserved in the ordinary sense. Instead, we have

$$T^{\mu\nu}{}_{;\mu} = 0, \tag{5.115}$$

in which the covariant derivative includes the effects of gravitation in a concealed way. We want to define an energy-momentum tensor $t^{\mu\nu}$ which explicitly includes gravitation and which is conserved in the ordinary sense,

$$\partial_\mu t^{\mu\nu} = 0. \tag{5.116}$$

We will do this for gravitating systems confined within a finite volume. At large distances, where gravity falls off, the spacetime metric approaches the Minkowski metric. We can then decompose

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{5.117}$$

where $h_{\mu\nu} \to 0$ as $r \to \infty$, but it can be large at finite distances. Let us now separate the part of the left-hand side of Einstein's equations which is linear in $h_{\mu\nu}$:

$$G^{\mu\nu} \equiv R^{\mu\nu} - g^{\mu\nu}\frac{R}{2} = G^{(1)\mu\nu} + \Delta G^{\mu\nu}. \tag{5.118}$$

The linear part of the Ricci tensor is

$$R^{(1)}{}_{\mu\nu} = \frac{\eta^{\alpha\beta}}{2}\left[\frac{\partial^2 h_{\alpha\beta}}{\partial x^\mu \partial x^\nu} + \frac{\partial^2 h_{\mu\nu}}{\partial x^\alpha \partial x^\beta} - \frac{\partial^2 h_{\mu\alpha}}{\partial x^\beta \partial x^\nu} - \frac{\partial^2 h_{\nu\alpha}}{\partial x^\mu \partial x^\beta}\right], \tag{5.119}$$

which we represent schematically as

$$R^{(1)}{}_{\mu\nu} = \langle \bullet\!\!-\!\!\bullet | \mu\nu \rangle + \langle \mu\nu | \bullet\!\!-\!\!\bullet \rangle - \langle \mu \sqcap \nu \rangle - \langle \nu \sqcap \mu \rangle. \tag{5.120}$$

Also,

$$R^{(1)} = \eta^{\mu\nu} R^{(1)}{}_{\mu\nu} = 2\left(\langle \bullet\!\!-\!\!\bullet | \bullet\!\!-\!\!\bullet \rangle - \langle \sqcap \rangle\right). \tag{5.121}$$

In total,

$$\begin{aligned} G^{(1)\mu\nu} &= R^{(1)\mu\nu} - \eta^{\mu\nu}\frac{R^{(1)}}{2} \\ &= \langle \bullet\!\!-\!\!\bullet | \mu\nu \rangle + \langle \mu\nu | \bullet\!\!-\!\!\bullet \rangle - \langle \mu \sqcap \nu \rangle - \langle \nu \sqcap \mu \rangle - \eta^{\mu\nu} \langle \bullet\!\!-\!\!\bullet | \bullet\!\!-\!\!\bullet \rangle + \eta^{\mu\nu} \langle \sqcap \rangle \\ &= \frac{\partial}{\partial x^\rho} \langle \bullet\!\!-\!\!\bullet | \nu \rangle \eta^{\mu\rho} + \frac{\partial}{\partial x^\rho} \langle \mu\nu | \rho \rangle - \frac{\partial}{\partial x^\rho} \langle \mu\rho | \nu \rangle - \frac{\partial}{\partial x^\rho} \langle \nu \sqcap \rangle \eta^{\mu\rho} \\ &\qquad - \frac{\partial}{\partial x^\rho} \langle \bullet\!\!-\!\!\bullet | \rho \rangle \eta^{\mu\nu} + \frac{\partial}{\partial x^\rho} \langle \rho \sqcap \rangle \eta^{\mu\nu} \\ &= \frac{\partial}{\partial x^\rho} \left(A^{\mu\nu\rho} - A^{\mu\rho\nu}\right), \end{aligned} \tag{5.122}$$

with

$$A^{\mu\nu\rho} = \eta^{\mu\rho} \langle \bullet\!\!-\!\!\bullet | \nu \rangle + \langle \mu\nu | \rho \rangle + \eta^{\mu\nu} \langle \rho \sqcap \rangle, \tag{5.123}$$

or, explicitly,

$$A^{\mu\nu\rho} = \frac{1}{2}\left\{\eta^{\nu\beta}\eta^{\mu\rho}\frac{\partial h^\alpha{}_\alpha}{\partial x^\beta} + \frac{\partial h^{\mu\nu}}{\partial x^\beta}\eta^{\rho\beta} + \eta^{\mu\nu}\frac{\partial h^{\rho\alpha}}{\partial x^\alpha}\right\}. \tag{5.124}$$

Thus,

$$G^{(1)\mu\nu} \equiv R^{(1)\mu\nu} - \frac{\eta^{\mu\nu}}{2}R^{(1)} = \frac{\partial}{\partial x^\rho}\left(A^{\mu\nu\rho} - A^{\mu\rho\nu}\right). \tag{5.125}$$

Differentiating with $\frac{\partial}{\partial x^\nu}$, we have

$$\frac{\partial G^{(1)\mu\nu}}{\partial x^\nu} \equiv \frac{\partial}{\partial x^\nu}\left(R^{(1)\mu\nu} - \frac{\eta^{\mu\nu}R^{(1)}}{2}\right) = \underbrace{\frac{\partial^2}{\partial x^\nu \partial x^\rho}}_{\text{sym.}} \underbrace{(A^{\mu\nu\rho} - A^{\mu\rho\nu})}_{\text{antisym.}} = 0. \tag{5.126}$$

Let us now return to Einstein's equations and rewrite them as follows (setting $\Lambda = 0$):

$$\begin{aligned} G^{\mu\nu} &\equiv R^{\mu\nu} - \frac{g^{\mu\nu}}{2}R = -8\pi G T^{\mu\nu} \\ \Leftrightarrow\ G^{(1)\mu\nu} &= -8\pi G T^{\mu\nu} + G^{(1)\mu\nu} - G^{\mu\nu} \\ \Leftrightarrow\ G^{(1)\mu\nu} &= -8\pi G t^{\mu\nu}, \end{aligned} \tag{5.127}$$

with

$$t^{\mu\nu} = T^{\mu\nu} + \frac{1}{8\pi G}(G^{\mu\nu} - G^{(1)\mu\nu}) = T^{\mu\nu} + \frac{1}{8\pi G}\Delta G^{\mu\nu}. \tag{5.128}$$

Differentiating the rewritten Einstein equations (5.127), we obtain

$$\begin{aligned} \frac{\partial}{\partial x^\nu}G^{(1)\mu\nu} &= -8\pi G \frac{\partial}{\partial x^\nu}t^{\mu\nu} \\ 0 &= -8\pi G \frac{\partial}{\partial x^\nu}t^{\mu\nu} \\ \boxed{\frac{\partial}{\partial x^\nu}t^{\mu\nu} = 0} & \end{aligned} \tag{5.129}$$

showing that $t^{\mu\nu}$ is conserved. It is composed of the energy-momentum tensor of matter, of the electromagnetic field and of all other non-gravitational fields, and of the term $\Delta G^{\mu\nu} = G^{\mu\nu} - G^{(1)\mu\nu}$ in which the part of $G^{\mu\nu}$ which is linear in $h_{\mu\nu}$ is removed. Therefore, $\Delta G^{\mu\nu}$ is at least quadratic in $h_{\mu\nu}$. In a particle physics interpretation, such nonlinear terms correspond to self-interactions of particles, which in this case are the gravitons.

At large distances, we assume that $h_{\mu\nu}$ approaches the Newtonian limit. It should then behave as the Newtonian potential, falling off as

$$h_{\mu\nu} \xrightarrow{r\to\infty} \frac{1}{r}. \tag{5.130}$$

The derivatives of $h_{\mu\nu}$ should behave as

$$\frac{\partial h_{\mu\nu}}{\partial x^\nu} \xrightarrow{r\to\infty} \frac{1}{r^2}, \tag{5.131}$$

and thus

$$\Delta G^{\mu\nu} \xrightarrow{r\to\infty} \frac{1}{r^4}. \tag{5.132}$$

With these fall-offs, which guarantee convergence, we can integrate the continuity equation

$$\frac{\partial t^{\mu 0}}{\partial t} + \frac{\partial t^{\mu j}}{\partial x^j} = 0, \tag{5.133}$$

over a large volume V, bounded by a surface S far away. We find

$$\int_V \mathrm{d}^3x\, \frac{\partial t^{\mu 0}}{\partial t} = -\int_V \mathrm{d}^3x\, \frac{\partial t^{\mu j}}{\partial x^j} = -\int_S \mathrm{d}S\, n_j t^{\mu j}. \tag{5.134}$$

On the left-hand side we identify the time derivative of the total four-momentum of the system,

$$P^\mu = \int_V \mathrm{d}^3x\, t^{\mu 0}. \tag{5.135}$$

On the right-hand side we identify the flux of four-momentum through S. Using (5.127), the four-momentum itself can be written as

$$\begin{aligned} -8\pi G\, P^\mu &= \int_V \mathrm{d}^3x\, \frac{\partial}{\partial x^\rho}\left(A^{\mu 0\rho} - A^{\mu\rho 0}\right) \\ &= \int_V \mathrm{d}^3x\, \frac{\partial}{\partial x^j}\left(A^{\mu 0 j} - A^{\mu j 0}\right) \\ &= \int_S \mathrm{d}S\, n_j \left(A^{\mu 0 j} - A^{\mu j 0}\right), \end{aligned} \tag{5.136}$$

where we used the divergence theorem in the last step. For $\mu = 0$,

$$-8\pi G P^0 = \int_S \mathrm{d}S\, n_j \left(A^{00j} - A^{0j0}\right), \tag{5.137}$$

which, after inserting the expression for $A^{\mu\nu\rho}$, gives

$$P^0 = -\frac{1}{16\pi G}\int_S \mathrm{d}S\, n_j \left[\frac{\partial h^{jk}}{\partial x^k} - \frac{\partial h^{kk}}{\partial x^j}\right]. \tag{5.138}$$

For $\mu = i = 1, 2, 3$, we have

$$-8\pi G P^i = \int_S \mathrm{d}S\, n_j \left(A^{i0j} - A^{ij0}\right), \tag{5.139}$$

which similarly gives

$$P^i = \frac{1}{16\pi G}\int_S \mathrm{d}S\, n_j \left[\frac{\partial h^{ij}}{\partial t} + \frac{\partial h^{i0}}{\partial x^j} - \delta_{ij}\left(\frac{\partial h^{0k}}{\partial x^k} + \frac{\partial h^{kk}}{\partial t}\right)\right]. \tag{5.140}$$

Angular momentum is derived similarly. As $t^{\mu\nu}$ is symmetric, the following tensor is conserved,

$$M^{\mu\nu\rho} = t^{\mu\nu}x^\rho - t^{\mu\rho}x^\nu. \tag{5.141}$$

Indeed,

$$\begin{aligned} \frac{\partial}{\partial x^\mu} M^{\mu\nu\rho} &= \cancelto{0}{\frac{\partial t^{\mu\nu}}{\partial x^\mu}} x^\rho + t^{\mu\nu}\delta^\rho{}_\mu - \cancelto{0}{\frac{\partial t^{\mu\rho}}{\partial x^\mu}} x^\nu - t^{\mu\rho}\delta^\nu{}_\mu \\ &= t^{\rho\nu} - t^{\nu\rho} = 0. \end{aligned} \tag{5.142}$$

Hence, the angular momentum

$$J^{\rho\nu} = \int \mathrm{d}^3x\, M^{0\rho\nu} \tag{5.143}$$

is conserved.

6

The Schwarzschild solution of Einstein's field equations

Einstein's equations (with $\Lambda = 0$, negligible on the scales of this chapter)

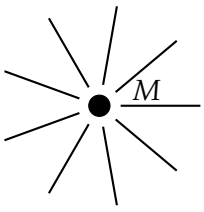
$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = -8\pi G T_{\mu\nu}, \tag{6.1}$$

or, equivalently,

$$R_{\mu\nu} = -8\pi G \left(T_{\mu\nu} - \frac{g_{\mu\nu}}{2} T^{\rho}{}_{\rho} \right), \tag{6.2}$$

are non-linear partial differential equations, which are hard to solve analytically. In general, they can be tackled numerically (with expensive computations) and, when appropriate, with approximate methods, typically expanding around the limit of weak gravitational fields.

We will now focus on a particular case in which an analytic solution, the so-called Schwarzschild solution, can be found. It corresponds to the physical situation of a spherical or point-like object of mass M, which produces an isotropic gravitational field (Fig. 6.1).



$T_{\mu\nu} = 0$ except at the centre,
static and isotropic metric.

Figure 6.1. A spherical mass M producing an isotropic gravitational field.

We will solve Einstein's equations away from the massive object, where the energy-momentum tensor vanishes. The equations then reduce to

$$R_{\mu\nu} = 0. \tag{6.3}$$

Proper time intervals will be given in terms of the metric as

$$\mathrm{d}\tau^2 = g_{\mu\nu}(x) \mathrm{d}x^{\mu} \mathrm{d}x^{\nu}. \tag{6.4}$$

Since the configuration is static, we expect

$$\frac{\partial}{\partial t} g_{\mu\nu} = 0, \tag{6.5}$$

and thus

$$g_{\mu\nu}(x) = g_{\mu\nu}(\vec{x}). \tag{6.6}$$

Due to isotropy, the proper time interval should be rotationally invariant. Its spatial part can then only be built from the rotational invariants

$$\mathrm{d}\vec{x}^2,\ \vec{x}^2 \equiv r^2,\ \vec{x}\cdot\mathrm{d}\vec{x}. \tag{6.7}$$

The most general ansatz is therefore

$$\mathrm{d}\tau^2 = \alpha(r)\mathrm{d}t^2 - \beta(r)\mathrm{d}t(\vec{x}\cdot\mathrm{d}\vec{x}) - \gamma(r)(\vec{x}\cdot\mathrm{d}\vec{x})^2 - \delta(r)\mathrm{d}\vec{x}^2. \tag{6.8}$$

We can now change to spherical coordinates,

$$x^3 = r\cos\theta,\ x^2 = r\sin\theta\sin\varphi,\ x^1 = r\sin\theta\cos\varphi. \tag{6.9}$$

This yields

$$\begin{aligned} \vec{x}^2 &= r^2, \\ \vec{x}\cdot\mathrm{d}\vec{x} &= r\mathrm{d}r, \\ \mathrm{d}\vec{x}^2 &= \mathrm{d}r^2 + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2\right). \end{aligned} \tag{6.10}$$

The metric becomes

$$\mathrm{d}\tau^2 = \alpha(r)\mathrm{d}t^2 - \beta(r)r\mathrm{d}r\mathrm{d}t - \gamma(r)\mathrm{d}r^2 - \delta(r)\left(\mathrm{d}r^2 + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\mathrm{d}\varphi^2\right). \tag{6.11}$$

Let us now change the time coordinate,

$$t \to t' = t + f(r). \tag{6.12}$$

Then,

$$\mathrm{d}t \to \mathrm{d}t' = \frac{\partial(t+f(r))}{\partial t}\mathrm{d}t + \frac{\partial(t+f(r))}{\partial r}\mathrm{d}r = \mathrm{d}t + f'(r)\mathrm{d}r. \tag{6.13}$$

and the metric becomes

$$\begin{aligned} \mathrm{d}\tau^2 &= \alpha(r)(\mathrm{d}t' - f'(r)\mathrm{d}r)^2 - \beta(r)r(\mathrm{d}t' - f'(r)\mathrm{d}r)\mathrm{d}r \\ &\quad - (\gamma(r)+\delta(r))\,\mathrm{d}r^2 - \delta(r)r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2) \\ &= \alpha(r)\mathrm{d}t'^2 - \mathrm{d}t'\mathrm{d}r\left(2\alpha(r)f'(r) + r\beta(r)\right) - \omega(r)\mathrm{d}r^2 \\ &\quad - \delta(r)r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2), \end{aligned} \tag{6.14}$$

with $\omega(r) \equiv \gamma(r) + \delta(r) - \alpha(r)f'^2(r) - r\beta(r)f'(r)$. We can now pick $f(r)$ such that

$$2\alpha(r)f'(r) + r\beta(r) = 0. \tag{6.15}$$

With this choice of coordinate system, the metric simplifies to

$$\mathrm{d}\tau^2 = \alpha(r)\mathrm{d}t'^2 - \omega(r)\mathrm{d}r^2 - \delta(r)r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2). \tag{6.16}$$

We can now perform one more change of variables,

$$r \longrightarrow r' = r\sqrt{\delta(r)}, \tag{6.17}$$

and rename the primed coordinates to unprimed ones. The metric then simplifies further to its standard form,

$$\mathrm{d}\tau^2 = B(r)\mathrm{d}t^2 - A(r)\mathrm{d}r^2 - r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2). \tag{6.18}$$

We can now read off the components of the metric from the above equation. The non-vanishing components are

$$g_{tt} = B(r),\ g_{rr} = -A(r),\ g_{\theta\theta} = -r^2,\ g_{\varphi\varphi} = -r^2\sin^2\theta. \tag{6.19}$$

All other components vanish. The metric must satisfy Einstein's field equations, which in vacuum take the form

$$R_{\mu\nu} = 0. \tag{6.20}$$

In what follows, we will construct the Ricci tensor which corresponds to the metric of Eq. (6.19). The inverse of the metric is also diagonal with elements

$$g^{tt} = 1/B(r),\ g^{rr} = -1/A(r),\ g^{\theta\theta} = -1/r^2,\ g^{\varphi\varphi} = -\frac{1}{r^2\sin^2\theta}. \tag{6.21}$$

The non-vanishing elements of the affine connection,

$$\Gamma^{\mu}_{\alpha\beta} = \frac{1}{2}g^{\mu\rho}\left[\frac{\partial g_{\alpha\rho}}{\partial x^{\beta}} + \frac{\partial g_{\rho\beta}}{\partial x^{\alpha}} - \frac{\partial g_{\alpha\beta}}{\partial x^{\rho}}\right], \tag{6.22}$$

are

$$\Gamma^{t}_{tr} = \Gamma^{t}_{rt} = \frac{\frac{d}{dr}B(r)}{2B(r)} \tag{6.23}$$

$$\Gamma^{r}_{tt} = \frac{\frac{d}{dr}B(r)}{2A(r)} \tag{6.24}$$

$$\Gamma^{r}_{rr} = \frac{\frac{d}{dr}A(r)}{2A(r)} \tag{6.25}$$

$$\Gamma^{r}_{\theta\theta} = -\frac{r}{A(r)} \tag{6.26}$$

$$\Gamma^{r}_{\varphi\varphi} = -\frac{r\left(\sin^2(\theta)\right)}{A(r)} \tag{6.27}$$

$$\Gamma^{\theta}_{r\theta} = \Gamma^{\theta}_{\theta r} = \frac{1}{r} \tag{6.28}$$

$$\Gamma^{\varphi}_{r\varphi} = \Gamma^{\varphi}_{\varphi r} = \frac{1}{r} \tag{6.29}$$

$$\Gamma^{\theta}_{\varphi\varphi} = -\sin(\theta)\cos(\theta) \tag{6.30}$$

$$\Gamma^{\varphi}_{\varphi\theta} = \Gamma^{\varphi}_{\theta\varphi} = \frac{\cos(\theta)}{\sin(\theta)} \tag{6.31}$$

With the affine connection at hand, we find that the Ricci tensor

$$R_{\mu\nu} = \frac{\partial\Gamma^{\alpha}_{\mu\alpha}}{\partial x^{\nu}} + \Gamma^{\beta}_{\mu\alpha}\Gamma^{\alpha}_{\beta\nu} - \frac{\partial\Gamma^{\alpha}_{\mu\nu}}{\partial x^{\alpha}} - \Gamma^{\beta}_{\mu\nu}\Gamma^{\alpha}_{\beta\alpha}, \tag{6.32}$$

has the following non-vanishing components

$$R_{\varphi\varphi} = \sin^2\theta\, R_{\theta\theta} \tag{6.33}$$

$$R_{\theta\theta} = \frac{r\left(\frac{d}{dr}B(r)\right)}{2A(r)\,B(r)} - \frac{r\left(\frac{d}{dr}A(r)\right)}{2A(r)^2} - 1 + \frac{1}{A(r)} \tag{6.34}$$

$$R_{rr} = -\frac{\left(\frac{d}{dr}B(r)\right)^2}{4B(r)^2} + \frac{\frac{d^2}{dr^2}B(r)}{2B(r)} - \frac{\left(\frac{d}{dr}A(r)\right)\left(\frac{d}{dr}B(r)\right)}{4A(r)\,B(r)} - \frac{\frac{d}{dr}A(r)}{A(r)\,r} \tag{6.35}$$

$$R_{tt} = \frac{\left(\frac{d}{dr}B(r)\right)\left(\frac{d}{dr}A(r)\right)}{4A(r)^2} - \frac{\frac{d^2}{dr^2}B(r)}{2A(r)} + \frac{\left(\frac{d}{dr}B(r)\right)^2}{4A(r)\,B(r)} - \frac{\frac{d}{dr}B(r)}{A(r)\,r} \tag{6.36}$$

$$\tag{6.37}$$

We notice that

$$-rA(r)\left(R_{tt}A(r) + R_{rr}B(r)\right) = \frac{d(A(r)B(r))}{dr}. \tag{6.38}$$

Since the left-hand side of the above vanishes due to Einstein's equations ($R_{rr} = R_{tt} = 0$), we obtain

$$A(r) = \frac{\kappa}{B(r)}, \quad \kappa = \text{constant}. \tag{6.39}$$

Substituting into $R_{\theta\theta}$, we find that

$$R_{\theta\theta} = \frac{\frac{d}{dr}\left(B(r)\,r\right)}{\kappa} - 1 = 0, \tag{6.40}$$

which has the solution

$$B(r) = \frac{c_1}{r} + \kappa. \tag{6.41}$$

The constants can be fixed by the expected behaviour at large distances, given by the Newtonian limit,

$$B(r) = g_{tt} \overset{r\to\infty}{\approx} 1 + 2\phi(r) = 1 - \frac{2GM}{r}. \tag{6.42}$$

We then have $c_1 = -2GM$ and $\kappa = 1$. This completely specifies the metric. The final result is the so-called Schwarzschild solution,

$$\mathrm{d}\tau^2 = \left[1 - \frac{r_S}{r}\right]\mathrm{d}t^2 - \frac{\mathrm{d}r^2}{1 - \frac{r_S}{r}} - r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2), \tag{6.43}$$

where the length

$$r_S \equiv 2GM \tag{6.44}$$

is the Schwarzschild radius, characterising the strength of the gravitational field due to the mass M located at the centre of our isotropic space.

Remark. The solution applies only outside the body ($r > R$), where $T_{\mu\nu} = 0$. The assumption of a static metric is in fact not needed (Birkhoff's theorem, Section 6.1), and the singular behaviour of g_{tt} and g_{rr} at $r = r_S$ is a coordinate artefact (Section 6.7).

6.1 Birkhoff's theorem

In our derivation we have assumed from the very beginning that the metric is static. This is a natural assumption for the field of a star which does not change with time. But what about a star which pulsates radially, or a spherical shell of matter which collapses? Remarkably, the answer is the same. This is the content of Birkhoff's theorem:

Every spherically symmetric solution of the vacuum field equations $R_{\mu\nu} = 0$ is, in a suitable coordinate system, the Schwarzschild metric of Eq. (6.43). In particular, it is static.

Let us sketch the proof, following the steps that we used above, but now keeping an arbitrary time dependence. Rotational invariance alone allows us to write the most general spherically symmetric proper time interval as

$$\mathrm{d}\tau^2 = C(t,r)\mathrm{d}t^2 - 2D(t,r)\mathrm{d}t\mathrm{d}r - H(t,r)\mathrm{d}r^2 - F(t,r)\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2\right), \tag{6.45}$$

where the four functions C, D, H, F may now depend on time as well. We simplify it in two steps.

(i) We define a new radial coordinate through $r'^2 = F(t,r)$, so that the angular part becomes $r'^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2)$. The area of a sphere of fixed r' is then $4\pi r'^2$, which gives the radial coordinate an invariant meaning. (This is possible as long as F is not constant in spacetime; the exceptional case does not admit an asymptotically flat solution.)

(ii) We remove the off-diagonal term with a new time coordinate. Writing $C\mathrm{d}t^2 - 2D\mathrm{d}t\mathrm{d}r = \frac{1}{C}\left(C\mathrm{d}t - D\mathrm{d}r\right)^2 - \frac{D^2}{C}\mathrm{d}r^2$, we choose an integrating factor $\chi(t,r)$ such that $\mathrm{d}t' = \chi\left(C\mathrm{d}t - D\mathrm{d}r\right)$ is an exact differential. Then $C\mathrm{d}t^2 - 2D\mathrm{d}t\mathrm{d}r = \mathrm{d}t'^2/(\chi^2 C) - (D^2/C)\mathrm{d}r^2$.

Renaming again the primed coordinates to unprimed ones, the metric takes the form

$$\mathrm{d}\tau^2 = e^{2a(t,r)}\mathrm{d}t^2 - e^{2b(t,r)}\mathrm{d}r^2 - r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2\right). \tag{6.46}$$

This is the standard form found above, with $B = e^{2a}$ and $A = e^{2b}$, but with functions which may depend on t. We now compute the Ricci tensor exactly as we did for the static case. Due to the time dependence, there is now a new non-vanishing off-diagonal component, and we find

$$R_{tr} = -\frac{2}{r}\frac{\partial b}{\partial t}, \tag{6.47}$$

$$e^{2(b-a)}R_{tt} + R_{rr} = -\frac{2}{r}\left(\frac{\partial a}{\partial r} + \frac{\partial b}{\partial r}\right), \tag{6.48}$$

$$R_{\theta\theta} = e^{-2b}\left[1 + r\left(\frac{\partial a}{\partial r} - \frac{\partial b}{\partial r}\right)\right] - 1. \tag{6.49}$$

(For a, b independent of t the second line is the combination $-rA\left(R_{tt}A + R_{rr}B\right) = \frac{\mathrm{d}(AB)}{\mathrm{d}r}$ that we used above.) We now impose the vacuum equations $R_{\mu\nu} = 0$ one at a time:

1. $R_{tr} = 0$ gives $\partial b/\partial t = 0$, so that $b = b(r)$ is time independent.

2. The second equation gives $\partial(a+b)/\partial r = 0$, i.e.

$$a(t,r) = -b(r) + h(t), \tag{6.50}$$

with an arbitrary function $h(t)$.

3. Substituting into $R_{\theta\theta} = 0$ we obtain $e^{-2b}(1-2rb') = 1$, or

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(re^{-2b}\right) = 1 \quad \Rightarrow \quad e^{-2b} = 1 - \frac{r_S}{r}, \tag{6.51}$$

where we have called the integration constant r_S.

The metric is therefore

$$\mathrm{d}\tau^2 = e^{2h(t)}\left(1-\frac{r_S}{r}\right)\mathrm{d}t^2 - \frac{\mathrm{d}r^2}{1-\frac{r_S}{r}} - r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2\right). \tag{6.52}$$

The only remnant of the time dependence is the factor $e^{2h(t)}$, which we can absorb in a new time coordinate $\mathrm{d}t' = e^{h(t)}\mathrm{d}t$. We arrive again at the Schwarzschild metric, Eq. (6.43), and the Newtonian limit identifies $r_S = 2GM$. (One can check that the remaining components R_{tt}, R_{rr} vanish as well, as they must due to the Bianchi identities.) □

Birkhoff's theorem has important physical consequences:

- The exterior gravitational field of a spherically symmetric body is static, even if the body pulsates, explodes or collapses radially. A spherically symmetric motion of matter therefore emits no gravitational waves; gravitational radiation needs a deviation from spherical symmetry (a time-dependent quadrupole moment, as we will see in Chapter 7).
- The field outside the body depends only on its total mass M, and not on how the mass is distributed radially; this is the relativistic analogue of Newton's shell theorem.
- Inside an empty spherical cavity at the centre of a spherically symmetric mass distribution, regularity at $r = 0$ requires $r_S = 0$, and spacetime is flat.

6.2 Trajectories

The metric of Eq. (6.43) can be used to approximate the gravitational field of the Sun in our planetary system, ignoring the gravity of the planets. To calculate the trajectory of a particle in this gravitational field, we use the geodesic equation

$$\frac{d^2x^\mu}{dp^2} + \Gamma^\mu_{\nu\rho}\frac{dx^\nu}{dp}\frac{dx^\rho}{dp} = 0, \tag{6.53}$$

with the coordinates

$$x^\mu = \{t, r, \theta, \varphi\}, \tag{6.54}$$

and the parameter p is the variable that we use to parametrise the trajectory $x^\mu(p)$ of the particle. For a start, we leave the evolution parameter unspecified. Once we have solved

the equations of motion in terms of the arbitrary parameter p, we will relate p to proper time via

$$\begin{aligned}\frac{d\tau^2}{dp^2} &= g_{\mu\nu}\frac{dx^\mu}{dp}\frac{dx^\nu}{dp} \\ &= B(r)\left(\frac{dt}{dp}\right)^2 - A(r)\left(\frac{dr}{dp}\right)^2 - r^2\left(\left(\frac{d\theta}{dp}\right)^2 + \sin^2\theta\left(\frac{d\varphi}{dp}\right)^2\right).\end{aligned} \tag{6.55}$$

This procedure will show that for a massive particle, p can be chosen to be proportional to proper time τ. However, for massless particles proper time intervals $d\tau$ vanish and a different variable needs to be chosen.

The geodesic equations (6.53) have the explicit form

$$\frac{d^2}{dp^2}t(p) + \frac{\left(\frac{d}{dr}B(r)\right)\left(\frac{d}{dp}t(p)\right)\left(\frac{d}{dp}r(p)\right)}{B(r)} = 0 \tag{6.56}$$

$$\begin{aligned}&\frac{d^2}{dp^2}r(p) + \frac{\left(\frac{d}{dr}B(r)\right)\left(\frac{d}{dp}t(p)\right)^2}{2A(r)} + \frac{\left(\frac{d}{dr}A(r)\right)\left(\frac{d}{dp}r(p)\right)^2}{2A(r)} \\ &\quad - \frac{r\left(\frac{d}{dp}\theta(p)\right)^2}{A(r)} - \frac{r\left(\sin^2(\theta)\right)\left(\frac{d}{dp}\varphi(p)\right)^2}{A(r)} = 0\end{aligned} \tag{6.57}$$

$$\frac{d^2}{dp^2}\theta(p) + \frac{2\left(\frac{d}{dp}r(p)\right)\left(\frac{d}{dp}\theta(p)\right)}{r} - \sin(\theta)\cos(\theta)\left(\frac{d}{dp}\varphi(p)\right)^2 = 0 \tag{6.58}$$

$$\frac{d^2}{dp^2}\varphi(p) + \frac{2\left(\frac{d}{dp}r(p)\right)\left(\frac{d}{dp}\varphi(p)\right)}{r} + \frac{2\cos(\theta)\left(\frac{d}{dp}\theta(p)\right)\left(\frac{d}{dp}\varphi(p)\right)}{\sin(\theta)} = 0 \tag{6.59}$$

We can simplify the problem of solving these equations by using isotropy: we can choose the coordinates such that the orbit lies in the plane $\theta = \frac{\pi}{2}$. We obtain

$$\frac{d^2}{dp^2}t(p) + \frac{\left(\frac{d}{dr}B(r)\right)\left(\frac{d}{dp}t(p)\right)\left(\frac{d}{dp}r(p)\right)}{B(r)} = 0 \tag{6.60}$$

$$\frac{d^2}{dp^2}r(p) + \frac{\left(\frac{d}{dr}B(r)\right)\left(\frac{d}{dp}t(p)\right)^2}{2A(r)} + \frac{\left(\frac{d}{dr}A(r)\right)\left(\frac{d}{dp}r(p)\right)^2}{2A(r)} - \frac{r\left(\frac{d}{dp}\varphi(p)\right)^2}{A(r)} = 0 \tag{6.61}$$

$$\frac{d^2}{dp^2}\varphi(p) + \frac{2\left(\frac{d}{dp}r(p)\right)\left(\frac{d}{dp}\varphi(p)\right)}{r} = 0 \tag{6.62}$$

By recognising that

$$\frac{dB(r)}{dp} = \frac{dB(r)}{dr}\frac{dr(p)}{dp},$$

we find that Eq. (6.60) takes the integrable form

$$\frac{d}{dp}\left[B(r)\frac{dt}{dp}\right] = 0. \tag{6.63}$$

This yields

$$\frac{dt}{dp} = \frac{f}{B(r)}, \tag{6.64}$$

where f is a constant. We will see that the value of this constant is physically immaterial.

Similarly, Eq. (6.62) takes the integrable form

$$\frac{d}{dp}\left(r^2\frac{d\varphi}{dp}\right) = 0, \tag{6.65}$$

which yields a constant of motion,

$$f\,J = r^2\frac{d\varphi}{dp}. \tag{6.66}$$

The integration constant J will turn out to be the analogue of the magnitude of the Newtonian angular momentum (the factor f on the left-hand side of Eq. (6.66) facilitates this identification).

Substituting the derivatives $\frac{d\varphi}{dp}$ and $\frac{dt}{dp}$ from Eqs. (6.66) and (6.64) into Eq. (6.61), we obtain an equation which can be brought to the integrable form (Exercise)

$$\frac{d}{dp}\left[A(r)\left(\frac{dr}{dp}\right)^2 + \frac{J^2f^2}{r^2} - \frac{f^2}{B}\right] = 0. \tag{6.67}$$

We therefore have

$$A(r)\left(\frac{dr}{dp}\right)^2\frac{1}{f^2} + \frac{J^2}{r^2} - \frac{1}{B} = -E, \tag{6.68}$$

where E is a further constant of motion.

We can now relate the parameter p to proper time more explicitly. Substituting $\theta = \frac{\pi}{2}$ and the results for $d\varphi/dp$, dr/dp, dt/dp from Eqs. (6.66), (6.64) and (6.68) into Eq. (6.55), we obtain

$$\left(\frac{d\tau}{dp}\right)^2 = Ef^2. \tag{6.69}$$

As anticipated, the evolution parameter p and proper time are linearly related. The positivity of the left-hand side of the above equation sets the constraint

$$E \geq 0. \tag{6.70}$$

For photons and other massless particles, $d\tau = 0$, which corresponds to

$$E_\gamma = 0. \tag{6.71}$$

For massive particles, we must have $E > 0$.

We can now trade the evolution parameter p for the time t,

$$\frac{dr(p)}{dp} = \frac{dr}{dt}\frac{dt}{dp} = \dot{r}\,\frac{f}{B(r)}, \tag{6.72}$$

and similarly,

$$\frac{d\varphi(p)}{dp} = \frac{d\varphi}{dt}\frac{dt}{dp} = \dot{\varphi}\,\frac{f}{B(r)}. \tag{6.73}$$

Substituting into Eqs. (6.66) and (6.68), we obtain

$$\frac{\dot{\varphi} r^2}{B(r)} = J, \tag{6.74}$$

and

$$\frac{A(r)}{B^2(r)}\dot{r}^2 + \left(\frac{J^2}{r^2} - \frac{1}{B(r)}\right) = -E. \tag{6.75}$$

From the above equations, we can derive the orbit equation, i.e. the dependence of the radius on the angle,

$$\frac{A(r)}{r^4}\left(\frac{dr}{d\varphi}\right)^2 + \left(\frac{1}{r^2} - \frac{1}{J^2 B(r)}\right) = -\frac{E}{J^2}. \tag{6.76}$$

To understand the physical meaning of the constants of motion E and J, let us expand our results around the Newtonian limit $r_S/r \to 0$, recalling that

$$B(r) = \frac{1}{A(r)} = 1 - \frac{r_S}{r}.$$

We find (Exercise)

$$J \approx r^2\,\dot{\varphi} \qquad \text{(Newtonian angular momentum)}, \tag{6.77}$$

and

$$\frac{1}{2}\dot{r}^2 + \frac{J^2}{2r^2} - \frac{\frac{r_S}{2}}{r} \approx \frac{1-E}{2} \qquad \text{(Newtonian energy)} \tag{6.78}$$

6.3 Gravitational redshift in the Schwarzschild field

In Chapter 2 we compared the rates of clocks at rest at different points of a static gravitational field, Eq. (2.45), and we found, for weak fields, a frequency shift $\nu_B \approx \nu_A(1+\phi(x_B)-\phi(x_A))$. We can now redo this computation exactly in the Schwarzschild field, using the constants of motion that we have just found.

Consider an observer who stays at rest at a fixed position (r, θ, φ); we call such an observer static. Its four-velocity follows from $\mathrm{d}\tau^2 = B(r)\mathrm{d}t^2$,

$$u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \left(\frac{1}{\sqrt{B(r)}}, 0, 0, 0\right). \tag{6.79}$$

A photon travels along a null geodesic $x^\mu(p)$, and its four-momentum is proportional to the tangent vector, $P^\mu = \lambda\,\frac{\mathrm{d}x^\mu}{\mathrm{d}p}$, with λ a constant. What frequency does the static observer measure? In the locally inertial frame of the observer, where $u^\mu = (1,0,0,0)$, the energy of the photon is the time component of its momentum, $h\nu = P^0 = \eta_{\mu\nu}P^\mu u^\nu$. The right-hand side is a scalar, so in our coordinates

$$h\nu = g_{\mu\nu}P^\mu u^\nu = \lambda\, B(r)\frac{\mathrm{d}t}{\mathrm{d}p}\,\frac{1}{\sqrt{B(r)}}. \tag{6.80}$$

But along any geodesic $B(r)\frac{\mathrm{d}t}{\mathrm{d}p} = f$ is a constant of motion, Eq. (6.64). Therefore

$$h\nu(r) = \frac{\lambda f}{\sqrt{B(r)}}, \qquad \text{i.e.} \qquad \nu(r)\sqrt{B(r)} = \text{constant along the light ray.} \tag{6.81}$$

A photon emitted by a static source at r_e with frequency ν_e (measured by the source) is received by a static observer at r_o with frequency

$$\boxed{\frac{\nu_o}{\nu_e} = \sqrt{\frac{B(r_e)}{B(r_o)}} = \sqrt{\frac{1-\frac{r_S}{r_e}}{1-\frac{r_S}{r_o}}}} \tag{6.82}$$

This is exactly the ratio $\sqrt{g_{00}(x_e)/g_{00}(x_o)}$ that we found in Eq. (2.45) and below, but now derived without any approximation and without the extra argument about the arrival times of successive wave crests: that argument is replaced by the conservation of f, which holds because the metric does not depend on t. (In the language of isometries, Chapter 9, $\xi^\mu = (1,0,0,0)$ is a Killing vector of the Schwarzschild metric and $f = g_{\mu\nu}\xi^\mu\frac{\mathrm{d}x^\nu}{\mathrm{d}p}$ is the associated conserved quantity.) For $r_S \ll r_e, r_o$ we recover the weak-field result,

$$\frac{\nu_o}{\nu_e} \approx 1 - \frac{GM}{r_e} + \frac{GM}{r_o} = 1 + \phi(r_e) - \phi(r_o). \tag{6.83}$$

Light climbing out of the potential well ($r_e < r_o$) is redshifted, light falling into it is blueshifted.

Let us look at some numbers.

- Sun. For light emitted at the solar surface and received far away, $\frac{\nu_o-\nu_e}{\nu_e} \approx -\frac{GM_\odot}{R_\odot c^2} = -\frac{1.477\,\mathrm{km}}{6.96\times10^5\,\mathrm{km}} \approx -2.1\times10^{-6}$, in agreement with our rough estimate in Chapter 2. The solar line shifts are contaminated by Doppler shifts of convective motions, but the gravitational redshift has been confirmed in the solar spectrum.

- Earth: the Pound–Rebka experiment. Near the surface of the Earth, for a height difference $h \ll R_\oplus$, $\phi(r_o) - \phi(r_e) \approx gh/c^2$. In 1960 Pound and Rebka sent 14.4 keV γ-rays of ^{57}Fe up and down the $h = 22.5\,\mathrm{m}$ tower of the Jefferson Laboratory at Harvard and measured the shift with the Mössbauer effect. The expected shift is tiny,
$$\frac{|\Delta\nu|}{\nu} = \frac{gh}{c^2} = \frac{9.81\,\mathrm{m\,s^{-2}}\times 22.5\,\mathrm{m}}{(3.0\times10^8\,\mathrm{m\,s^{-1}})^2} \approx 2.5\times10^{-15}, \tag{6.84}$$
and they found 1.05 ± 0.10 times this value (improved to the 1% level by Pound and Snider in 1965).

- Clocks in orbit. Today the effect is part of everyday technology: the clocks of the GPS satellites ($r \approx 26\,600\,\mathrm{km}$) run faster than clocks on the ground by about 46 μs per day due to the gravitational shift, partly compensated by about 7 μs per day of special-relativistic time dilation.

Finally, Eq. (6.82) has a striking feature: as the source approaches $r_e \to r_S$, the received frequency goes to zero. Light emitted from close to the Schwarzschild radius reaches distant observers infinitely redshifted. We will understand the meaning of this in Section 6.7.

6.4 Light deflection

For light, $E = 0$, and the orbit equation (6.76) gives

$$\varphi = \pm \int \frac{\sqrt{A}\mathrm{d}r}{r^2} \left(\frac{1}{J^2 B(r)} - \frac{1}{r^2} \right)^{-1/2} . \tag{6.85}$$

Let us assume a minimum distance $r = r_0$ ($\frac{\mathrm{d}r}{\mathrm{d}\varphi}\Big|_{r=r_0} = 0$) from the massive centre. Then,

$$\frac{1}{J^2 B(r_0)} - \frac{1}{r_0^2} = 0 \;\Rightarrow\; J^2 = \frac{r_0^2}{B(r_0)}, \tag{6.86}$$

and

$$\begin{aligned} \varphi &= \pm \int \sqrt{A(r)} \frac{\mathrm{d}r}{r^2} \left[\frac{B(r_0)}{r_0^2 B(r)} - \frac{1}{r^2} \right]^{-1/2} \\ &= \pm \int \frac{\mathrm{d}r}{r} \frac{r_0}{\sqrt{r^2 - r_0^2}} \left\{ 1 + \frac{r_S}{2r_0} \left[\frac{r(r + r_0) + r_0^2}{r(r + r_0)} \right] + O(r_S^2) \right\} \\ &= \pm \left[\text{const} - \arctan\left(\frac{r_0}{\sqrt{r^2 - r_0^2}} \right) + \frac{(2 + \frac{r_0}{r})(1 - \frac{r_0}{r})}{\sqrt{1 - \frac{r_0^2}{r^2}}} \frac{r_S}{2r_0} + O(r_S^2) \right] . \end{aligned} \tag{6.87}$$

With the above, we can compute (Exercise) the deflection angle $\Delta\varphi$ of light passing by the Sun (Fig. 6.2).

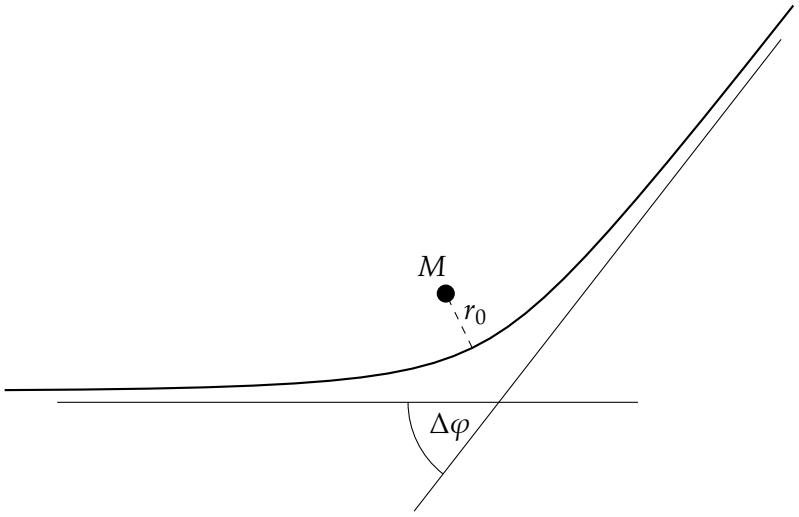


Figure 6.2. Deflection of a light ray passing a mass M at minimum distance r_0.

We find

$$\Delta\varphi = \frac{2r_S}{r_0} = \frac{4MG}{r_0}. \tag{6.88}$$

For a ray grazing the Sun, $\Delta\varphi \approx 1.75''$.

Remark. Here $\Delta\varphi \equiv 2|\varphi(r_0) - \varphi(\infty)| - \pi$, and r_0 is the coordinate radius of closest approach; to first order in r_S it coincides with the impact parameter $b = J$ (since $J^2 = r_0^2/B(r_0)$), so $\Delta\varphi = 4GM/(c^2 b)$. The value 1.75″ is for a ray grazing the solar limb, $r_0 = R_\odot$ (with $GM_\odot/c^2 = 1.477\,\mathrm{km}$, $R_\odot = 6.96 \times 10^5\,\mathrm{km}$ one gets 1.751″); it is twice the "Newtonian" value $2GM/(c^2 b)$ (Weinberg Sec. 8.5).

6.5 Time delay of radar echoes

The deflection of light tells us that light rays are bent by the gravitational field. In 1964 I. Shapiro pointed out that they are also delayed: a radar signal sent from the Earth to a planet (or a spacecraft) and reflected back takes longer when it passes close to the Sun. This is the "fourth test" of general relativity, and it has become the most precise one in the solar system.

We compute the coordinate time that light needs to travel from the point of closest approach r_0 to a point at radius r. For light $E = 0$, and Eq. (6.75) becomes

$$\frac{A(r)}{B^2(r)}\dot{r}^2 + \frac{J^2}{r^2} - \frac{1}{B(r)} = 0. \tag{6.89}$$

At the closest approach $\dot{r} = 0$, which fixes, as in the light deflection computation, $J^2 = r_0^2/B(r_0)$. Solving for $\mathrm{d}t/\mathrm{d}r = 1/\dot{r}$, we find

$$t(r, r_0) = \int_{r_0}^{r} \mathrm{d}r' \sqrt{\frac{A(r')}{B(r')}} \left[1 - \frac{B(r')}{B(r_0)}\frac{r_0^2}{r'^2}\right]^{-1/2}. \tag{6.90}$$

Without gravity, $A = B = 1$ and $t(r, r_0) = \sqrt{r^2 - r_0^2}$, which is just the length of a straight line (Pythagoras). Let us now expand the integrand to first order in r_S. We have $\sqrt{A/B} = 1/B \approx 1 + \frac{r_S}{r}$ and

$$1 - \frac{B(r)}{B(r_0)}\frac{r_0^2}{r^2} \approx 1 - \frac{r_0^2}{r^2}\left[1 + r_S\left(\frac{1}{r_0} - \frac{1}{r}\right)\right] = \left(1 - \frac{r_0^2}{r^2}\right)\left[1 - \frac{r_S r_0}{r(r + r_0)}\right]. \tag{6.91}$$

Therefore

$$t(r, r_0) \approx \int_{r_0}^{r} \frac{\mathrm{d}r'}{\sqrt{1 - r_0^2/r'^2}} \left[1 + \frac{r_S}{r'} + \frac{r_S\, r_0}{2r'(r' + r_0)}\right]. \tag{6.92}$$

The integrals are elementary (one can check the result by differentiating it with respect to r), and we obtain

$$t(r, r_0) = \sqrt{r^2 - r_0^2} + r_S \ln\left(\frac{r + \sqrt{r^2 - r_0^2}}{r_0}\right) + \frac{r_S}{2}\sqrt{\frac{r - r_0}{r + r_0}} + O(r_S^2). \tag{6.93}$$

The two extra terms are the time delay with respect to straight-line propagation at the speed of light. Physically, the "coordinate speed" of light, $|\mathrm{d}r/\mathrm{d}t| = B(r)$ for radial rays, is smaller than one where the field is strong.

Consider now a radar signal sent from the Earth, at $r_\oplus$, to a planet at r_P on the other side of the Sun ("superior conjunction"), reflected and received back on Earth. The ray grazes the Sun at $r_0 \ll r_\oplus, r_P$. The round trip takes $2\,[t(r_\oplus, r_0) + t(r_P, r_0)]$, and the excess with respect to the flat-space value is, for $r_0 \ll r_\oplus, r_P$,

$$\boxed{\Delta t \approx 2r_S \left[\ln\left(\frac{4r_\oplus r_P}{r_0^2}\right) + 1\right] = 4GM\left[\ln\left(\frac{4r_\oplus r_P}{r_0^2}\right) + 1\right]} \tag{6.94}$$

(The Earth clock measures proper time $\mathrm{d}\tau = \sqrt{B(r_\oplus)}\mathrm{d}t$, which differs from t by a relative 10^{-8}, a negligible correction to Δt.)

For Venus ($r_P = 1.08 \times 10^8\,\mathrm{km}$, $r_\oplus = 1.50 \times 10^8\,\mathrm{km}$) and a ray grazing the solar limb, $r_0 = R_\odot = 6.96 \times 10^5\,\mathrm{km}$, we have $4GM_\odot/c^3 = 19.7\,\mu\mathrm{s}$ and $\ln\!\left(4r_\oplus r_P/R_\odot^2\right) \approx 11.8$, so that

$$\Delta t \approx 19.7\,\mu\mathrm{s} \times 12.8 \approx 2.5 \times 10^{-4}\,\mathrm{s}. \tag{6.95}$$

In practice the radar beam cannot pass arbitrarily close to the Sun, and the delays at superior conjunction are of order $200\,\mu$s, compared to a total round-trip time of about half an hour. Note that Δt is defined with respect to the coordinate distance and is not directly observable. What one measures is the characteristic logarithmic variation of the echo time as the impact parameter r_0 changes through the conjunction. Shapiro and collaborators measured it with radar echoes from Mercury and Venus in the late 1960s and early 1970s, confirming the prediction at the level of a few per cent; the Viking landers on Mars brought this to 0.1% in 1979.

Parametrised post-Newtonian (PPN) test. To quantify how well such experiments test general relativity, it is useful to parametrise the metric of the Sun more generally, to the order in GM/r that is relevant (Eddington–Robertson expansion, Weinberg Sec. 8.3),

$$\begin{aligned}\mathrm{d}\tau^2 = {} & \left[1 - \frac{2GM}{r} + 2(\beta-\gamma)\frac{G^2M^2}{r^2} + \ldots\right]\mathrm{d}t^2 - \left[1 + 2\gamma\frac{GM}{r} + \ldots\right]\mathrm{d}r^2 \\ & - r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2\right),\end{aligned} \tag{6.96}$$

where general relativity predicts $\beta = \gamma = 1$. Repeating the computation of Eq. (6.90) with these A and B, one finds that the coefficient of the logarithm in Eq. (6.93) becomes $(1+\gamma)GM$ instead of $2GM$, so that the delay (and also the light deflection) scales with

$$\frac{1+\gamma}{2}. \tag{6.97}$$

The most precise measurement to date used the Cassini spacecraft on its way to Saturn in 2002: rather than the delay itself, the frequency shift of the radio link, i.e. the time derivative of the delay, was measured as the line of sight passed the Sun. The result (Bertotti, Iess and Tortora, 2003) is

$$\gamma - 1 = (2.1 \pm 2.3) \times 10^{-5}, \tag{6.98}$$

i.e. $(1+\gamma)/2$ agrees with the general relativity value 1 at the 10^{-5} level.

6.6 Bound orbits

A bound orbit has two radial turning points, the minimum distance r_- and the maximum distance r_+ from the central mass (Fig. 6.3). The angle swept from r_- to r follows from the orbit equation (6.76),

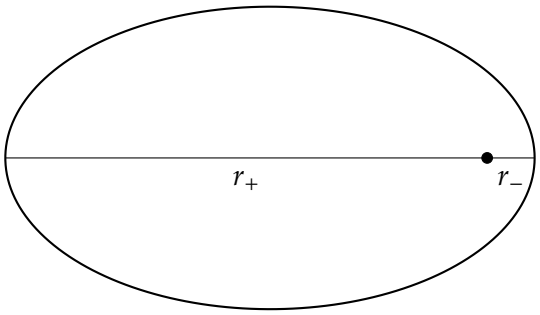


Figure 6.3. A bound orbit with minimum and maximum distances r_- and r_+ from the central mass.

$$\varphi(r) = \varphi(r_-) + \int_{r_-}^{r} \frac{\sqrt{A(r')}}{r'^2} \mathrm{d}r' \left[\frac{1}{J^2 B(r')} - \frac{E}{J^2} - \frac{1}{r'^2} \right]^{-1/2} . \tag{6.99}$$

We can express E and J^2 in terms of r_+ and r_-, and then expand for small r_S. One finds

$$\Delta\varphi = 2|\varphi(r_+) - \varphi(r_-)| = 2\pi + 3\pi MG \left(\frac{1}{r_+} + \frac{1}{r_-} \right) . \tag{6.100}$$

For Mercury $\Delta\varphi_{\text{precession}} = (\Delta\varphi - 2\pi) \approx \frac{0.1035''}{\text{revolution}}$. In a century (about 415 revolutions), this gives a precession of $42.98'' \approx 43''$.

Remark. Since $\frac{1}{r_+} + \frac{1}{r_-} = \frac{2}{a(1-e^2)}$, one has $\Delta\varphi - 2\pi = 6\pi GM/(c^2 a(1-e^2))$. This is only the relativistic part; the observed perihelion advance of Mercury is about $574''$/century, of which about $531''$ are due to perturbations by the other planets (Weinberg Sec. 8.6; Will, Living Rev. Rel. 17 (2014) 4).

6.7 The Schwarzschild radius and the horizon

For the Sun the Schwarzschild radius is $r_S = 2GM_\odot/c^2 \approx 2.95\,\text{km}$, deep inside the Sun, where our vacuum solution does not apply. For the Earth it is about 9 mm. But nothing forbids a body to be more compact than its Schwarzschild radius, and then the region $r \le r_S$ is part of the vacuum solution. What happens there? The metric of Eq. (6.43) looks alarming at two places: g_{tt} vanishes and g_{rr} diverges at $r = r_S$, and both diverge at $r = 0$.

6.7.1 Coordinate singularities and curvature singularities

A divergence of metric components may simply be due to a bad choice of coordinates. A familiar example is the flat plane in polar coordinates, $\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\varphi^2$, where $g^{\varphi\varphi} = 1/r^2$ diverges at $r = 0$, although nothing is special there. To decide whether spacetime itself is singular, we should look at quantities which do not depend on the coordinates, i.e. scalars

built from the curvature. The Ricci scalar vanishes in vacuum, so the simplest non-trivial one is the Kretschmann scalar. For the Schwarzschild metric one finds (a straightforward but lengthy computation with the Christoffel symbols listed above)

$$R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} = \frac{12\,r_S^2}{r^6}. \tag{6.101}$$

It is perfectly finite at $r = r_S$, but it diverges at $r = 0$. Tidal forces, which are measured by the Riemann tensor, are therefore finite at the Schwarzschild radius; for a black hole of large mass they are even very weak there, since $12r_S^2/r^6 = 12/r_S^4$ at $r = r_S$. In contrast, $r = 0$ is a genuine curvature singularity, where general relativity itself ceases to be applicable. We will now show that $r = r_S$ is merely a coordinate singularity, by following an observer who falls in, and then by finding coordinates in which the metric is regular there.

6.7.2 Radial free fall

Consider a massive particle which is released from rest at $r = r_0$ and falls radially, $J = 0$. From Eq. (6.68) with $J = 0$ and $\frac{\mathrm{d}r}{\mathrm{d}p} = 0$ at $r = r_0$, we find $E = 1/B(r_0)$. Using $\left(\frac{\mathrm{d}\tau}{\mathrm{d}p}\right)^2 = Ef^2$ and $AB = 1$, Eq. (6.68) gives

$$\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 = \frac{1}{E} - B(r) = \frac{r_S}{r} - \frac{r_S}{r_0}. \tag{6.102}$$

Amusingly, this is exactly Newton's law for the radial fall from rest, $\frac{1}{2}\dot{r}^2 - \frac{GM}{r} = -\frac{GM}{r_0}$, provided that we use the proper time of the falling particle and the Schwarzschild radial coordinate. Nothing happens at $r = r_S$. The solution is a cycloid (Exercise),

$$r = \frac{r_0}{2}\left(1 + \cos\eta\right), \qquad \tau = \frac{r_0}{2}\sqrt{\frac{r_0}{r_S}}\left(\eta + \sin\eta\right), \qquad 0 \le \eta \le \pi, \tag{6.103}$$

and the particle reaches the singularity $r = 0$ after the finite proper time

$$\tau(r_0 \to 0) = \frac{\pi}{2}\sqrt{\frac{r_0^3}{r_S}}. \tag{6.104}$$

Let us now describe the same fall with the Schwarzschild time t, which is the proper time of a static observer at large distances. From Eq. (6.64), $\frac{\mathrm{d}t}{\mathrm{d}\tau} = \frac{\mathrm{d}t/\mathrm{d}p}{\mathrm{d}\tau/\mathrm{d}p} = \frac{1}{B(r)\sqrt{E}} = \frac{\sqrt{B(r_0)}}{B(r)}$, and therefore

$$\frac{\mathrm{d}t}{\mathrm{d}r} = -\frac{\sqrt{B(r_0)}}{B(r)\sqrt{\frac{r_S}{r} - \frac{r_S}{r_0}}} \overset{r\to r_S}{\approx} -\frac{r_S}{r - r_S} \quad \Rightarrow \quad t \approx -r_S \ln\frac{r - r_S}{r_S} + \text{constant}. \tag{6.105}$$

The coordinate time diverges logarithmically as the particle approaches r_S: for the distant observer, the particle never crosses the Schwarzschild radius, it approaches it exponentially, $r - r_S \propto e^{-t/r_S}$. At the same time, by Eq. (6.82), the light that the particle emits reaches the distant observer more and more redshifted (and more and more rarely), so that the particle appears to freeze and fade away. Both descriptions are correct; they only show that t is a bad coordinate at $r = r_S$.

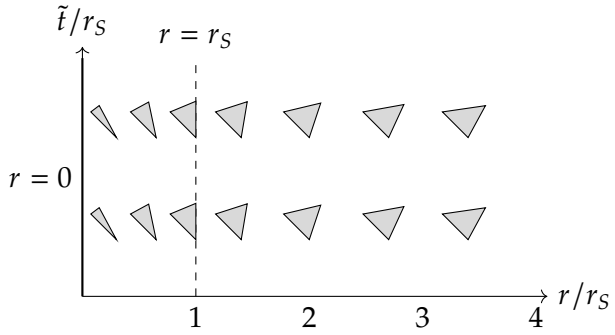


Figure 6.4. Light cones in ingoing Eddington–Finkelstein coordinates, $\tilde{t} = v - r$. The left edge of each cone is an ingoing light ray, the right edge an outgoing one. At $r = r_S$ the outgoing ray stays at constant r; inside, both edges point towards $r = 0$.

6.7.3 Eddington–Finkelstein coordinates

To find better coordinates, let us adapt them to light rays. For radial light rays, $\mathrm{d}\tau = \mathrm{d}\theta = \mathrm{d}\varphi = 0$, and Eq. (6.43) gives $\mathrm{d}t = \pm \mathrm{d}r/B(r)$. We define the "tortoise coordinate"

$$r^* = r + r_S \ln\left|\frac{r}{r_S} - 1\right|, \qquad \frac{\mathrm{d}r^*}{\mathrm{d}r} = \frac{1}{1 - \frac{r_S}{r}}, \tag{6.106}$$

so that radial light rays satisfy $t \mp r^* =$ constant: outgoing rays have constant $u \equiv t - r^*$ and ingoing rays have constant $v \equiv t + r^*$. (Near r_S, $r^* \to -\infty$: an ingoing ray needs infinite t to reach r_S, like the falling particle.) We now replace t by the "advanced time" v. With $\mathrm{d}t = \mathrm{d}v - \mathrm{d}r/B(r)$, the proper time interval becomes

$$\mathrm{d}\tau^2 = \left(1 - \frac{r_S}{r}\right)\mathrm{d}v^2 - 2\mathrm{d}v\mathrm{d}r - r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2\right). \tag{6.107}$$

These are the ingoing Eddington–Finkelstein coordinates. All metric components are finite at $r = r_S$ and the determinant of the metric is $-r^4\sin^2\theta \neq 0$, so the metric is perfectly regular there; it is singular only at $r = 0$. The coordinates (v, r, θ, φ) cover smoothly the region $r > r_S$ and the region $0 < r < r_S$.

Let us look at the radial light rays in these coordinates. Setting $\mathrm{d}\tau = 0$ in Eq. (6.107) gives two solutions:

$$\mathrm{d}v = 0 \quad \text{(ingoing)}, \qquad \frac{\mathrm{d}r}{\mathrm{d}v} = \frac{1}{2}\left(1 - \frac{r_S}{r}\right) \quad \text{("outgoing")}. \tag{6.108}$$

For $r > r_S$ the second family moves outwards. At $r = r_S$ it stays at constant radius, and for $r < r_S$ it moves inwards as well. The light cones tip over as we approach the Schwarzschild radius, as shown in Fig. 6.4, where we have used $\tilde{t} = v - r$ as time coordinate, so that ingoing rays are at 45° ($\mathrm{d}r/\mathrm{d}\tilde{t} = -1$) and outgoing rays have $\mathrm{d}r/\mathrm{d}\tilde{t} = (r - r_S)/(r + r_S)$. Since a massive particle moves inside its light cone, for $r < r_S$ every particle, whatever its engine, must move towards smaller r: inside r_S, r behaves as a time coordinate, and $r = 0$ lies in the future of every observer. No signal from $r < r_S$ can ever reach $r > r_S$.

6.7.4 Kruskal–Szekeres coordinates

The ingoing coordinates treat ingoing and outgoing rays asymmetrically. Using both u and v, the Schwarzschild metric takes the form $\mathrm{d}\tau^2 = \left(1 - \frac{r_S}{r}\right)\mathrm{d}u\mathrm{d}v - r^2\mathrm{d}\Omega^2$, with

$d\Omega^2 \equiv d\theta^2 + \sin^2\theta d\varphi^2$, which is again singular at r_S because $u \to +\infty$ there. Kruskal and Szekeres (1960) cured this with the exponential maps $U = -e^{-u/(2r_S)}$ and $V = e^{v/(2r_S)}$, and defined $T = (V+U)/2$, $X = (V-U)/2$. For $r > r_S$ this gives

$$T = \sqrt{\frac{r}{r_S} - 1}\, e^{r/(2r_S)} \sinh\frac{t}{2r_S}, \qquad X = \sqrt{\frac{r}{r_S} - 1}\, e^{r/(2r_S)} \cosh\frac{t}{2r_S}. \tag{6.109}$$

A direct computation (Exercise) shows that the metric becomes

$$d\tau^2 = \frac{4r_S^3}{r} e^{-r/r_S} \left(dT^2 - dX^2\right) - r^2 \left(d\theta^2 + \sin^2\theta d\varphi^2\right), \tag{6.110}$$

where r is now a function of T and X, defined implicitly by

$$X^2 - T^2 = \left(\frac{r}{r_S} - 1\right) e^{r/r_S}, \qquad \text{and} \qquad \frac{T}{X} = \tanh\frac{t}{2r_S}. \tag{6.111}$$

This metric has remarkable properties:

- The prefactor $4r_S^3 e^{-r/r_S}/r$ is finite and non-zero at $r = r_S$ (where it equals $4r_S^2/e$). The metric is regular everywhere except at $r = 0$.
- In the (T, X) plane the metric is proportional to the two-dimensional Minkowski metric. Radial light rays are therefore straight lines at 45°, $T = \pm X +$ constant, as in special relativity, and the causal structure can be read off directly.
- The right-hand side of Eq. (6.111) is a smooth function of r, and it takes the value -1 at $r = 0$. We can thus extend the coordinates to the whole region $T^2 - X^2 < 1$. The Schwarzschild radius $r = r_S$ corresponds to the lines $T = \pm X$, and the singularity $r = 0$ to the hyperbola $T^2 - X^2 = 1$. Inside, $0 < r < r_S$, the transformation reads $T = \sqrt{1 - \frac{r}{r_S}}\, e^{r/(2r_S)} \cosh\frac{t}{2r_S}$, $X = \sqrt{1 - \frac{r}{r_S}}\, e^{r/(2r_S)} \sinh\frac{t}{2r_S}$, and one checks that it also leads to Eq. (6.110).

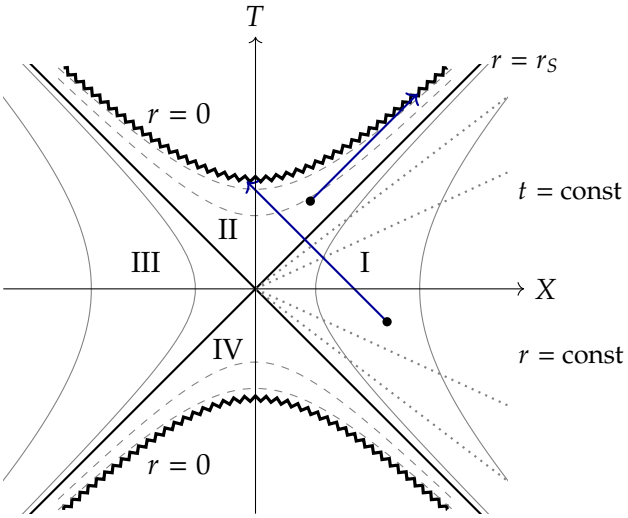


Figure 6.5. Kruskal–Szekeres diagram (θ, φ suppressed). Solid gray: $r = 1.1\,r_S$, $1.5\,r_S$; dashed: $r = 0.8\,r_S$, $0.5\,r_S$; dotted: $t = \pm r_S, \pm 2r_S$; zigzag: the singularity $r = 0$. The light rays (arrows) start in region I and in region II; both end on the future singularity.

The Kruskal–Szekeres diagram, Fig. 6.5, contains four regions. Region I ($X > |T|$) is our original exterior, $r > r_S$, with the lines of constant t through the origin and the lines of constant r as hyperbolae; $t \to +\infty$ is the line $T = X$. Region II ($T > |X|$) is the interior $0 < r < r_S$, which we have reached already with Eddington–Finkelstein coordinates. Regions III and IV are a second exterior and a "white hole" (the time reverse of region II); they appear in the eternal, mathematically maximally extended, solution, but not for a black hole formed by the collapse of a star, where the vacuum metric describes only the region outside the star's surface.

6.7.5 What is a black hole?

We can now state precisely what is special about $r = r_S$. Any light ray or massive particle in region II stays in region II and ends on the singularity $r = 0$, since the light cones in the diagram are at 45° and the singularity is a spacelike line in the future. Nothing which is in region II can ever send a signal to region I. The boundary $T = X$ of the region from which signals can escape to large distances is called the event horizon, and the region inside it is a black hole. The event horizon is not a material surface: an observer falling through it does not notice anything locally special (the curvature is finite, Eq. (6.101)). But once inside, the observer reaches $r = 0$ within a proper time not longer than $\pi r_S/2 = \pi GM$ (this is Eq. (6.104) for $r_0 = r_S$, which is the maximum), about 1.5×10^{-4} s for $M = 10\,M_\odot$ and about one minute for $M = 4 \times 10^6\,M_\odot$.

By Birkhoff's theorem, a spherical star which collapses below its Schwarzschild radius leaves exactly this geometry outside. Real stars rotate; the black holes they form are described by the Kerr solution, which we will not discuss, but which shares the features of the horizon and the inevitable fall inside it.

Observational status. Black holes are by now observed in several ways. Stars orbiting the compact radio source Sgr A* at the centre of the Milky Way reveal a mass of about $4 \times 10^6\,M_\odot$ within a region smaller than the solar system (Nobel Prize in Physics 2020 to R. Genzel and A. Ghez, together with R. Penrose for the theory of black hole formation). Since 2015 the LIGO and Virgo detectors observe gravitational waves from the inspiral and merger of black holes, starting with GW150914 (two black holes of about 36 and 29 $M_\odot$ merging into one of about 62 $M_\odot$); the signals after the merger agree with the ringing of a Kerr black hole. The Event Horizon Telescope has imaged the shadow of the black holes M87* (2019, $6.5 \times 10^9\,M_\odot$) and Sgr A* (2022), with sizes in agreement with general relativity (for a non-rotating black hole the shadow has diameter $3\sqrt{3}\,r_S$).

6.8 Exercises

Exercise.. (Radar echo from Mars.) (a) Derive Eq. (6.93) by differentiating its right-hand side with respect to r and comparing with the expanded integrand. (b) Show that for $r_0 \ll r_\oplus, r_P$ the round-trip delay is given by Eq. (6.94). (c) Evaluate it for Mars at superior conjunction ($r_P = 2.28 \times 10^8$ km) and a ray grazing the Sun. (d) Repeat the computation of Eq. (6.93) with $B = 1 - 2GM/r$, $A = 1 + 2\gamma GM/r$ and show that the logarithmic term is multiplied by $(1+\gamma)/2$. [Answer for (c): $\Delta t \approx 19.7\,\mu\text{s} \times 13.5 \approx 2.7 \times 10^{-4}$ s.]

Exercise.. (Falling into a black hole.) (a) Verify that the cycloid $r = \frac{r_0}{2}(1 + \cos\eta)$, $\tau = \frac{r_0}{2}\sqrt{r_0/r_S}\,(\eta + \sin\eta)$ solves Eq. (6.102), and derive Eq. (6.104). (b) An observer falls from rest at $r_0 = 10\,r_S$ into a black hole of $10\,M_\odot$. How long does the fall to $r = 0$ take on the observer's watch? How long does the part from r_S to 0 take? (c) Show that the Schwarzschild time t needed to reach r_S is infinite. [Answers: (b) 4.9 ms; about 0.07 ms.]

Exercise.. (Gravitational redshift of white dwarfs and neutron stars.) (a) The white dwarf Sirius B has $M = 1.02\,M_\odot$ and $R = 0.0080\,R_\odot$. Compute the redshift $z = \nu_e/\nu_o - 1$ of its spectral lines for a distant observer, and express it as an equivalent Doppler velocity cz. (Measurements with the Hubble Space Telescope find about 80 km/s.) (b) Compute z exactly and in the weak-field approximation for a neutron star with $M = 1.4\,M_\odot$ and $R = 12$ km. Is the weak-field approximation adequate? (c) Show, using Eqs. (6.109) and (6.111), that the Kruskal–Szekeres transformation turns the Schwarzschild metric into Eq. (6.110). [Answers: (a) $z \approx 2.7 \times 10^{-4}$, $cz \approx 81$ km/s; (b) $z = 0.235$ exactly, 0.172 in the weak-field approximation.]

7

Gravitational waves

The Einstein equations (with $\Lambda = 0$)

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = -8\pi G T_{\mu\nu}, \tag{7.1}$$

or, equivalently,

$$R_{\mu\nu} = -8\pi G S_{\mu\nu}, \tag{7.2}$$

with

$$S_{\mu\nu} = T_{\mu\nu} - \frac{g_{\mu\nu}}{2} T^{\alpha}{}_{\alpha}, \tag{7.3}$$

admit wave solutions. We can see this by looking at these equations in the weak-field approximation, in which the metric is the Minkowski metric plus a small perturbation,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{7.4}$$

with $h_{\mu\nu}$ small. The Ricci tensor is

$$\begin{aligned} R_{\mu\nu} &= \frac{\partial \Gamma^{\alpha}{}_{\mu\alpha}}{\partial x^{\nu}} - \frac{\partial \Gamma^{\alpha}{}_{\mu\nu}}{\partial x^{\alpha}} + O(h^2) \\ &= \frac{\eta^{\alpha\rho}}{2} \left[\frac{\partial}{\partial x^{\nu}} \left(\cancel{\frac{\partial h_{\mu\rho}}{\partial x^{\alpha}}} + \frac{\partial h_{\rho\alpha}}{\partial x^{\mu}} - \frac{\partial h_{\mu\alpha}}{\partial x^{\rho}} \right) - \frac{\partial}{\partial x^{\alpha}} \left(\cancel{\frac{\partial h_{\mu\rho}}{\partial x^{\nu}}} + \frac{\partial h_{\nu\rho}}{\partial x^{\mu}} - \frac{\partial h_{\mu\nu}}{\partial x^{\rho}} \right) \right] + O(h^2) \\ &= \frac{\eta^{\alpha\rho}}{2} \left[\frac{\partial^2 h_{\rho\alpha}}{\partial x^{\mu} \partial x^{\nu}} + \frac{\partial^2 h_{\mu\nu}}{\partial x^{\rho} \partial x^{\alpha}} - \frac{\partial^2 h_{\mu\alpha}}{\partial x^{\rho} \partial x^{\nu}} - \frac{\partial^2 h_{\nu\rho}}{\partial x^{\mu} \partial x^{\alpha}} \right] + O(h^2). \end{aligned} \tag{7.5}$$

To fix the ambiguity of choosing a coordinate system ("gauge"), we impose the harmonic coordinate condition

$$\begin{aligned} g^{\alpha\beta} \Gamma^{\mu}{}_{\alpha\beta} &= 0 \\ (\eta^{\alpha\beta} + h^{\alpha\beta}) \frac{1}{2} (\eta^{\mu\rho} + h^{\mu\rho}) \left[\frac{\partial h_{\alpha\rho}}{\partial x^{\beta}} + \frac{\partial h_{\beta\rho}}{\partial x^{\alpha}} - \frac{\partial h_{\alpha\beta}}{\partial x^{\rho}} \right] &= 0 \\ \eta^{\mu\rho} \eta^{\alpha\beta} \left[2 \frac{\partial h_{\alpha\rho}}{\partial x^{\beta}} - \frac{\partial h_{\alpha\beta}}{\partial x^{\rho}} \right] &= O(h^2) \\ \eta^{\alpha\beta} \frac{\partial h_{\alpha\rho}}{\partial x^{\beta}} &= \frac{1}{2} \frac{\partial h_{\alpha\beta}}{\partial x^{\rho}} \eta^{\alpha\beta} + O(h^2). \end{aligned} \tag{7.6}$$

Thus,

$$\eta^{\alpha\beta}\frac{\partial^2 h_{\alpha\mu}}{\partial x^\nu \partial x^\beta} = \frac{1}{2}\frac{\partial^2 h_{\alpha\beta}}{\partial x^\mu \partial x^\nu}\eta^{\alpha\beta}, \tag{7.7}$$

and

$$\eta^{\alpha\beta}\frac{\partial^2 h_{\beta\nu}}{\partial x^\mu \partial x^\alpha} = \frac{1}{2}\frac{\partial^2 h_{\alpha\beta}}{\partial x^\mu \partial x^\nu}\eta^{\alpha\beta}. \tag{7.8}$$

With these two relations the expression for the linearised Ricci tensor simplifies to

$$R_{\mu\nu} = \frac{1}{2}\eta^{\alpha\beta}\frac{\partial^2 h_{\mu\nu}}{\partial x^\alpha \partial x^\beta} = \frac{1}{2}\Box h_{\mu\nu}. \tag{7.9}$$

For a system of harmonic coordinates, Einstein's linearised equations become

$$\Box h_{\mu\nu} = -16\pi G S_{\mu\nu}, \tag{7.10}$$

$$\frac{\partial h^\alpha{}_\mu}{\partial x^\alpha} = \frac{1}{2}\frac{\partial h^\alpha{}_\alpha}{\partial x^\mu}. \tag{7.11}$$

7.1 Solutions of the linearised equations

Equation (7.10) has the same form as the equation for the four-vector potential of electrodynamics in the Lorenz gauge. We can therefore use the retarded Green's function

$$G_{\text{ret}}(x-x') = \frac{1}{2\pi}\,\delta\left((x-x')^2\right)\Theta(x^0 > x'^0) = \frac{\delta\left(x^0 - x'^0 - |\vec{x}-\vec{x}\,'|\right)}{4\pi|\vec{x}-\vec{x}\,'|}, \tag{7.12}$$

which satisfies

$$\Box_x G_{\text{ret}}(x-x') = \delta(x-x'). \tag{7.13}$$

Then, the solution is

$$h_{\mu\nu}(x) = -16\pi G\int \mathrm{d}^4x'\, G_{\text{ret}}(x-x')S_{\mu\nu}(x'), \tag{7.14}$$

as

$$\Box_x h_{\mu\nu} = -16\pi G\int \mathrm{d}^4x'\,\Big(\underbrace{\Box_x G_{\text{ret}}(x-x')}_{=\delta(x-x')}\Big)S_{\mu\nu}(x') = -16\pi G S_{\mu\nu}(x). \tag{7.15}$$

Performing the x'^0 integration in (7.14), we have

$$\begin{aligned} h_{\mu\nu}(x) &= -8G\int \mathrm{d}^3x'\, S_{\mu\nu}(\vec{x}\,', t')\mathrm{d}t'\frac{\delta(t' - (t - |\vec{x}-\vec{x}\,'|))}{2|\vec{x}-\vec{x}\,'|} \\ h_{\mu\nu}(\vec{x}, t) &= -4G\int \mathrm{d}^3x'\,\frac{S_{\mu\nu}(\vec{x}\,', t-|\vec{x}-\vec{x}\,'|)}{|\vec{x}-\vec{x}\,'|}. \end{aligned} \tag{7.16}$$

Remark. The overall sign is fixed by the Newtonian limit: for a static source $S_{00} = \rho/2$ and $h_{00} = -2G\int \mathrm{d}^3x'\,\rho/|\vec{x}-\vec{x}'| = 2\phi$, i.e. $g_{00} = 1 + 2\phi$ in signature $(+,-,-,-)$. Weinberg's $h_{\mu\nu} = +4G\int S_{\mu\nu}/|\vec{x}-\vec{x}'|$ (Weinberg, Sec. 10.1) holds for his signature $(-,+,+,+)$, where $h_{\mu\nu}$ has the opposite sign. Quantities quadratic in h (energy flux, radiated power) are

unaffected. The retarded solution above satisfies the harmonic coordinate condition of (7.11). Indeed, we have

$$T^{\mu\nu}{}_{;\mu} = 0 \;\Rightarrow\; \frac{\partial T^{\mu\nu}}{\partial x^\mu} + \underbrace{\Gamma T}_{\text{non-linear}} = 0 \;\Rightarrow\; \frac{\partial T^{\mu\nu}}{\partial x^\mu} = 0, \tag{7.17}$$

ignoring non-linearities. Then,

$$\frac{\partial}{\partial x^\mu} S^{\mu\nu} = \partial_\mu \left(T^{\mu\nu} - \frac{\eta^{\mu\nu}}{2} T^\alpha{}_\alpha \right) = -\frac{\eta^{\mu\nu}}{2} \frac{\partial}{\partial x^\mu} T^\alpha{}_\alpha. \tag{7.18}$$

Also,

$$S^\mu{}_\mu = T^\mu{}_\mu - \frac{\delta^\mu{}_\mu}{2} T^\alpha{}_\alpha = -T^\alpha{}_\alpha. \tag{7.19}$$

Thus,

$$\begin{aligned} \frac{\partial S^{\mu\nu}}{\partial x^\mu} &= \frac{1}{2} \frac{\partial S^\alpha{}_\alpha}{\partial x^\mu} \eta^{\mu\nu} \\ \frac{\partial S^\mu{}_\nu}{\partial x^\mu} &= \frac{1}{2} \frac{\partial S^\alpha{}_\alpha}{\partial x^\nu}. \end{aligned} \tag{7.20}$$

Then,

$$\begin{aligned} \frac{\partial h^\mu{}_\nu}{\partial x^\mu} &= -16\pi G \frac{\partial}{\partial x^\mu} \int \mathrm{d}^4x'\, G_{\text{ret}}(x - x') S^\mu{}_\nu(x') \\ &= -16\pi G \int \mathrm{d}^4x'\, \frac{\partial G_{\text{ret}}(x - x')}{\partial x^\mu} S^\mu{}_\nu(x') \\ &= +16\pi G \int \mathrm{d}^4x' \left(\frac{\partial}{\partial x'^\mu} G_{\text{ret}}(x - x') \right) S^\mu{}_\nu(x') \\ &= \underbrace{-16\pi G\, G_{\text{ret}}(x - x') S^\mu{}_\nu(x')}_{=0 \text{ for } x = x'} + - 16\pi G \int \mathrm{d}^4x'\, G_{\text{ret}} \frac{\partial S^\mu{}_\nu(x')}{\partial x'^\mu} \\ &= -16\pi G \int \mathrm{d}^4x'\, G_{\text{ret}}(x - x') \frac{1}{2} \frac{\partial S^\mu{}_\mu(x')}{\partial x'^\nu} \\ &= \cdots \\ &= \frac{\partial}{\partial x^\nu} \left[-16\pi G \int \mathrm{d}^4x'\, G_{\text{ret}}(x - x') \frac{S^\mu{}_\mu(x')}{2} \right] \\ &= \frac{1}{2} \frac{\partial h^\mu{}_\mu(x)}{\partial x^\nu}. \checkmark \end{aligned} \tag{7.21}$$

The general solution is the sum of the solution (7.16) and a solution of the homogeneous equations

$$\begin{aligned} \Box h_{\mu\nu} &= 0 \\ \partial_\mu h^\mu{}_\nu &= \frac{1}{2} \partial_\nu h^\mu{}_\mu. \end{aligned} \tag{7.22}$$

These can be solved with plane waves. Since the metric $h_{\mu\nu}$ must be real, we seek solutions of the form

$$h_{\mu\nu} = \varepsilon_{\mu\nu} e^{ik\cdot x} + \varepsilon^*_{\mu\nu} e^{-ik\cdot x}, \tag{7.23}$$

where

$$k \cdot x \equiv \eta_{\mu\nu} k^\mu x^\nu = k_\mu x^\mu. \tag{7.24}$$

Substituting into $\Box h_{\mu\nu} = 0$ gives $k_\mu k^\mu \equiv k^2 = 0$. From

$$\frac{\partial h^\mu{}_\nu}{\partial x^\mu} = \frac{1}{2} \frac{\partial h^\mu{}_\mu}{\partial x^\nu}, \tag{7.25}$$

we have

$$\varepsilon^\mu{}_\nu k_\mu = \frac{1}{2} \varepsilon^\mu{}_\mu k_\nu. \tag{7.26}$$

Let us now count the independent components of the tensor $\varepsilon_{\mu\nu}$, which is known as the polarisation tensor.

- Due to the symmetry of $h_{\mu\nu}$, the polarisation tensor $\varepsilon_{\mu\nu}$ is also symmetric.
- The ten components of $\varepsilon_{\mu\nu}$ are related via the four equations of (7.26). Thus $\varepsilon_{\mu\nu}$ has at most six independent components.

Let us now investigate further the freedom that we have to perform coordinate transformations. Under transformations

$$x^\mu \to x'^\mu = x^\mu + \omega^\mu(x), \tag{7.27}$$

the metric transforms as

$$g_{\mu\nu} \to g'_{\mu\nu}(x') = \frac{\partial x^\rho}{\partial x'^\mu} \frac{\partial x^\sigma}{\partial x'^\nu} g_{\rho\sigma}(x). \tag{7.28}$$

In the primed coordinate system, the metric should still deviate only slightly from the Minkowski metric. Thus, we only consider coordinate transformations with $\omega^\mu \sim h^{\mu\nu}$ small. From

$$\begin{aligned}
\frac{\partial x'^\mu}{\partial x^\rho} \frac{\partial x^\rho}{\partial x'^\nu} &= \delta^\mu{}_\nu \\
\to \left(\delta^\mu{}_\rho + \frac{\partial \omega^\mu}{\partial x^\rho} \right) \frac{\partial x^\rho}{\partial x'^\nu} &= \delta^\mu{}_\nu \\
\Rightarrow \left(\delta^\alpha{}_\mu - \frac{\partial \omega^\alpha}{\partial x^\mu} \right) \left(\delta^\mu{}_\rho + \frac{\partial \omega^\mu}{\partial x^\rho} \right) \frac{\partial x^\rho}{\partial x'^\nu} &= \delta^\mu{}_\nu \left(\delta^\alpha{}_\mu - \frac{\partial \omega^\alpha}{\partial x^\mu} \right) \\
\left(\delta^\alpha{}_\rho - \frac{\partial \omega^\alpha}{\partial x^\rho} + \frac{\partial \omega^\alpha}{\partial x^\rho} + \mathcal{O}(\omega^2) \right) \frac{\partial x^\rho}{\partial x'^\nu} &= \delta^\alpha{}_\nu - \frac{\partial \omega^\alpha}{\partial x^\nu} \\
\Rightarrow \frac{\partial x^\alpha}{\partial x'^\nu} &\approx \delta^\alpha{}_\nu - \frac{\partial \omega^\alpha}{\partial x^\nu}.
\end{aligned} \tag{7.29}$$

Then,

$$\begin{aligned}
g'_{\mu\nu}(x') &= \left(\delta^\alpha{}_\nu - \frac{\partial \omega^\alpha}{\partial x^\nu} \right) \left(\delta^\beta{}_\mu - \frac{\partial \omega^\beta}{\partial x^\mu} \right) (\eta_{\alpha\beta} + h_{\alpha\beta}) \\
&\approx \eta_{\mu\nu} + h_{\mu\nu} - \frac{\partial \omega_\mu}{\partial x^\nu} - \frac{\partial \omega_\nu}{\partial x^\mu}.
\end{aligned} \tag{7.30}$$

We have already required the harmonic coordinate condition, $g^{\alpha\beta}\Gamma^\mu{}_{\alpha\beta} = 0$, on $h_{\mu\nu}$. A generic coordinate transformation could spoil that. However, in the linear approximation

it is still possible to find $x \to x + \omega$ with $\omega^\mu(x)$ such that the harmonic coordinate condition is satisfied in the new frame too. Indeed, we require

$$\begin{aligned} & g'_{\alpha\beta}\Gamma'^{\mu}{}_{\alpha\beta} = 0 \\ \Rightarrow\ & \frac{\partial h'^{\mu}{}_{\nu}}{\partial x^\mu} = \frac{1}{2}\frac{\partial h'^{\mu}{}_{\mu}}{\partial x^\nu} \\ \Rightarrow\ & \frac{\partial}{\partial x^\mu}\left(h^{\mu}{}_{\nu} - \frac{\partial \omega^\mu}{\partial x^\nu} - \frac{\partial \omega_\nu}{\partial x_\mu}\right) \equiv \frac{1}{2}\frac{\partial}{\partial x^\nu}\left(h^{\mu}{}_{\mu} - \frac{\partial \omega^\mu}{\partial x^\mu} - \frac{\partial \omega_\mu}{\partial x_\mu}\right) \\ & -\Box\omega_\nu(x) - \frac{\partial^2 \omega^\mu(x)}{\partial x^\mu \partial x^\nu} = -\frac{1}{2}\frac{\partial^2 \omega^\mu(x)}{\partial x^\mu \partial x^\nu} - \frac{1}{2}\frac{\partial^2 \omega^\mu(x)}{\partial x^\mu \partial x^\nu} \\ & \Rightarrow\ \Box\omega_\nu(x) = 0. \end{aligned} \tag{7.31}$$

Thus, the harmonic coordinate condition does not fix the coordinate system completely in the linearised theory. For example, we can choose

$$\omega_\mu(x) = \mathrm{i}\varepsilon_\mu e^{+\mathrm{i}k\cdot x} - \mathrm{i}\varepsilon^*_\mu e^{-\mathrm{i}k\cdot x}. \tag{7.32}$$

Then $h_{\mu\nu} = \varepsilon_{\mu\nu}e^{\mathrm{i}k\cdot x} + \varepsilon^*{}_{\mu\nu}e^{-\mathrm{i}k\cdot x}$ becomes

$$h'_{\mu\nu} = \varepsilon'_{\mu\nu}e^{\mathrm{i}k\cdot x} + \varepsilon'^*{}_{\mu\nu}e^{-\mathrm{i}k\cdot x}, \tag{7.33}$$

with

$$\varepsilon'_{\mu\nu} = \varepsilon_{\mu\nu} + k_\mu\varepsilon_\nu + k_\nu\varepsilon_\mu. \tag{7.34}$$

Both $h_{\mu\nu}$ and $h'_{\mu\nu}$ are physically equivalent. However, the coordinate transformation introduces four arbitrary parameters, the components of ε_μ. These parameters can only affect combinations of the components of the polarisation tensor which have no physical significance. In conclusion, only two of the six components of $\varepsilon_{\mu\nu}$ are physical.

To illustrate this, let us choose the wave to propagate along the positive z-axis, $k^\mu = k(1,0,0,1)$. Then

$$k_\mu\varepsilon^{\mu}{}_{\nu} = \frac{1}{2}k_\nu\varepsilon^{\mu}{}_{\mu}, \tag{7.35}$$

or, equivalently,

$$k^\mu\varepsilon_{\mu\nu} = \frac{1}{2}\varepsilon_{\mu\rho}\eta^{\rho\mu}\eta_{\nu\alpha}k^\alpha. \tag{7.36}$$

This yields

$$\begin{aligned} \varepsilon_{01} &= -\varepsilon_{13} \\ \varepsilon_{02} &= -\varepsilon_{23} \\ \varepsilon_{03} &= -\frac{1}{2}(\varepsilon_{00} + \varepsilon_{33}) \\ \varepsilon_{22} &= -\varepsilon_{11}, \end{aligned} \tag{7.37}$$

which indeed reduces the number of independent components of $\varepsilon_{\mu\nu}$ to six. Let us now perform the change of coordinates described above giving $h_{\mu\nu} \to h'_{\mu\nu}$, where both satisfy the harmonic coordinate condition. Explicitly, the polarisation tensor changes as $\varepsilon_{\mu\nu} \to \varepsilon'_{\mu\nu} = \varepsilon_{\mu\nu} + k_\mu\varepsilon_\nu + k_\nu\varepsilon_\mu$. The components

$$\begin{aligned} \varepsilon_{11} &\to \varepsilon'_{11} = \varepsilon_{11} \\ \varepsilon_{12} &\to \varepsilon'_{12} = \varepsilon_{12}, \end{aligned} \tag{7.38}$$

remain unchanged, while

$$\begin{aligned} \varepsilon_{13} &\to \varepsilon'_{13} = \varepsilon_{13} - k\varepsilon_1 \\ \varepsilon_{23} &\to \varepsilon'_{23} = \varepsilon_{23} - k\varepsilon_2 \\ \varepsilon_{00} &\to \varepsilon'_{00} = \varepsilon_{00} + 2k\varepsilon_0 \\ \varepsilon_{33} &\to \varepsilon'_{33} = \varepsilon_{33} - 2k\varepsilon_3. \end{aligned} \tag{7.39}$$

We can choose

$$\varepsilon_1 = \frac{\varepsilon_{13}}{k},\ \varepsilon_2 = \frac{\varepsilon_{23}}{k},\ \varepsilon_3 = \frac{\varepsilon_{33}}{2k},\ \varepsilon_0 = -\frac{\varepsilon_{00}}{2k}. \tag{7.40}$$

In these coordinates the wave has only two degrees of freedom,

$$\varepsilon'_{11} = \varepsilon_{11}, \quad \varepsilon'_{12} = \varepsilon_{12}, \tag{7.41}$$

while all other independent components of the polarisation tensor are in this way made to vanish.

Let us now perform a rotation by an angle θ about the z-axis (the direction of $\vec{k} = k\hat{z}$). The polarisation tensor then changes as

$$\varepsilon_{\mu\nu} \to \varepsilon'_{\mu\nu} = R_\mu{}^\rho R_\nu{}^\sigma \varepsilon_{\rho\sigma}, \tag{7.42}$$

with

$$\begin{aligned} R_1{}^1 &= \cos\theta \qquad R_1{}^2 = \sin\theta \\ R_2{}^1 &= -\sin\theta \quad R_2{}^2 = \cos\theta \\ R_3{}^3 &= R_0{}^0 = 1 \quad \text{and all other } R_\mu{}^\nu = 0. \end{aligned} \tag{7.43}$$

Defining

$$\begin{aligned} e_\pm &= \varepsilon_{11} \mp \mathrm{i}\varepsilon_{12} \\ f_\pm &= \varepsilon_{31} \mp \mathrm{i}\varepsilon_{32}, \end{aligned} \tag{7.44}$$

we have that

$$\begin{aligned} e'_\pm &= e^{\pm 2\mathrm{i}\theta} e_\pm \qquad \to \text{“helicity } \pm 2\text{”} \\ f'_\pm &= e^{\pm \mathrm{i}\theta} f_\pm \qquad \to \text{“helicity } \pm 1\text{”} \\ \varepsilon'_{33} &= \varepsilon_{33},\ \varepsilon'_{00} = \varepsilon_{00} \quad \to \text{“scalar”}. \end{aligned} \tag{7.45}$$

7.2 Analogy with the electromagnetic field

Recall the corresponding relations for an electromagnetic plane wave in the Lorenz gauge,

$$A_\mu(x) = \varepsilon_\mu e^{+\mathrm{i}k\cdot x} + \varepsilon^*_\mu e^{-\mathrm{i}k\cdot x}, \tag{7.46}$$

$$\left.\begin{aligned} \Box A_\mu &= 0 \\ \partial_\mu A^\mu &= 0 \end{aligned}\right\} \to \begin{aligned} k^2 &= 0 \\ k_\mu \varepsilon^\mu &= 0, \end{aligned} \tag{7.47}$$

$$\begin{aligned} A_\mu &\to A_\mu + \partial_\mu \phi, \\ \phi(x) &= \mathrm{i}\varepsilon e^{+\mathrm{i}k\cdot x} - \mathrm{i}\varepsilon^* e^{-\mathrm{i}k\cdot x}. \end{aligned} \tag{7.48}$$

Then,

$$A_\mu(x) \to A'_\mu(x) = \varepsilon'_\mu e^{ik\cdot x} + \varepsilon'^*_\mu e^{-ik\cdot x}, \tag{7.49}$$

with

$$\varepsilon'_\mu = \varepsilon_\mu - \varepsilon k_\mu. \tag{7.50}$$

For $k^\mu = k(1,0,0,1)$,

$$\varepsilon_\mu k^\mu = 0 \;\Rightarrow\; \varepsilon^0 - \varepsilon^3 = 0 \;\Rightarrow\; \varepsilon_0 = -\varepsilon_3. \tag{7.51}$$

Also, by

$$\varepsilon_\mu \to \varepsilon'_\mu = \varepsilon_\mu - \varepsilon k_\mu, \tag{7.52}$$

we get

$$\begin{aligned} \varepsilon'_0 &= \varepsilon_0 - \varepsilon k \\ \varepsilon'_3 &= \varepsilon_3 + \varepsilon k = -\varepsilon_0 + \varepsilon k \\ \varepsilon'_1 &= \varepsilon_1 \\ \varepsilon'_2 &= \varepsilon_2. \end{aligned} \tag{7.53}$$

Setting $\varepsilon = \frac{\varepsilon_0}{k}$ we get

$$\varepsilon'_0 = 0,\ \varepsilon'_3 = 0, \tag{7.54}$$

and only ε_1 and ε_2 are physical.

Under a rotation,

$$\varepsilon'_\mu = R_\mu{}^\nu(\theta)\varepsilon_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix} = \begin{pmatrix} \varepsilon_0 \\ \varepsilon_1\cos\theta + \varepsilon_2\sin\theta \\ -\varepsilon_1\sin\theta + \varepsilon_2\cos\theta \\ \varepsilon_3 \end{pmatrix}. \tag{7.55}$$

Thus,

$$\left.\begin{aligned} \varepsilon'_0 &= \varepsilon_0 \\ \varepsilon'_3 &= \varepsilon_3 \end{aligned}\right\} \to \text{ scalar}, \tag{7.56}$$

and with

$$\varepsilon_\pm = \frac{\varepsilon_1 \mp i\varepsilon_2}{2},\ \varepsilon_+ + \varepsilon_- = \varepsilon_1,\ \varepsilon_+ - \varepsilon_- = -i\varepsilon_2, \tag{7.57}$$

we have

$$\varepsilon'_\pm = \frac{\varepsilon'_1 \mp i\varepsilon'_2}{2} = e^{\pm i\theta}\varepsilon_\pm, \tag{7.58}$$

the physical degrees of freedom with helicity ± 1.

7.3 Energy and momentum of plane waves

Recall that we can rewrite Einstein's equations

$$R_{\mu\nu} - \frac{g_{\mu\nu}}{2}R = -8\pi G T_{\mu\nu}, \tag{7.59}$$

as

$$R^{(1)}_{\mu\nu} - \frac{\eta_{\mu\nu}}{2}R^{(1)} = -8\pi G t_{\mu\nu}, \tag{7.60}$$

where $R^{(1)}{}_{\mu\nu}$ is the linear part of the Ricci tensor, $R^{(1)} \equiv \eta^{\mu\nu} R^{(1)}{}_{\mu\nu}$ is the linear part of the curvature scalar, while

$$-8\pi G t_{\mu\nu} = -8\pi G T_{\mu\nu} - \left(R_{\mu\nu} - \frac{g_{\mu\nu}}{2} R\right) + R^{(1)}{}_{\mu\nu} - \frac{\eta_{\mu\nu}}{2} R^{(1)}. \tag{7.61}$$

We have shown that $\partial_\mu \left(R^{(1)}{}_{\mu\nu} - \frac{\eta_{\mu\nu}}{2} R^{(1)}\right) = 0$, which implies that

$$\partial_\mu t^{\mu\nu} = 0, \tag{7.62}$$

and we have thus identified $t^{\mu\nu}$ with the energy-momentum tensor of matter and gravitation. For plane waves, $T_{\mu\nu} = 0$ (empty space), and we can therefore write

$$-8\pi G t_{\mu\nu} = -R_{\mu\nu} + \frac{g_{\mu\nu}}{2} R + R^{(1)}{}_{\mu\nu} - \frac{\eta_{\mu\nu}}{2} R^{(1)}. \tag{7.63}$$

We can decompose the Ricci tensor in powers of $h_{\mu\nu}$, where $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$. We write

$$R_{\mu\nu} = R^{(1)}{}_{\mu\nu} + R^{(2)}{}_{\mu\nu} + \cdots. \tag{7.64}$$

Then,

$$\begin{aligned} 8\pi G t_{\mu\nu} &= \left(R^{(1)}{}_{\mu\nu} + R^{(2)}{}_{\mu\nu} + \cdots\right) - \frac{\eta_{\mu\nu} + h_{\mu\nu}}{2} \left(\eta^{\alpha\beta} - h^{\alpha\beta}\right) \left(R^{(1)}{}_{\alpha\beta} + R^{(2)}{}_{\alpha\beta} + \cdots\right) \\ &\quad - R^{(1)}{}_{\mu\nu} + \frac{\eta_{\mu\nu}}{2} \eta^{\alpha\beta} R^{(1)}{}_{\alpha\beta} \\ &\approx R^{(2)}{}_{\mu\nu} - \frac{\eta_{\mu\nu}\eta^{\alpha\beta}}{2} R^{(2)}{}_{\alpha\beta} - \frac{h_{\mu\nu}\eta^{\alpha\beta} - \eta_{\mu\nu} h^{\alpha\beta}}{2} R^{(1)}{}_{\alpha\beta} + O(h^3). \end{aligned} \tag{7.65}$$

Plane-wave solutions satisfy the linearised Einstein equations,

$$R^{(1)}{}_{\mu\nu} = 0. \tag{7.66}$$

Therefore, the energy-momentum tensor of a gravitational wave is of second order in the weak-field approximation and reads

$$\begin{aligned} t_{\mu\nu} &\approx \frac{1}{8\pi G} \left[R^{(2)}{}_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} \eta^{\alpha\beta} R^{(2)}{}_{\alpha\beta}\right] \\ &= \frac{1}{8\pi G} \left[\delta^\alpha{}_\mu \delta^\beta{}_\nu - \frac{1}{2} \eta_{\mu\nu} \eta^{\alpha\beta}\right] R^{(2)}{}_{\alpha\beta}. \end{aligned} \tag{7.67}$$

An explicit calculation (Exercise) gives

$$\begin{aligned} R^{(2)}{}_{\mu\kappa} &= -\frac{1}{2} h^{\lambda\nu} \left[\frac{\partial^2 h_{\lambda\nu}}{\partial x^\kappa \partial x^\mu} - \frac{\partial^2 h_{\mu\nu}}{\partial x^\kappa \partial x^\lambda} \right. \\ &\quad \left. - \frac{\partial^2 h_{\lambda\kappa}}{\partial x^\nu \partial x^\mu} + \frac{\partial^2 h_{\mu\kappa}}{\partial x^\nu \partial x^\lambda}\right] \\ &\quad + \frac{1}{4} \left[2\frac{\partial h^\nu{}_\sigma}{\partial x^\nu} - \frac{\partial h^\nu{}_\nu}{\partial x^\sigma}\right] \left[\frac{\partial h^\sigma{}_\mu}{\partial x^\kappa} + \frac{\partial h^\sigma{}_\kappa}{\partial x^\mu} - \frac{\partial h_{\mu\kappa}}{\partial x^\sigma}\right] \\ &\quad - \frac{1}{4} \left[\frac{\partial h_{\sigma\kappa}}{\partial x^\lambda} + \frac{\partial h_{\sigma\lambda}}{\partial x^\kappa} - \frac{\partial h_{\lambda\kappa}}{\partial x^\sigma}\right] \left[\frac{\partial h^\sigma{}_\mu}{\partial x_\lambda} + \frac{\partial h^{\sigma\lambda}}{\partial x^\mu} - \frac{\partial h^\lambda{}_\mu}{\partial x_\sigma}\right]. \end{aligned} \tag{7.68}$$

In the above we raise and lower indices with the Minkowski part of the metric only. The first term does not vanish: it is a contraction of h with the linearised Riemann tensor, not with $R^{(1)}_{\lambda\nu}$. For the plane wave below its period average is $k_\mu k_\kappa\left(\varepsilon^{*\lambda\rho}\varepsilon_{\lambda\rho} - \frac{1}{2}|\varepsilon^\lambda{}_\lambda|^2\right)$, while the last line averages to $-\frac{1}{2}k_\mu k_\kappa\left(\varepsilon^{*\lambda\rho}\varepsilon_{\lambda\rho} - \frac{1}{2}|\varepsilon^\lambda{}_\lambda|^2\right)$; both must be kept. The second line vanishes due to the harmonic coordinate condition

$$\frac{1}{2}\frac{\partial h^\nu{}_\nu}{\partial x^\sigma} = \frac{\partial h^\nu{}_\sigma}{\partial x^\nu}. \tag{7.69}$$

The first and the last lines need to be computed explicitly, substituting the wave solution

$$h_{\mu\nu} = \varepsilon_{\mu\nu}e^{-\mathrm{i}k\cdot x} + \varepsilon^*{}_{\mu\nu}e^{+\mathrm{i}k\cdot x}. \tag{7.70}$$

We will need products of the type

$$\begin{aligned}\frac{\partial h_{\alpha\beta}}{\partial x^\gamma}\frac{\partial h_{\mu\nu}}{\partial x^\rho} &= \mathrm{i}k_\gamma\left(-\varepsilon_{\alpha\beta}e^{-\mathrm{i}k\cdot x} + \varepsilon^*{}_{\alpha\beta}e^{+\mathrm{i}k\cdot x}\right)\mathrm{i}k_\rho\left(-\varepsilon_{\mu\nu}e^{-\mathrm{i}k\cdot x} + \varepsilon^*{}_{\mu\nu}e^{+\mathrm{i}k\cdot x}\right)\\ &= -k_\gamma k_\rho\left[\varepsilon_{\alpha\beta}\varepsilon_{\mu\nu}e^{-2\mathrm{i}k\cdot x} + \varepsilon^*{}_{\alpha\beta}\varepsilon^*{}_{\mu\nu}e^{+2\mathrm{i}k\cdot x} - \varepsilon_{\alpha\beta}\varepsilon^*{}_{\mu\nu} - \varepsilon^*{}_{\alpha\beta}\varepsilon_{\mu\nu}\right].\end{aligned} \tag{7.71}$$

We can now average the energy and momentum over the period of the wave. The averaging will require integrals of the form

$$\int_0^T \mathrm{d}t\, e^{\pm 4\pi\mathrm{i}t/T} = 0. \tag{7.72}$$

Then, we are left with

$$\left\langle\frac{\partial h_{\alpha\beta}}{\partial x^\gamma}\frac{\partial h_{\mu\nu}}{\partial x^\rho}\right\rangle = k_\gamma k_\rho\left(\varepsilon_{\alpha\beta}\varepsilon^*{}_{\mu\nu} + \varepsilon^*{}_{\alpha\beta}\varepsilon_{\mu\nu}\right). \tag{7.73}$$

Taking the average then yields

$$\left\langle R^{(2)}{}_{\mu\nu}\right\rangle = \cdots = \frac{k_\mu k_\nu}{2}\left(\varepsilon^{*\lambda\rho}\varepsilon_{\lambda\rho} - \frac{1}{2}\left|\varepsilon^\lambda{}_\lambda\right|^2\right). \tag{7.74}$$

Then,

$$8\pi G\left\langle t_{\mu\nu}\right\rangle \approx \left\langle R^{(2)}{}_{\mu\nu}\right\rangle - \frac{\eta_{\mu\nu}}{2}\underbrace{\eta^{\alpha\beta}\left\langle R^{(2)}{}_{\alpha\beta}\right\rangle}_{\propto k^2=0} = \frac{k_\mu k_\nu}{2}\left(\varepsilon^{*\lambda\rho}\varepsilon_{\lambda\rho} - \frac{1}{2}\left|\varepsilon^\lambda{}_\lambda\right|^2\right). \tag{7.75}$$

Notice that $\left\langle t_{\mu\nu}\right\rangle$ is gauge invariant. Under

$$\varepsilon_{\mu\nu} \to \varepsilon'{}_{\mu\nu} = \varepsilon_{\mu\nu} + k_\mu\varepsilon_\nu + k_\nu\varepsilon_\mu, \tag{7.76}$$

we have

$$\varepsilon^{*\mu\nu}\varepsilon_{\mu\nu} \to \varepsilon^{*\mu\nu}\varepsilon_{\mu\nu} + 2k_\mu k_\nu\varepsilon^{*\mu}\varepsilon^\nu + 4k_\mu\mathrm{Re}[\varepsilon^*{}_\nu\varepsilon^{\mu\nu}] = \varepsilon^{*\mu\nu}\varepsilon_{\mu\nu} + 2\left|k\cdot\varepsilon\right|^2 + 2\mathrm{Re}[k_\nu\varepsilon^{*\nu}\varepsilon^\mu{}_\mu], \tag{7.77}$$

and

$$\begin{aligned}\varepsilon^\mu{}_\mu &\to \varepsilon^\mu{}_\mu + 2k_\mu\varepsilon^\mu\\ \left|\varepsilon^\mu{}_\mu\right|^2 &\to \left|\varepsilon^\mu{}_\mu\right|^2 + 4\mathrm{Re}[k_\nu\varepsilon^{*\nu}\varepsilon^\mu{}_\mu] + 4\left|k\cdot\varepsilon\right|^2,\end{aligned} \tag{7.78}$$

which leaves

$$\varepsilon^{*\mu\nu}\varepsilon_{\mu\nu} - \frac{1}{2}\left|\varepsilon^{\lambda}{}_{\lambda}\right|^2 \to \varepsilon^{*\mu\nu}\varepsilon_{\mu\nu} - \frac{1}{2}\left|\varepsilon^{\lambda}{}_{\lambda}\right|^2, \tag{7.79}$$

invariant. If we take $k^\mu = k(1,0,0,1)$ and choose the gauge so that $\varepsilon_{11}, \varepsilon_{12} \neq 0$ and all other $\varepsilon_{\mu\nu} = 0$, we have

$$\left\langle t_{\mu\nu} \right\rangle = \frac{k_\mu k_\nu}{8\pi G}\left[|\varepsilon_{11}|^2 + |\varepsilon_{12}|^2\right] = \frac{k_\mu k_\nu}{16\pi G}\left[|e_+|^2 + |e_-|^2\right], \tag{7.80}$$

with

$$e_\pm = \varepsilon_{11} \mp \mathrm{i}\varepsilon_{12}. \tag{7.81}$$

Exercise: Compare with the corresponding expression for an electromagnetic wave.

The continuity equation gives

$$\partial_\mu t^{\mu\nu} = 0 \;\Rightarrow\; \frac{\partial t^{0\nu}}{\partial t} + \frac{\partial t^{i\nu}}{\partial x^i} = 0. \tag{7.82}$$

For $\nu = 0$ we have

$$\begin{aligned} \frac{\partial}{\partial t}\left\langle t^{00}\right\rangle &= -\frac{\partial}{\partial x^i}\left\langle t^{0i}\right\rangle \\ \Rightarrow\; \frac{\partial}{\partial t}\int_V \mathrm{d}^3x\, t^{00} &= -\int_V \mathrm{d}^3x\, \frac{\partial}{\partial x^i} t^{0i} \\ \Rightarrow\; \frac{\partial E}{\partial t} &= -\int_{S(V)} \mathrm{d}S\, \hat{n}_i t^{0i}, \end{aligned} \tag{7.83}$$

where the right-hand side is the energy flux through the surface $S(V)$.

We will soon see that sources (mass distributions) confined to finite regions of space give rise to spherical gravitational waves. For a volume V whose surface $S(V)$ is far away from the source, this flux is the total radiated power,

$$P_{\mathrm{rad}} = -\frac{\partial E}{\partial t} = \int_{S(V)} \mathrm{d}S\, \hat{n}_i t^{0i}. \tag{7.84}$$

Remark. Here $\hat{n}_i$ are the Euclidean components of the outward unit normal (as in Gauss's theorem); on a sphere, $\hat{n}_i = \hat{x}^i$. Note that in electrodynamics the analogue of t^{00} is the energy density of an electromagnetic wave, and that of t^{0i} is the Poynting vector. For a sphere of radius r, $\mathrm{d}S = r^2\mathrm{d}\Omega$ and

$$\begin{aligned} P_{\mathrm{rad}} &= \int \mathrm{d}\Omega\, \frac{\mathrm{d}P_{\mathrm{rad}}}{\mathrm{d}\Omega} \\ &= \int \mathrm{d}\Omega\, r^2 \hat{x}^i \left\langle t^{0i}\right\rangle \\ \Rightarrow\; \frac{\mathrm{d}P_{\mathrm{rad}}}{\mathrm{d}\Omega} &= r^2 \hat{x}^i \left\langle t^{0i}\right\rangle. \end{aligned} \tag{7.85}$$

7.4 Generation of gravitational waves

For a system with energy-momentum tensor $T_{\mu\nu}$, the metric

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{7.86}$$

in the weak field approximation is given by

$$h_{\mu\nu}(\vec{x},t) = -4G \int \mathrm{d}^3x' \, \frac{S_{\mu\nu}(\vec{x}', t - |\vec{x} - \vec{x}'|)}{|\vec{x} - \vec{x}'|}, \tag{7.87}$$

where

$$S_{\mu\nu} = T_{\mu\nu} - \frac{\eta_{\mu\nu}}{2} T^{\alpha}{}_{\alpha}. \tag{7.88}$$

We restrict our study to sources $S_{\mu\nu}(\vec{x}', t')$ which are confined to a region $|\vec{x}'| \leq R$ and calculate the metric $h_{\mu\nu}(\vec{x}, t)$ at distances $|\vec{x}| \equiv r \gg R$. Then,

$$\begin{aligned} |\vec{x} - \vec{x}'| &= r \left(1 - \frac{2\hat{x} \cdot \vec{x}'}{r} + \frac{\vec{x}'^2}{r^2}\right)^{1/2} \\ |\vec{x} - \vec{x}'| &\approx r \left(1 - \frac{\hat{x} \cdot \vec{x}'}{r}\right) + O\left(\frac{R^2}{r}\right). \end{aligned} \tag{7.89}$$

We can analyse the source $S_{\mu\nu}(\vec{x}, t)$ in Fourier space,

$$S_{\mu\nu}(\vec{x}, t) = \int_0^\infty \mathrm{d}\omega \, e^{-\mathrm{i}\omega t} S_{\mu\nu}(\vec{x}, \omega) + \text{c.c.}, \tag{7.90}$$

where the c.c. stands for the complex conjugate. Then, the metric becomes

$$\begin{aligned} h_{\mu\nu}(\vec{x}, t) &= -4G \int_0^\infty \mathrm{d}\omega \int \mathrm{d}^3x' \, \frac{S_{\mu\nu}(\vec{x}', \omega) e^{-\mathrm{i}\omega(t - |\vec{x} - \vec{x}'|)}}{|\vec{x} - \vec{x}'|} + \text{c.c.} \\ &= -4G \int_0^\infty \mathrm{d}\omega \int \mathrm{d}^3x' \, \frac{S_{\mu\nu}(\vec{x}', \omega)}{r} e^{-\mathrm{i}\left[\omega(t - r + \vec{x}' \cdot \hat{x}) + O(\frac{\omega R^2}{r})\right]} + \text{c.c.} + O\left(\frac{R}{r}\right) \\ &\approx \int_0^\infty \mathrm{d}\omega \, e^{-\mathrm{i}\omega(t-r)} \left[-\frac{4G}{r} \int \mathrm{d}^3x' \, S_{\mu\nu}(\vec{x}', \omega) e^{-\mathrm{i}\omega \hat{x} \cdot \vec{x}'}\right] + \text{c.c.} \end{aligned} \tag{7.91}$$

In the above we have assumed that

(1) $\omega r \gg 1$, i.e. $r \gg \frac{1}{\omega}$ (long wavelengths do not dominate),

(2) $\frac{\omega R^2}{r}$ is small,

(3) $r \gg R$.

Combined, these conditions restrict our calculation to distances

$$r \gg \max\{R, \omega R^2, \frac{1}{\omega}\}. \tag{7.92}$$

Let us define the wave vector

$$\vec{k}(\omega) \equiv \omega \hat{x}. \tag{7.93}$$

Then,

$$\begin{aligned} h_{\mu\nu} &\approx \int_0^\infty \mathrm{d}\omega \, e^{-\mathrm{i}(\omega t - \vec{k} \cdot \vec{x})} \int \mathrm{d}^3x' \, S_{\mu\nu}(\vec{x}', \omega) e^{-\mathrm{i}\vec{x}' \cdot \vec{k}} \left(-\frac{4G}{r}\right) + \text{c.c.} \\ \Rightarrow \; h_{\mu\nu} &\approx \int_0^\infty \mathrm{d}\omega \, e^{-\mathrm{i}(\omega t - \vec{k} \cdot \vec{x})} \varepsilon_{\mu\nu}(\omega, \vec{k}) + \text{c.c.}, \end{aligned} \tag{7.94}$$

with the polarisation tensor

$$\varepsilon_{\mu\nu}(\omega,\vec{k}) \equiv -\frac{4G}{r} S_{\mu\nu}(\omega,\vec{k}), \tag{7.95}$$

and

$$S_{\mu\nu}(\omega,\vec{k}) = \int \mathrm{d}^3x'\, e^{-\mathrm{i}\vec{k}\cdot\vec{x}'} S_{\mu\nu}(\omega,\vec{x}'). \tag{7.96}$$

We have found that at large distances the metric takes the form of a superposition of (spherical, $\hat{k} = \hat{x}$) plane waves, with the polarisation tensor given above. We can now compute the energy-momentum tensor carried by the wave. We have already carried out this computation for a monochromatic wave. However, (7.94) is a superposition of frequencies, and the energy-momentum tensor will have additional interference terms of the form

$$\begin{aligned}\frac{\partial h_{\alpha\beta}}{\partial x^\gamma}\frac{\partial h_{\mu\nu}}{\partial x^\rho} \to \iint_0^\infty \mathrm{d}\omega \mathrm{d}\omega'\, k_\gamma k'_\rho \left[\varepsilon_{\alpha\beta}(k) e^{-\mathrm{i}(\omega t - \vec{k}\cdot\vec{x})} - \varepsilon^*{}_{\alpha\beta}(k) e^{+\mathrm{i}(\omega t - \vec{k}\cdot\vec{x})}\right] \times \\ \left[\varepsilon_{\mu\nu}(k') e^{-\mathrm{i}(\omega' t - \vec{k}'\cdot\vec{x})} - \varepsilon^*{}_{\mu\nu}(k') e^{+\mathrm{i}(\omega' t - \vec{k}'\cdot\vec{x})}\right].\end{aligned} \tag{7.97}$$

Integrating over time,

$$\int \mathrm{d}t\, \frac{\partial h_{\alpha\beta}}{\partial x^\gamma}\frac{\partial h_{\mu\nu}}{\partial x^\rho}, \tag{7.98}$$

gives rise to the time integrals

$$\begin{aligned}\int_{-\infty}^{+\infty} \mathrm{d}t\, e^{\pm\mathrm{i}(\omega+\omega')t} &= 2\pi\delta(\omega+\omega') = 0, \\ \int_{-\infty}^{+\infty} \mathrm{d}t\, e^{\pm\mathrm{i}(\omega-\omega')t} &= 2\pi\delta(\omega-\omega'),\end{aligned} \tag{7.99}$$

which eliminates all contributions from $\omega \neq \omega'$. Therefore, our result for the time-averaged energy-momentum tensor can be used here, without worrying about interference between different frequencies:

$$\int_{-\infty}^{+\infty} \mathrm{d}t\, t^{\mu\nu} = \int_0^\infty \mathrm{d}\omega\, \frac{k^\mu k^\nu}{8G}\left(\varepsilon^{*\alpha\beta}\varepsilon_{\alpha\beta} - \frac{1}{2}\left|\varepsilon^\alpha{}_\alpha\right|^2\right). \tag{7.100}$$

We can now calculate the radiated energy,

$$E = \int \mathrm{d}t\, P_{\mathrm{rad}} = \int \mathrm{d}t \int \mathrm{d}S\, \hat{n}_i t^{0i} = \int \mathrm{d}t \int \mathrm{d}\Omega\, r^2(\hat{x}^i t^{0i}). \tag{7.101}$$

Per unit solid angle, this reads

$$\begin{aligned}\frac{\mathrm{d}E}{\mathrm{d}\Omega} &= r^2 \int \mathrm{d}t\, (\hat{x}^i t^{0i}) = \cdots \\ &= 2G \int_0^\infty \mathrm{d}\omega\, \omega^2 \left[T^{*\mu\nu}(\vec{k},\omega) T_{\mu\nu}(\vec{k},\omega) - \frac{1}{2}\left|T^\mu{}_\mu(\vec{k},\omega)\right|^2\right],\end{aligned} \tag{7.102}$$

where

$$T_{\mu\nu}(\vec{k},\omega) \equiv \int \mathrm{d}^3x'\, e^{-\mathrm{i}\vec{k}\cdot\vec{x}'} T_{\mu\nu}(\vec{x}',\omega), \tag{7.103}$$

and

$$T_{\mu\nu}(\vec{x},t) = \int_0^\infty \mathrm{d}\omega \left[e^{-\mathrm{i}\omega t} T_{\mu\nu}(\vec{x},\omega) + e^{+\mathrm{i}\omega t} T^*_{\mu\nu}(\vec{x},\omega) \right]. \tag{7.104}$$

Application: radiation from a collision

Consider n particles moving with velocities $\vec{v}_n$ and carrying four-momenta $P_n^\mu \equiv (E_n, E_n\vec{v}_n)$ at times $t < 0$ before they collide at $t = 0$. After the collision the particles will have momenta $\bar{P}_n^\mu \equiv (\bar{E}_n, \bar{E}_n\bar{\vec{v}}_n)$. Due to momentum conservation we have

$$\sum_n (P_n - \bar{P}_n) = 0. \tag{7.105}$$

The energy-momentum tensor is

$$T^{\mu\nu}(\vec{x},t) = \sum_n \frac{P_n^\mu P_n^\nu}{E_n} \delta^{(3)}(\vec{x} - \vec{v}_n t)\Theta(-t) + \sum_n \frac{\bar{P}_n^\mu \bar{P}_n^\nu}{\bar{E}_n} \delta^{(3)}(\vec{x} - \bar{\vec{v}}_n t)\Theta(t). \tag{7.106}$$

We now use the integral representations of the Θ-function

$$\Theta(x) = \lim_{\varepsilon\to 0^+} \frac{1}{2\pi\mathrm{i}} \int_{-\infty}^{+\infty} \mathrm{d}\omega \, \frac{e^{+\mathrm{i}\omega x}}{\omega - \mathrm{i}\varepsilon}, \tag{7.107}$$

and of the δ-function

$$\delta^{(3)}(\vec{x}) = \int \frac{\mathrm{d}^3k}{(2\pi)^3} e^{+\mathrm{i}\vec{k}\cdot\vec{x}}. \tag{7.108}$$

Then,

$$\begin{aligned}
T^{\mu\nu}(\vec{x},t) &= \int \frac{\mathrm{d}\omega \mathrm{d}^3k}{(2\pi)^4 \mathrm{i}} \left[\sum_n \frac{P_n^\mu P_n^\nu}{E_n} e^{+\mathrm{i}\vec{k}\cdot(\vec{x}-\vec{v}_n t)} \frac{e^{+\mathrm{i}\omega(-t)}}{\omega - \mathrm{i}\varepsilon} \right. \\
&\qquad \left. - \sum_n \frac{\bar{P}_n^\mu \bar{P}_n^\nu}{\bar{E}_n} e^{+\mathrm{i}\vec{k}\cdot(\vec{x}-\bar{\vec{v}}_n t)} \frac{e^{-\mathrm{i}\omega t}}{\omega + \mathrm{i}\varepsilon} \right] \\
&= \int \frac{\mathrm{d}\omega \mathrm{d}^3k}{(2\pi)^4 \mathrm{i}} \left[\sum_n \frac{P_n^\mu P_n^\nu}{E_n} \frac{e^{-\mathrm{i}[(\omega+\vec{k}\cdot\vec{v}_n)t - \vec{k}\cdot\vec{x}]}}{\omega - \mathrm{i}\varepsilon} \right. \\
&\qquad \left. - \sum_n \frac{\bar{P}_n^\mu \bar{P}_n^\nu}{\bar{E}_n} \frac{e^{-\mathrm{i}[(\omega+\bar{\vec{v}}_n\cdot\vec{k})t - \vec{k}\cdot\vec{x}]}}{\omega + \mathrm{i}\varepsilon} \right] \\
&= \int \frac{\mathrm{d}^3k}{(2\pi)^4} e^{+\mathrm{i}\vec{k}\cdot\vec{x}} \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega}{\mathrm{i}} \left[\sum_n \frac{P_n^\mu P_n^\nu}{E_n} \frac{1}{\omega - \vec{k}\cdot\vec{v}_n - \mathrm{i}\varepsilon} \right. \\
&\qquad \left. - \sum_n \frac{\bar{P}_n^\mu \bar{P}_n^\nu}{\bar{E}_n} \frac{1}{\omega - \vec{k}\cdot\bar{\vec{v}}_n + \mathrm{i}\varepsilon} \right] e^{-\mathrm{i}\omega t}.
\end{aligned} \tag{7.109}$$

We can thus read off

$$T^{\mu\nu}(\vec{k},\omega) = \frac{1}{2\pi\mathrm{i}} \sum_n \frac{P_n^\mu P_n^\nu}{E_n(\omega - \vec{k}\cdot\vec{v}_n - \mathrm{i}\varepsilon)} - \frac{1}{2\pi\mathrm{i}} \sum_n \frac{\bar{P}_n^\mu \bar{P}_n^\nu}{\bar{E}_n(\omega - \vec{k}\cdot\bar{\vec{v}}_n + \mathrm{i}\varepsilon)}. \tag{7.110}$$

Then, indeed,

$$\begin{aligned} T^{\mu\nu}(\vec{x},t) &= \int \frac{\mathrm{d}^3k}{(2\pi)^3} e^{+\mathrm{i}\vec{k}\cdot\vec{x}} \int_{-\infty}^{+\infty} \mathrm{d}\omega\, e^{-\mathrm{i}\omega t} T^{\mu\nu}(\vec{k},\omega) \\ &= \int \frac{\mathrm{d}^3k}{(2\pi)^3} e^{+\mathrm{i}\vec{k}\cdot\vec{x}} \int_{0}^{+\infty} \mathrm{d}\omega \left[e^{-\mathrm{i}\omega t} T^{\mu\nu}(\vec{k},\omega) + e^{+\mathrm{i}\omega t} T^{\mu\nu}(\vec{k},-\omega) \right] \\ &= \int_0^\infty \mathrm{d}\omega \left[e^{-\mathrm{i}\omega t} \int \frac{\mathrm{d}^3k}{(2\pi)^3} T^{\mu\nu}(\vec{k},\omega) e^{+\mathrm{i}\vec{k}\cdot\vec{x}} + e^{+\mathrm{i}\omega t} \int \frac{\mathrm{d}^3k}{(2\pi)^3} T^{\mu\nu}(-\vec{k},-\omega) e^{-\mathrm{i}\vec{k}\cdot\vec{x}} \right] \\ &= \int_0^\infty \mathrm{d}\omega \left[e^{-\mathrm{i}\omega t} T^{\mu\nu}(\vec{x},\omega) + \text{c.c.} \right]. \end{aligned} \tag{7.111}$$

Notice also that

$$P_n \cdot k = E_n\omega - \vec{k}\cdot E_n\vec{v}_n = E_n(\omega - \vec{k}\cdot\vec{v}_n), \tag{7.112}$$

where we used $k^\mu = (\omega, \vec{k})$. Thus,

$$T^{\mu\nu}(\vec{k},\omega) = \frac{1}{2\pi\mathrm{i}} \sum_n \frac{P_n^\mu P_n^\nu}{P_n\cdot k - \mathrm{i}\varepsilon} - \frac{1}{2\pi\mathrm{i}} \sum_n \frac{\bar{P}_n^\mu \bar{P}_n^\nu}{\bar{P}_n\cdot k + \mathrm{i}\varepsilon}. \tag{7.113}$$

Moreover,

$$P_n\cdot k = E_n\omega(1 - \vec{v}_n\cdot\hat{x}) = E_n\omega(1 - v_n\cos\theta), \tag{7.114}$$

which never vanishes for massive particles. We can then drop the $\pm\mathrm{i}\varepsilon$ in the denominators and write the expression more compactly as

$$T^{\mu\nu}(\vec{k},\omega) = \frac{1}{2\pi\mathrm{i}} \sum_N \eta_N \frac{P_N^\mu P_N^\nu}{P_N\cdot k}, \tag{7.115}$$

where N runs over the particles of both the initial and the final state, and

$$\eta_N = \begin{cases} +1 & \text{initial state,} \\ -1 & \text{final state.} \end{cases} \tag{7.116}$$

Finally,

$$\begin{aligned} T^{*\mu\nu}T_{\mu\nu} &= \frac{1}{(2\pi)^2} \sum_{N,M} \eta_N\eta_M P_N^\mu P_N^\nu P_{M\mu} P_{M\nu} \frac{1}{(P_N\cdot k)(P_M\cdot k)} \\ &= \sum_{N,M} \frac{\eta_N\eta_M}{(2\pi)^2} \frac{(P_N\cdot P_M)^2}{(P_N\cdot k)(P_M\cdot k)}, \end{aligned} \tag{7.117}$$

and

$$\left|T^\mu{}_\mu\right|^2 = \frac{1}{(2\pi)^2} \sum_{N,M} \eta_N\eta_M \frac{P_N^2 P_M^2}{(P_N\cdot k)(P_M\cdot k)} = \frac{1}{(2\pi)^2} \sum_{N,M} \eta_N\eta_M \frac{m_N^2 m_M^2}{(P_N\cdot k)(P_M\cdot k)}. \tag{7.118}$$

So, the total radiated energy per solid angle is

$$\frac{\mathrm{d}E}{\mathrm{d}\Omega} = \int \mathrm{d}\omega\, \frac{G\omega^2}{2\pi^2} \sum_{N,M} \eta_N\eta_M \frac{(P_N\cdot P_M)^2 - \frac{1}{2}m_N^2 m_M^2}{(P_N\cdot k)(P_M\cdot k)}. \tag{7.119}$$

Observations:

(1) If $\bar{P}_n = P_n$ then $T_{\mu\nu}(\vec{k}, \omega) = 0$ and there is no radiation, $\frac{dE}{d\Omega} = 0$. To produce gravitational radiation, a particle needs to change its momentum, i.e. to be accelerated. This is also the case in electromagnetism, where radiation is produced by accelerated charges.

(2) For massless particles $v_N = 1$ and $P_N \cdot k = E_N \omega(1 - \cos\theta)$. For $\theta = 0$ this vanishes, which seems to create a singularity in our expression for the energy. This singularity, however, turns out to be spurious.

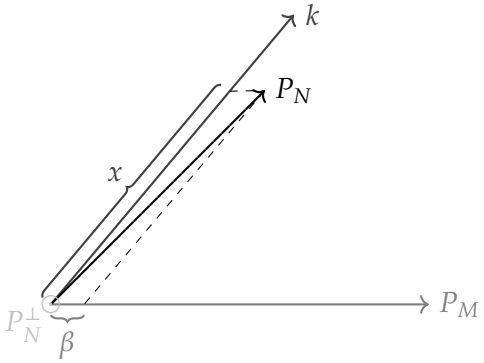


Figure 7.1. Decomposition of P_N when it is nearly parallel to k.

When P_N approaches the direction of k (Fig. 7.1), we can write

$$P_N = \underbrace{x}_{\text{large}} k + \underbrace{\beta}_{\text{small}} P_M + \underbrace{P_N^\perp}_{\text{small}}, \tag{7.120}$$

and $P_N \cdot k = x\cancelto{0}{k^2} + \beta P_M \cdot k + \cancelto{0}{P_N^\perp \cdot k}$.

Then

$$\frac{P_N \cdot P_M}{P_N \cdot k} \approx \frac{x\, P_M \cdot k}{\beta P_M \cdot k} = \frac{x}{\beta} \sim \text{large}. \tag{7.121}$$

However, in the sum this singularity vanishes. From $P_N \approx xk$ we get $P_N \cdot P_M \approx xP_M \cdot k \Rightarrow x = \frac{P_N \cdot P_M}{P_M \cdot k}$ and

$$\begin{aligned} \sum_{N,M} \eta_N \eta_M \frac{(P_N \cdot P_M)^2}{(P_N \cdot k)(P_M \cdot k)} &\approx \sum_{N,M} \eta_N \eta_M x \frac{P_N \cdot P_M}{P_N \cdot k} \\ &= \sum_N \frac{\eta_N x}{P_N \cdot k} \underbrace{\left(\sum_M \eta_M P_M \right)}_{=0} \cdot P_N . \end{aligned} \tag{7.122}$$

Due to momentum conservation the inner sum vanishes, and with it the singularity.

(3) In (7.119), by substituting $\vec{k} = \hat{x}\omega$, we find

$$\begin{aligned} \frac{dE}{d\Omega} &= \int_0^\infty d\omega \, \frac{G}{2\pi^2} \sum_{N,M} \eta_N \eta_M \frac{(P_N P_M)^2 - \frac{1}{2} m_N^2 m_M^2}{E_N E_M (1 - \vec{v}_N \cdot \hat{x})(1 - \vec{v}_M \cdot \hat{x})} \\ \Rightarrow \frac{dE}{d\Omega} &\propto \int_0^\infty d\omega, \end{aligned} \tag{7.123}$$

which is divergent. This divergence is removed if we take into account that the collision does not happen instantaneously at $t = 0$, but takes a finite time Δt to go from the initial to the final state. The integral then behaves as $\frac{dE}{d\Omega} \propto \int_0^{1/\Delta t} d\omega$.

The collision example describes a source which is anything but slowly moving. Most astrophysical sources of gravitational waves (binary stars, and binary black holes until

shortly before they merge) are instead non-relativistic: their velocities are small and their size R is small compared with the wavelength, $\omega R \ll 1$. In this case the formulae above simplify enormously. Before working this out, let us first understand what a gravitational wave actually does when it passes through a region of space.

7.5 Effect of a gravitational wave on test masses

Consider a plane wave travelling in the $\hat{z}$ direction, in the gauge in which only $\varepsilon_{11} = -\varepsilon_{22}$ and $\varepsilon_{12} = \varepsilon_{21}$ are non-zero (the transverse-traceless or TT gauge). It is convenient to use the mixed components $h^i{}_j = \eta^{ik}h_{kj} = -h_{ij}$, which in our signature are the ones with the "natural" sign, and to write them in terms of two real functions of $t - z$,

$$h^1{}_1 = -h^2{}_2 = h_+(t-z), \qquad h^1{}_2 = h^2{}_1 = h_\times(t-z), \tag{7.124}$$

all other components being zero. For a monochromatic wave $h_+ = -2\,\mathrm{Re}[\varepsilon_{11}e^{-\mathrm{i}\omega(t-z)}]$ and $h_\times = -2\,\mathrm{Re}[\varepsilon_{12}e^{-\mathrm{i}\omega(t-z)}]$. The proper time interval reads

$$\mathrm{d}\tau^2 = \mathrm{d}t^2 - \left(\delta_{ij} + h^i{}_j\right)\mathrm{d}x^i\mathrm{d}x^j = \mathrm{d}t^2 - (1+h_+)\mathrm{d}x^2 - (1-h_+)\mathrm{d}y^2 - 2h_\times \mathrm{d}x\mathrm{d}y - \mathrm{d}z^2. \tag{7.125}$$

Free particles stay at fixed TT coordinates. Take a particle at rest at $t = 0$, $u^\mu = \mathrm{d}x^\mu/\mathrm{d}\tau = (1,0,0,0)$. The geodesic equation, (2.12), gives at this instant

$$\frac{\mathrm{d}^2x^i}{\mathrm{d}\tau^2} = -\Gamma^i{}_{00} = -\frac{\eta^{ik}}{2}\left(2\frac{\partial h_{k0}}{\partial t} - \frac{\partial h_{00}}{\partial x^k}\right) = 0, \tag{7.126}$$

since $h_{00} = h_{0k} = 0$ in TT gauge. Hence the particle stays at rest, $u^\mu = (1,0,0,0)$ at all times, and since $g_{00} = 1$ its proper time is the coordinate time t. The TT coordinates are thus attached to freely falling test masses: the coordinate positions of free particles do not change when the wave passes. This does not mean that nothing happens. What changes is the proper distance between the particles. For two particles at $z = 0$ separated by the small coordinate vector ξ^i (with $\xi^3 = 0$) the proper distance is, from (7.125),

$$L^2 = \left(\delta_{ij} + h^i{}_j\right)\xi^i\xi^j \;\Rightarrow\; L \approx |\vec{\xi}|\left(1 + \frac{1}{2}h^i{}_j\hat{\xi}^i\hat{\xi}^j\right). \tag{7.127}$$

Equivalently, the physical displacement (the one measured with rulers in a local inertial frame of one of the particles) is

$$X^i = \xi^i + \frac{1}{2}h^i{}_j\xi^j, \qquad \frac{\mathrm{d}^2X^i}{\mathrm{d}t^2} = \frac{1}{2}\ddot{h}^i{}_jX^j + O(h^2). \tag{7.128}$$

The same result from the geodesic deviation equation. The statement above was made in a particular gauge; let us check it with a covariant tool. Consider two neighbouring geodesics $x^\mu(\tau)$ and $x^\mu(\tau) + \xi^\mu(\tau)$. Subtracting their geodesic equations to first order in ξ and working in a locally inertial frame at the point considered ($\Gamma = 0$ there, but not its derivatives), we find $\mathrm{d}^2\xi^\lambda/\mathrm{d}\tau^2 = -\partial_\sigma\Gamma^\lambda{}_{\nu\rho}\,\xi^\sigma u^\nu u^\rho$. In the same frame the second covariant derivative along the curve is $\mathrm{D}^2\xi^\lambda/\mathrm{D}\tau^2 = \mathrm{d}^2\xi^\lambda/\mathrm{d}\tau^2 + \partial_\rho\Gamma^\lambda{}_{\nu\sigma}\,u^\rho u^\nu\xi^\sigma$. Thus

$$\frac{\mathrm{D}^2\xi^\lambda}{\mathrm{D}\tau^2} = \left(\frac{\partial\Gamma^\lambda{}_{\nu\sigma}}{\partial x^\rho} - \frac{\partial\Gamma^\lambda{}_{\nu\rho}}{\partial x^\sigma}\right)u^\nu\xi^\sigma u^\rho = R^\lambda{}_{\nu\sigma\rho}\,u^\nu\xi^\sigma u^\rho, \tag{7.129}$$

where we recognised the Riemann tensor of Chapter 5 at a point where $\Gamma = 0$. Being a tensor equation, (7.129) holds in all coordinates. This is the geodesic deviation equation: relative accelerations of freely falling particles (tidal forces) are the physical, coordinate-independent manifestation of curvature. Two checks:

- Newtonian limit: for $g_{00} = 1 + 2\phi$ and particles at rest, $R^i{}_{0j0} = -\partial_i\partial_j\phi$, so $\mathrm{d}^2\xi^i/\mathrm{d}t^2 = -\partial_i\partial_j\phi\,\xi^j$, which is the Newtonian tidal acceleration.
- TT wave: from (7.125), to first order in h, $R^i{}_{0j0} = \partial_0\Gamma^i{}_{0j} = -\frac{1}{2}\ddot{h}_{ij} = \frac{1}{2}\ddot{h}^i{}_j$, which reproduces (7.128).

The Riemann tensor of the wave is gauge invariant in the linear approximation, which confirms that the oscillating proper distances are physical.

Polarisations. Consider a ring of free particles in the xy-plane, $\xi^i = \rho(\cos\theta, \sin\theta, 0)$, and the purely "plus" wave $h_+ = h\cos\omega t$, $h_\times = 0$ (at $z = 0$). Then

$$X^1 = \rho\cos\theta\left(1 + \frac{h}{2}\cos\omega t\right), \qquad X^2 = \rho\sin\theta\left(1 - \frac{h}{2}\cos\omega t\right): \tag{7.130}$$

the ring is stretched along x while it is squeezed along y, and half a period later the other way round. For the "cross" wave $h_+ = 0$, $h_\times = h\cos\omega t$ we get $X^1 = \rho(\cos\theta + \frac{h}{2}\cos\omega t\sin\theta)$, $X^2 = \rho(\sin\theta + \frac{h}{2}\cos\omega t\cos\theta)$, the same pattern rotated by 45°. This is the helicity-two nature of the wave seen geometrically: a rotation by 45° turns h_+ into $h_\times$, and a rotation by 90° turns h_+ into $-h_+$ (compare $e'_\pm = e^{\pm 2\mathrm{i}\theta}e_\pm$). Particles along the direction of propagation are not affected (the wave is transverse), and the area of the ring is unchanged to first order (the wave is traceless). Circularly polarised waves, $h_+ = h\cos\omega t$, $h_\times = \pm h\sin\omega t$, make the ellipse rotate. The two linear polarisations are illustrated in Fig. 7.2.

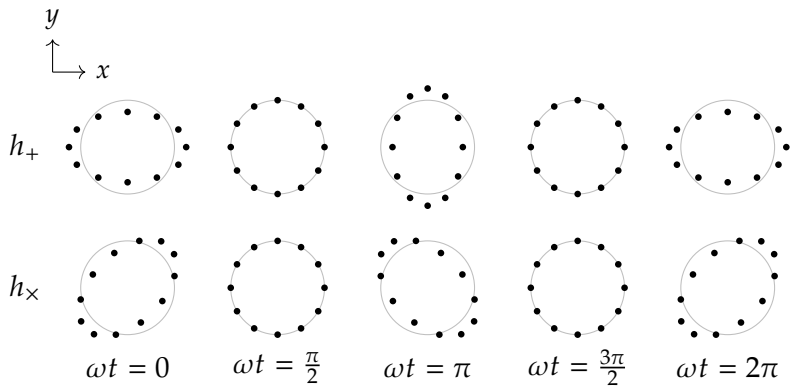


Figure 7.2. Deformation of a ring of free test masses in the xy-plane by a wave travelling along z, Eq. (7.128), for the plus (top) and cross (bottom) polarisations, with the (enormously exaggerated) amplitude $h = 0.5$.

7.6 The quadrupole formula

We now return to the retarded solution, (7.16),

$$h_{\mu\nu}(\vec{x}, t) = -4G\int \mathrm{d}^3x'\,\frac{S_{\mu\nu}(\vec{x}', t - |\vec{x} - \vec{x}'|)}{|\vec{x} - \vec{x}'|}, \tag{7.131}$$

and assume that the source is slowly moving. Its characteristic frequencies ω satisfy $\omega R \sim v \ll 1$, with v the typical velocity. Then, at large distances, $t - |\vec{x} - \vec{x}'| \approx t - r + \hat{x} \cdot \vec{x}'$ and the shift $\hat{x} \cdot \vec{x}' \lesssim R$ of the retarded time is negligible compared with the period (in Fourier space, $e^{-i\vec{k}\cdot\vec{x}'} \approx 1$). We obtain

$$h_{\mu\nu}(\vec{x}, t) \approx -\frac{4G}{r} \int \mathrm{d}^3x'\, S_{\mu\nu}(\vec{x}', t - r). \tag{7.132}$$

We have seen that only the transverse spatial components of the wave are physical; all others can be removed by a gauge transformation. So we need only the spatial components h_{ij}. With $T_{ij} = T^{ij}$ (two lowered spatial indices) and $\eta_{ij} = -\delta_{ij}$,

$$S_{ij} = T_{ij} - \frac{\eta_{ij}}{2} T^{\alpha}{}_{\alpha} = T^{ij} + \frac{\delta_{ij}}{2} T^{\alpha}{}_{\alpha}. \tag{7.133}$$

The second term is proportional to δ_{ij}; it will be removed by the projection on the transverse-traceless part below.

The stresses are second time derivatives of the mass distribution. The spatial stresses T^{ij} of a slowly moving source are not easy to know (for a binary star they include the gravitational binding). Fortunately we do not need them. Conservation of energy and momentum, $\partial_\mu T^{\mu\nu} = 0$, gives

$$\frac{\partial T^{00}}{\partial t} = -\frac{\partial T^{k0}}{\partial x^k}, \qquad \frac{\partial T^{0l}}{\partial t} = -\frac{\partial T^{kl}}{\partial x^k} \;\Rightarrow\; \frac{\partial^2 T^{00}}{\partial t^2} = \frac{\partial^2 T^{kl}}{\partial x^k \partial x^l}. \tag{7.134}$$

Multiplying by $x^i x^j$ and integrating by parts twice (the source is localised, so there are no surface terms),

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} \int \mathrm{d}^3x\, T^{00} x^i x^j = \int \mathrm{d}^3x\, T^{kl} \frac{\partial^2 (x^i x^j)}{\partial x^k \partial x^l} = 2 \int \mathrm{d}^3x\, T^{ij}. \tag{7.135}$$

(For a self-gravitating source $T^{\mu\nu}$ is to be read as the conserved total energy-momentum of matter and gravitation, $t^{\mu\nu}$ of the previous section; to the order at which we work, this amounts to Newtonian mechanics.) For a non-relativistic source $T^{00} \approx \rho$, the mass density. Defining the second moment of the mass distribution and its traceless part, the quadrupole moment,

$$I_{ij}(t) = \int \mathrm{d}^3x\, \rho(\vec{x}, t)\, x^i x^j, \qquad Q_{ij}(t) = I_{ij} - \frac{\delta_{ij}}{3} I_{kk} = \int \mathrm{d}^3x\, \rho \left(x^i x^j - \frac{\delta_{ij}}{3} r^2 \right), \tag{7.136}$$

we find

$$h_{ij}(\vec{x}, t) \approx -\frac{2G}{r} \ddot{I}_{ij}(t - r) + (\text{terms} \propto \delta_{ij}). \tag{7.137}$$

Transverse-traceless projection. For a wave propagating in the direction $\hat{n} = \hat{x}$ we define the projector onto the plane transverse to $\hat{n}$,

$$P_{ij} = \delta_{ij} - \hat{n}_i \hat{n}_j, \qquad P_{ij}\hat{n}_j = 0, \quad P_{ik}P_{kj} = P_{ij}, \quad P_{ii} = 2, \tag{7.138}$$

and, for any symmetric A_{ij}, its transverse-traceless part

$$A^{\mathrm{TT}}_{ij} = \Lambda_{ij,kl} A_{kl}, \qquad \Lambda_{ij,kl} = P_{ik}P_{jl} - \frac{1}{2} P_{ij}P_{kl}. \tag{7.139}$$

One checks that $A^{\mathrm{TT}}_{ij}\hat{n}_j = 0$, $A^{\mathrm{TT}}_{ii} = 0$, and that $\delta^{\mathrm{TT}}_{ij} = P_{ij} - \frac{1}{2}P_{ij}\cdot 2 = 0$: the trace terms drop out, and so $I^{\mathrm{TT}}_{ij} = Q^{\mathrm{TT}}_{ij}$. For $\hat{n} = \hat{z}$ the projection simply keeps the xy block and removes its trace, $A^{\mathrm{TT}}_{11} = -A^{\mathrm{TT}}_{22} = \frac{1}{2}(A_{11} - A_{22})$, $A^{\mathrm{TT}}_{12} = A_{12}$; these are exactly the two components which survived the gauge fixing of the plane wave. The radiation field is therefore

$$\boxed{h^{\mathrm{TT}}_{ij}(\vec{x},t) = -\frac{2G}{r}\,\ddot{Q}^{\mathrm{TT}}_{ij}(t-r)}, \qquad \text{i.e.} \qquad h^{i}{}_{j}{}^{\mathrm{TT}} = +\frac{2G}{r}\,\ddot{Q}^{\mathrm{TT}}_{ij}(t-r). \tag{7.140}$$

The minus sign in the first form is due to our signature: $h^{i}{}_{j}$ of (7.124) enters the spatial metric as $-g_{ij} = \delta_{ij} + h^{i}{}_{j}$, and it agrees with the familiar $+\frac{2G}{r}\ddot{Q}^{\mathrm{TT}}_{ij}$ of the literature in signature $(-,+,+,+)$. The overall sign is the one of the Newtonian check of the retarded solution, $h_{00} = -4G\int S_{00}/r = -2GM/r = 2\phi$. Note also that the wave is generated by the second time derivative of the quadrupole moment: there is no monopole radiation (mass is conserved, and so is $\int T^{00}$) and no dipole radiation (the momentum $\int T^{0i} = \frac{\mathrm{d}}{\mathrm{d}t}\int \rho\, x^i$ is conserved). This is unlike electromagnetism, where the electric dipole moment of a system of charges can oscillate.

Radiated power. For an outgoing wave in the direction $\hat{n}$ the energy flux is $\hat{n}^i\langle t^{0i}\rangle = \langle t^{00}\rangle$, since $\langle t^{\mu\nu}\rangle \propto k^\mu k^\nu$. In TT gauge our result $\langle t_{00}\rangle = \frac{k_0^2}{8\pi G}\left(|\varepsilon_{11}|^2 + |\varepsilon_{12}|^2\right)$ can be written in the time domain as

$$\langle t^{00}\rangle = \frac{1}{32\pi G}\left\langle \dot{h}^{\mathrm{TT}}_{ij}\dot{h}^{\mathrm{TT}}_{ij}\right\rangle = \frac{1}{16\pi G}\left\langle \dot{h}_+^2 + \dot{h}_\times^2\right\rangle, \tag{7.141}$$

as one verifies by inserting $h_{ij} = \varepsilon_{ij}e^{-\mathrm{i}\omega(t-z)}$ + c.c. and averaging. With (7.140) the power per solid angle is then

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} = r^2\langle t^{00}\rangle = \frac{G}{8\pi}\left\langle \ddot{Q}^{\mathrm{TT}}_{ij}\ddot{Q}^{\mathrm{TT}}_{ij}\right\rangle, \tag{7.142}$$

where $\ddot{Q}_{ij}$ is evaluated at the retarded time and the $1/r^2$ fall-off of the energy flux is compensated by the area $r^2\mathrm{d}\Omega$. Using (7.139), $\Lambda_{ij,kl}\Lambda_{ij,mn} = \Lambda_{kl,mn}$ and $Q_{ii} = 0$,

$$\ddot{Q}^{\mathrm{TT}}_{ij}\ddot{Q}^{\mathrm{TT}}_{ij} = \ddot{Q}_{ij}\ddot{Q}_{ij} - 2\hat{n}_i\ddot{Q}_{ij}\ddot{Q}_{jk}\hat{n}_k + \frac{1}{2}\left(\hat{n}_i\ddot{Q}_{ij}\hat{n}_j\right)^2. \tag{7.143}$$

The angular integrals $\int \mathrm{d}\Omega\,\hat{n}_i\hat{n}_j = \frac{4\pi}{3}\delta_{ij}$ and $\int \mathrm{d}\Omega\,\hat{n}_i\hat{n}_j\hat{n}_k\hat{n}_l = \frac{4\pi}{15}(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$ give $4\pi\left(1 - \frac{2}{3} + \frac{1}{15}\right)\ddot{Q}_{ij}\ddot{Q}_{ij} = \frac{8\pi}{5}\ddot{Q}_{ij}\ddot{Q}_{ij}$. Therefore the total radiated power is

$$\boxed{P = \frac{G}{5}\left\langle \dddot{Q}_{ij}\dddot{Q}_{ij}\right\rangle} \qquad \left(= \frac{G}{5c^5}\left\langle \dddot{Q}_{ij}\dddot{Q}_{ij}\right\rangle \text{ in ordinary units}\right). \tag{7.144}$$

This is Einstein's quadrupole formula (1918). It can also be obtained from the general expression for $\mathrm{d}E/\mathrm{d}\Omega$ derived above, by expanding $T_{\mu\nu}(\vec{k},\omega)$ to lowest order in ωR (Weinberg, Ch. 10). The factor $1/c^5$ makes gravitational radiation from laboratory sources hopelessly small: the natural unit of power is $c^5/G = 3.6\times 10^{52}\,\mathrm{W}$, and $P \sim \frac{c^5}{G}\left(\frac{r_S}{R}\right)^2\left(\frac{v}{c}\right)^6$ for a source of mass $M = r_S c^2/(2G)$, size R and velocity v. Only compact (r_S/R not small) and fast (v/c not small) astrophysical systems radiate appreciably.

Angular distribution. Unlike the total power, (7.142) depends on the direction through the projector. In contrast to the dipole pattern of electromagnetism, the radiation of a quadrupole source is in general non-zero in every direction. We will see an example for a binary below: there the power emitted along the orbital axis is eight times the power emitted in the orbital plane.

7.7 Binary systems

Consider two stars (or black holes) of masses m_1, m_2, total mass $M = m_1 + m_2$ and reduced mass $\mu = m_1 m_2/M$, on a circular Newtonian orbit of radius a (the separation) in the xy-plane. In the centre-of-mass frame the relative coordinate is $\vec{x}(t) = a(\cos\omega t, \sin\omega t, 0)$, with Kepler's third law $\omega^2 a^3 = GM$. The second moment of the mass distribution is that of a single particle of mass μ at $\vec{x}(t)$,

$$I_{ij} = \mu x^i x^j = \frac{\mu a^2}{2}\begin{pmatrix} 1+\cos 2\omega t & \sin 2\omega t & 0 \\ \sin 2\omega t & 1-\cos 2\omega t & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{7.145}$$

The time-dependent part oscillates at twice the orbital frequency, $\omega_{\rm gw} = 2\omega$: after half an orbit the mass distribution looks the same.

Waveform. Let us observe the binary at distance r, at an angle ι between the line of sight and the orbital axis. Applying (7.140), with h_+, $h_\times$ defined with respect to the polarisation axes $\hat{e}_\theta$, $\hat{e}_\varphi$ transverse to the line of sight ($\varphi = 0$),

$$\begin{aligned} h_+ &= -\frac{4G\mu\omega^2 a^2}{r}\,\frac{1+\cos^2\iota}{2}\,\cos(2\omega t_r)\,, \\ h_\times &= -\frac{4G\mu\omega^2 a^2}{r}\,\cos\iota\,\sin(2\omega t_r)\,, \end{aligned} \qquad t_r = t - r. \tag{7.146}$$

Along the axis ($\iota = 0$) the wave is circularly polarised; in the orbital plane ($\iota = \pi/2$) it is linearly polarised with half the amplitude. Eliminating a with Kepler's law and introducing the gravitational-wave frequency $f = \omega_{\rm gw}/(2\pi) = \omega/\pi$, the amplitude for $\iota = 0$ is

$$h_0 = \frac{4G\mu\omega^2 a^2}{r} = \frac{4}{r}(G\mathcal{M})^{5/3}(\pi f)^{2/3}, \qquad \mathcal{M} \equiv \mu^{3/5}M^{2/5} = \frac{(m_1 m_2)^{3/5}}{(m_1+m_2)^{1/5}}. \tag{7.147}$$

The masses enter only through the chirp mass $\mathcal{M}$.

Power. Inserting Q_{ij} into (7.144) (only the $\cos 2\omega t$, $\sin 2\omega t$ terms contribute, each derivative gives a factor 2ω),

$$P = \frac{32}{5}G\mu^2 a^4\omega^6 = \frac{32}{5}\frac{G^4\mu^2 M^3}{a^5} = \frac{32}{5}\left(G\mathcal{M}\,\pi f\right)^{10/3}\frac{1}{G}. \tag{7.148}$$

The angular distribution, from (7.142), is

$$\frac{\mathrm{d}P}{\mathrm{d}\Omega} = \frac{2G\mu^2 a^4\omega^6}{\pi}\left[\left(\frac{1+\cos^2\iota}{2}\right)^2 + \cos^2\iota\right], \tag{7.149}$$

which integrates to (7.148), and is 8 times larger along the axis than in the orbital plane. For the Earth–Sun system (7.148) gives about 200 W: completely negligible.

Chirp. The energy radiated is taken from the orbital energy $E = -\frac{GM\mu}{2a}$. With $\mathrm{d}E/\mathrm{d}t = -P$ the orbit shrinks, $\dot{a} = -\frac{64}{5}G^3\mu M^2/a^3$, and the frequency increases. Using $f = \frac{1}{\pi}\sqrt{GM/a^3}$, i.e. $\dot{f}/f = -\frac{3}{2}\dot{a}/a$, we find

$$\dot{f} = \frac{96}{5}\pi^{8/3}(G\mathcal{M})^{5/3} f^{11/3}. \tag{7.150}$$

Again only the chirp mass appears, so that measuring f and $\dot f$ determines it, $G\mathcal{M} = \left[\frac{5}{96}\pi^{-8/3} f^{-11/3}\dot f\right]^{3/5}$ (this is how its name arose). Integrating (7.150), the time left until coalescence (formally $f \to \infty$) is

$$\tau = \frac{5}{256}\,(G\mathcal{M})^{-5/3}\,(\pi f)^{-8/3}, \qquad \text{i.e.} \qquad f(\tau) = \frac{1}{\pi}\left(\frac{5}{256\,\tau}\right)^{3/8}(G\mathcal{M})^{-5/8}. \tag{7.151}$$

Both the frequency and, by (7.147), the amplitude $h_0 \propto f^{2/3}$ grow as the merger approaches: the signal is a "chirp". Our Newtonian treatment breaks down when the orbital velocity $v = (\pi G M f)^{1/3}$ becomes comparable to 1, i.e. for the last few orbits, where numerical relativity takes over.

The Hulse–Taylor pulsar. The orbital period $P_b = 2\pi/\omega = 2/f$ decreases at the rate $\dot P_b = -2\dot f/f^2$. For circular orbits this is $\dot P_b = -\frac{192\pi}{5}\,(2\pi G\mathcal{M}/P_b)^{5/3}$. Real binaries are often eccentric; Peters and Mathews computed the quadrupole radiation of a Kepler ellipse of eccentricity e and found that it is enhanced by the factor

$$\dot P_b = -\frac{192\pi}{5}\left(\frac{2\pi G\mathcal{M}}{P_b}\right)^{5/3} f(e), \qquad f(e) = \frac{1 + \frac{73}{24}e^2 + \frac{37}{96}e^4}{(1-e^2)^{7/2}}. \tag{7.152}$$

The binary pulsar PSR B1913+16, discovered by R. Hulse and J. Taylor in 1974, consists of two neutron stars with $m_p = 1.4398\,M_\odot$ and $m_c = 1.3886\,M_\odot$ (these masses are themselves determined from relativistic effects in the pulsar timing: periastron precession and time dilation), orbital period $P_b = 0.322997\,\mathrm{d} = 7.75\,\mathrm{h}$ and eccentricity $e = 0.6171$. The semi-major axis is $a = 1.95\times10^9\,\mathrm{m} \approx 2.8\,R_\odot$ and the velocities are $v \sim 10^{-3}$, so that the quadrupole approximation is excellent. With $\mathcal{M} = 1.231\,M_\odot$, $GM_\odot/c^3 = 4.925\,\mu\mathrm{s}$ and $f(e) = 11.86$, (7.152) gives

$$\dot P_b^{\mathrm{GR}} = -2.403\times10^{-12}, \tag{7.153}$$

i.e. the period decreases by about 76 μs per year. The observed value, after correcting for the (small) kinematic effect of the relative acceleration of the pulsar and the Sun in the Galaxy, is $\dot P_b^{\mathrm{obs}} = (-2.398 \pm 0.004)\times10^{-12}$, in agreement at the 0.2% level. After three decades of observation the accumulated shift of the time of periastron passage is about 40 s. This was the first (indirect) evidence for gravitational waves, recognised with the Nobel Prize in Physics of 1993. Integrating the Peters equations for $a(t)$ and $e(t)$, the two neutron stars will merge in about 3×10^8 years. **References.** Quadrupole formula and binaries: Weinberg, Ch. 10; M. Maggiore, Gravitational Waves, Vol. 1, Chs. 3–4; P. C. Peters and J. Mathews, Phys. Rev. **131** (1963) 435; P. C. Peters, Phys. Rev. **136** (1964) B1224. PSR B1913+16: R. A. Hulse and J. H. Taylor, Astrophys. J. **195** (1975) L51; J. M. Weisberg and Y. Huang, Astrophys. J. **829** (2016) 55 (masses, orbital elements, $\dot P_b^{\mathrm{obs}}/\dot P_b^{\mathrm{GR}} = 0.9983 \pm 0.0016$).

7.8 Detection of gravitational waves

A passing wave changes the proper distances between free test masses by the fraction $\delta L/L = \frac{1}{2}h^i{}_j\hat\xi^i\hat\xi^j$. The idea of an interferometric detector is to compare the lengths of two perpendicular arms. Mirrors suspended as pendulums act, at frequencies well above the pendulum frequency, as free masses in the horizontal directions. For a wave travelling along z and arms of length L along x and y, a light ray along the x-arm has, from (7.125), $\mathrm{d}t = (1 + \frac{1}{2}h_+)\mathrm{d}x$, and along the y-arm $\mathrm{d}t = (1 - \frac{1}{2}h_+)\mathrm{d}y$ (for wavelengths much longer

than the arms, so that h_+ is constant during the light travel time). The round-trip times then differ by

$$\Delta T = 2L\,h_+ \qquad \Longleftrightarrow \qquad \Delta L \equiv L_x - L_y = h_+ L, \tag{7.154}$$

which is measured as a change of the interference at the output of a Michelson interferometer. For a general direction of incidence the signal is a combination $F_+ h_+ + F_\times h_\times$ with antenna-pattern factors $|F_{+,\times}| \le 1$ which depend on the orientation of the detector relative to the source.

The numbers are daunting. With $h \sim 10^{-21}$ and $L = 4\,\mathrm{km}$ the change in arm length is $\Delta L \sim 4 \times 10^{-18}\,\mathrm{m}$, about $1/200$ of the radius of a proton. The two LIGO detectors (Hanford and Livingston, USA, $L = 4\,\mathrm{km}$) and Virgo (Cascina, Italy, $L = 3\,\mathrm{km}$) reach this sensitivity in the band ~ 20–$2000\,\mathrm{Hz}$, thanks to Fabry–Perot cavities in the arms (the light bounces some hundred times), high laser power, seismic isolation and quantum-optics techniques (squeezed light). Comparing several detectors far apart removes local disturbances and allows one to locate the source on the sky.

GW150914. On 14 September 2015 the two LIGO detectors recorded, 7 ms apart, a signal whose frequency swept from 35 to 150 Hz in about 0.2 s, with a peak strain of 1.0×10^{-21}. Let us check that our formulae reproduce its main features.

- From f and $\dot f$, (7.150) gives a chirp mass $\mathcal{M} \approx 30\,M_\odot$ (in the detector frame; the source-frame value is smaller by a factor $1 + z = 1.09$). For two equal masses $M = 2^{6/5}\mathcal{M} \approx 70\,M_\odot$.
- The time to coalescence from $f = 35\,\mathrm{Hz}$ is $\tau = 0.17\,\mathrm{s}$, in agreement with the duration of the signal.
- At $f = 150\,\mathrm{Hz}$ Kepler's law gives a separation $a = (GM/\pi^2 f^2)^{1/3} \approx 350\,\mathrm{km}$, less than twice the Schwarzschild radius $2GM \approx 210\,\mathrm{km}$ of the total mass, and an orbital velocity $v \approx 0.5$. Only black holes are compact enough to come this close without touching (neutron stars would be too light).
- At a distance $r = 410\,\mathrm{Mpc}$, (7.147) gives $h_0 \approx 2 \times 10^{-21}$ at 150 Hz; the factor two with respect to the observed peak is accounted for by the orientation of the source, the antenna pattern of the detectors, and the breakdown of the quadrupole approximation close to the merger.

The full analysis, with waveforms from post-Newtonian theory and numerical relativity, gives two black holes of 36 and 29 $M_\odot$ merging into a black hole of 62 $M_\odot$ at a distance of $410^{+160}_{-180}\,\mathrm{Mpc}$ ($z \approx 0.09$). The difference, $3\,M_\odot c^2 \approx 5 \times 10^{47}\,\mathrm{J}$, was radiated in gravitational waves, with a peak power of $3.6 \times 10^{49}\,\mathrm{W}$, about $10^{-3} c^5/G$ and more than the combined light output of all stars in the observable universe. The Nobel Prize in Physics 2017 was awarded to R. Weiss, B. Barish and K. Thorne for the detection. By 2020 about ninety mergers had been observed (catalogue GWTC-3), and many more since, among them the binary neutron star GW170817, followed 1.7 s later by a gamma-ray burst and by a kilonova observed across the electromagnetic spectrum; it established that gravitational waves travel at the speed of light to within one part in 10^{15}. **Other frequency bands.** Ground-based interferometers are limited at low frequencies by seismic and gravity-gradient noise. Supermassive black-hole binaries (10^5–$10^9\,M_\odot$) radiate at much lower frequencies. The space interferometer LISA (ESA/NASA, three spacecraft with 2.5×10^6 km arms, planned

launch in the 2030s) will cover 10^{-4}–10^{-1} Hz: massive black-hole mergers, extreme-mass-ratio inspirals and thousands of galactic white-dwarf binaries. At nanohertz frequencies pulsar timing arrays (PTA) use millisecond pulsars as clocks: a passing wave modulates the arrival times of their pulses with a characteristic angular correlation between pulsars (the Hellings–Downs curve). In 2023 the NANOGrav, EPTA/InPTA, PPTA and CPTA collaborations reported evidence for a stochastic background at these frequencies, most likely from the population of supermassive black-hole binaries. References: B. P. Abbott et al. (LIGO Scientific and Virgo Collaborations), Phys. Rev. Lett. **116** (2016) 061102 (GW150914), and Ann. Phys. (Berlin) **529** (2017) 1600209 (the basic physics of GW150914); Phys. Rev. Lett. **119** (2017) 161101 (GW170817); R. Abbott et al., Phys. Rev. X **13** (2023) 041039 (catalogue GWTC-3); P. Amaro-Seoane et al., arXiv:1702.00786 (LISA); G. Agazie et al. (NANOGrav), Astrophys. J. Lett. **951** (2023) L8; Maggiore, Vol. 1, Ch. 9 (detectors).

7.9 Exercises

Exercise.. (Laboratory sources.) (a) A rod of mass m and length L rotates with angular velocity Ω about an axis through its centre, perpendicular to it. Show with (7.144) that it radiates $P = \frac{2}{45}\frac{G}{c^5} m^2 L^4 \Omega^6$. (b) Evaluate it for a steel beam with $m = 500\,\mathrm{t}$, $L = 20\,\mathrm{m}$, $\Omega = 30\,\mathrm{s}^{-1}$ (close to the point where it would break apart). (c) Evaluate (7.148) for the Earth–Sun system. [Answers: (b) $P \approx 4 \times 10^{-29}\,\mathrm{W}$; (c) $P \approx 200\,\mathrm{W}$.]

Exercise.. (Interferometer response.) (a) Starting from (7.125), compute the round-trip time of light along an arm of length L in the direction $(\cos\alpha, \sin\alpha, 0)$, for a wave along z with constant h_+, $h_\times$ during the light travel time. (b) Show that a detector with arms along x and y responds only to h_+, and a detector with arms rotated by 45° only to $h_\times$. (c) Show that the detector does not respond at all to a wave travelling along one of its arms with polarisation $h_\times$ relative to the plane of the arms. [Answer (a): $T = 2L\left(1 + \frac{1}{2}h_+ \cos 2\alpha + \frac{1}{2}h_\times \sin 2\alpha\right)$.]

Exercise.. (Chirp and coalescence.) (a) Derive (7.150) and the time to coalescence $\tau(f)$ from (7.148), Kepler's law and $E = -GM\mu/(2a)$. (b) Show that the number of gravitational-wave cycles between frequency f and coalescence is $N = \int f\,\mathrm{d}t = \frac{1}{32\pi^{8/3}}(G\mathcal{M})^{-5/3} f^{-5/3}$, and evaluate it for $\mathcal{M} = 30\,M_\odot$ from $f = 35\,\mathrm{Hz}$ and for a binary neutron star ($\mathcal{M} = 1.2\,M_\odot$) from $f = 20\,\mathrm{Hz}$, the lower edge of the LIGO band. (c) For PSR B1913+16, evaluate (7.152) with the data given in the text and check that $\dot{P}_b = -2.40 \times 10^{-12}$. [Answers: (b) $N \approx 10$ and $N \approx 5 \times 10^3$.]

8

Action principle for gravity

The Einstein equations (for $\Lambda = 0$; the cosmological constant is added below) have the form

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + 8\pi G T_{\mu\nu} = 0, \tag{8.1}$$

where $R_{\mu\nu}$ is the Ricci tensor, $R = g^{\alpha\beta} R_{\alpha\beta}$ the curvature scalar and $T_{\mu\nu}$ is the energy-momentum tensor. We will show that the above can be derived as the Euler–Lagrange equations of an action integral

$$S = \int \mathrm{d}^4 x \sqrt{g} \mathcal{L}, \tag{8.2}$$

where the Lagrangian density is the sum of a gravitational and a matter part,

$$\mathcal{L} = \mathcal{L}_G + \mathcal{L}_M. \tag{8.3}$$

In the above, $\mathcal{L}_M$ generates the energy-momentum tensor term in Einstein's equations.

The action must be invariant under general coordinate transformations,

$$\begin{aligned} &x \to x' \\ &S' = \int \mathrm{d}^4 x' \sqrt{g'} \mathcal{L}' = \int \mathrm{d}^4 x \sqrt{g} \mathcal{L}. \end{aligned} \tag{8.4}$$

We have already seen that the measure $\mathrm{d}^4 x \sqrt{g}$ is invariant under such transformations,

$$\mathrm{d}^4 x' \sqrt{g'} = \mathrm{d}^4 x \sqrt{g}. \tag{8.5}$$

8.1 The gravitational part

We focus first on the gravitational term $\mathcal{L}_G$, ignoring (or setting to zero) the energy-momentum tensor. To maintain invariance under coordinate transformations, $\mathcal{L}_G$ must be a scalar,

$$\mathcal{L}_G \xrightarrow{x \to x'} \mathcal{L}'_G = \mathcal{L}_G. \tag{8.6}$$

A distinguished scalar that we know of is the curvature scalar. **Remark.** R is not the only scalar one can build (R^2, $R_{\mu\nu} R^{\mu\nu}$, ... are also scalars). What singles out $\sqrt{g} R$ (plus a constant, -2Λ) in $D = 4$ is that it is the only choice built from the metric whose

field equations contain at most second derivatives of $g_{\mu\nu}$ (Lovelock's theorem; Weinberg Sec. 7.1 for the analogous uniqueness of the field equations). Higher-curvature terms are suppressed at long distances. Let us then take

$$\begin{aligned} S_G &= \int \mathrm{d}^4x\sqrt{g}\mathcal{L} \\ &= \text{const}\int \mathrm{d}^4x\sqrt{g}R \\ &= \text{const}\int \mathrm{d}^4x\sqrt{g}g^{\alpha\beta}R_{\alpha\beta}. \end{aligned} \tag{8.7}$$

We require that the action is stationary under variations of the metric field,

$$g_{\mu\nu} \to g_{\mu\nu} + \delta g_{\mu\nu}. \tag{8.8}$$

We have

$$\delta S_G = \text{const}\int \mathrm{d}^4x\,\delta(\sqrt{g}R), \tag{8.9}$$

with

$$\begin{aligned} \delta(\sqrt{g}R) &= (\delta\sqrt{g})R + \sqrt{g}\delta(g^{\alpha\beta}R_{\alpha\beta}) \\ &= R\delta\sqrt{g} + \sqrt{g}(\delta g^{\alpha\beta})R_{\alpha\beta} + \sqrt{g}g^{\alpha\beta}\delta R_{\alpha\beta}, \end{aligned} \tag{8.10}$$

where $g = -\det\left(g_{\mu\nu}\right)$. Let us consider the variation

$$\begin{aligned} \delta\log g &= \delta\log\left(-\det\left(g_{\mu\nu}\right)\right) \\ &= \log\left(-\det\left(g_{\mu\nu}+\delta g_{\mu\nu}\right)\right) - \log\left(-\det\left(g_{\mu\nu}\right)\right) \\ &= \log\left(\frac{\det\left(g_{\mu\nu}+\delta g_{\mu\nu}\right)}{\det\left(g_{\mu\nu}\right)}\right) = \log\left(\det\left(g^{\mu\rho}(g_{\rho\nu}+\delta g_{\rho\nu})\right)\right) \\ &= \log\left(\det\left(\delta^{\mu}{}_{\nu} + g^{\rho\mu}\delta g_{\rho\nu}\right)\right) = \log\left(1+\mathrm{Tr}\left(g^{\rho\mu}\delta g_{\rho\nu}\right)+O(\delta g^2)\right) \\ &= \log\left(1+\mathrm{Tr}\left(g^{\rho\mu}\delta g_{\rho\nu}\right)\right)+O(\delta g^2) \\ &= \log\left(1+g^{\mu\rho}\delta g_{\mu\rho}\right)+O(\delta g^2) \\ &= g^{\mu\rho}\delta g_{\mu\rho}+O(\delta g^2). \end{aligned} \tag{8.11}$$

Thus,

$$\begin{aligned} \delta\log g &= g^{\mu\rho}\delta g_{\mu\rho} \\ \Rightarrow\ \delta g^{1/2} &= \frac{1}{2\sqrt{g}}\delta g = \frac{1}{2}\sqrt{g}\delta\log g \\ \Rightarrow\ \delta\sqrt{g} &= \frac{1}{2}\sqrt{g}g^{\mu\rho}\delta g_{\mu\rho}. \end{aligned} \tag{8.12}$$

Also, from

$$\begin{aligned} g^{\alpha\beta}g_{\beta\gamma} &= \delta^{\alpha}{}_{\gamma} \\ 0 = \delta\left(g^{\alpha\beta}g_{\beta\gamma}\right) &= \left(\delta g^{\alpha\beta}\right)g_{\beta\gamma} + g^{\alpha\beta}\delta g_{\beta\gamma} \\ \left(\delta g^{\alpha\beta}\right)g_{\beta\gamma} &= -\left(\delta g_{\beta\gamma}\right)g^{\alpha\beta} \\ \left(\delta g^{\alpha\beta}\right)g_{\beta\gamma}g^{\gamma\delta} &= -\left(\delta g_{\beta\gamma}\right)g^{\alpha\beta}g^{\gamma\delta} \\ \Rightarrow\ \delta g^{\alpha\delta} &= -g^{\alpha\beta}g^{\delta\gamma}\delta g_{\beta\gamma}. \end{aligned} \tag{8.13}$$

The final variation we require is of the Ricci tensor. Recall that

$$R_{\mu\nu} = \frac{\partial \Gamma^{\alpha}{}_{\mu\alpha}}{\partial x^{\nu}} - \frac{\partial \Gamma^{\alpha}{}_{\mu\nu}}{\partial x^{\alpha}} + \Gamma^{\alpha}{}_{\mu\beta}\Gamma^{\beta}{}_{\nu\alpha} - \Gamma^{\alpha}{}_{\mu\nu}\Gamma^{\beta}{}_{\beta\alpha}. \tag{8.14}$$

To compute the variation of the Ricci tensor, we will require the variation of the affine connection. While the affine connection $\Gamma^{\alpha}{}_{\mu\nu}$ is not a tensor, it turns out that its variation is a tensor. To show this, let us look at the covariant derivative of the variation of the metric,

$$\delta g_{\mu\nu;\rho} = \left(\delta g_{\mu\nu}\right)_{;\rho} = \frac{\partial \delta g_{\mu\nu}}{\partial x^{\rho}} - \Gamma^{\alpha}{}_{\mu\rho}\delta g_{\alpha\nu} - \Gamma^{\alpha}{}_{\nu\rho}\delta g_{\alpha\mu}. \tag{8.15}$$

Then,

$$\begin{aligned}
\delta g_{\mu\rho;\nu} + \delta g_{\nu\rho;\mu} - \delta g_{\mu\nu;\rho} &= \frac{\partial \delta g_{\mu\rho}}{\partial x^{\nu}} + \frac{\partial \delta g_{\nu\rho}}{\partial x^{\mu}} - \frac{\partial \delta g_{\mu\nu}}{\partial x^{\rho}} + \cancel{\Gamma^{\alpha}{}_{\mu\rho}\delta g_{\alpha\nu}} + \cancel{\Gamma^{\alpha}{}_{\rho\nu}\delta g_{\alpha\mu}} \\
&\quad - \Gamma^{\alpha}{}_{\mu\nu}\delta g_{\alpha\rho} - \cancel{\Gamma^{\alpha}{}_{\rho\nu}\delta g_{\alpha\mu}} - \Gamma^{\alpha}{}_{\mu\nu}\delta g_{\alpha\rho} - \cancel{\Gamma^{\alpha}{}_{\rho\mu}\delta g_{\alpha\nu}} \\
&= \frac{\partial \delta g_{\mu\rho}}{\partial x^{\nu}} + \frac{\partial \delta g_{\nu\rho}}{\partial x^{\mu}} - \frac{\partial \delta g_{\mu\nu}}{\partial x^{\rho}} - 2\Gamma^{\alpha}{}_{\mu\nu}\delta g_{\alpha\rho} \\
&= \delta\left(\frac{\partial g_{\mu\rho}}{\partial x^{\nu}} + \frac{\partial g_{\nu\rho}}{\partial x^{\mu}} - \frac{\partial g_{\mu\nu}}{\partial x^{\rho}}\right) - 2\delta\left(\Gamma^{\alpha}{}_{\mu\nu}g_{\alpha\rho}\right) + g_{\alpha\rho}2\delta\Gamma^{\alpha}{}_{\mu\nu}.
\end{aligned} \tag{8.16}$$

Moreover,

$$\delta\left(\Gamma^{\alpha}{}_{\mu\nu}g_{\alpha\rho}\right) = \delta\Bigg(\underbrace{g_{\alpha\rho}g^{\alpha\lambda}}_{=\delta^{\lambda}{}_{\rho}}\frac{1}{2}\left(\frac{\partial g_{\lambda\mu}}{\partial x^{\nu}} + \frac{\partial g_{\lambda\nu}}{\partial x^{\mu}} - \frac{\partial g_{\mu\nu}}{\partial x^{\lambda}}\right)\Bigg) = \frac{1}{2}\delta\left(\frac{\partial g_{\rho\mu}}{\partial x^{\nu}} + \frac{\partial g_{\rho\nu}}{\partial x^{\mu}} - \frac{\partial g_{\mu\nu}}{\partial x^{\rho}}\right). \tag{8.17}$$

Thus, we have shown that

$$\begin{aligned}
2g_{\alpha\rho}\delta\Gamma^{\alpha}{}_{\mu\nu} &= \delta g_{\mu\rho;\nu} + \delta g_{\nu\rho;\mu} - \delta g_{\mu\nu;\rho} \\
\delta\Gamma^{\lambda}{}_{\mu\nu} &= \frac{1}{2}g^{\lambda\rho}\left(\delta g_{\mu\rho;\nu} + \delta g_{\nu\rho;\mu} - \delta g_{\mu\nu;\rho}\right).
\end{aligned} \tag{8.18}$$

Therefore, unlike the affine connection $\Gamma^{\alpha}{}_{\mu\nu}$ which is not a tensor, its variation $\delta\Gamma^{\alpha}{}_{\mu\nu}$ is a tensor. Let us now take the variation of the Ricci tensor in (8.14):

$$\begin{aligned}
\delta R_{\mu\nu} &= \frac{\partial \delta\Gamma^{\alpha}{}_{\mu\alpha}}{\partial x^{\nu}} - \frac{\partial \delta\Gamma^{\alpha}{}_{\mu\nu}}{\partial x^{\alpha}} + \left(\delta\Gamma^{\alpha}{}_{\mu\beta}\right)\Gamma^{\beta}{}_{\nu\alpha} + \Gamma^{\alpha}{}_{\mu\beta}\delta\Gamma^{\beta}{}_{\nu\alpha} - \left(\delta\Gamma^{\alpha}{}_{\mu\nu}\right)\Gamma^{\beta}{}_{\beta\alpha} - \Gamma^{\alpha}{}_{\mu\nu}\delta\Gamma^{\beta}{}_{\beta\alpha} \\
&= \underbrace{\frac{\partial \delta\Gamma^{\alpha}{}_{\mu\alpha}}{\partial x^{\nu}} - \Gamma^{\alpha}{}_{\mu\nu}\delta\Gamma^{\beta}{}_{\beta\alpha}}_{=(\delta\Gamma^{\alpha}{}_{\mu\alpha})_{;\nu}} \underbrace{- \frac{\partial \delta\Gamma^{\alpha}{}_{\mu\nu}}{\partial x^{\alpha}} + \left(\delta\Gamma^{\alpha}{}_{\mu\beta}\right)\Gamma^{\beta}{}_{\nu\alpha} + \Gamma^{\alpha}{}_{\mu\beta}\delta\Gamma^{\beta}{}_{\nu\alpha} - \left(\delta\Gamma^{\alpha}{}_{\mu\nu}\right)\Gamma^{\beta}{}_{\beta\alpha}}_{=-\delta\Gamma^{\alpha}{}_{\mu\nu;\alpha}} \\
&= \delta\Gamma^{\alpha}{}_{\mu\alpha;\nu} - \delta\Gamma^{\alpha}{}_{\mu\nu;\alpha}.
\end{aligned} \tag{8.19}$$

The above is the so-called Palatini equation. Here we have recognised the covariant derivatives of the variation of the affine connection (which is a tensor). Contracting with $g^{\mu\nu}$, we have

$$\begin{aligned}
g^{\mu\nu}\delta R_{\mu\nu} &= g^{\mu\nu}\delta\Gamma^{\alpha}{}_{\mu\alpha;\nu} - g^{\mu\nu}\delta\Gamma^{\alpha}{}_{\mu\nu;\alpha} \\
&= \left(g^{\mu\nu}\delta\Gamma^{\alpha}{}_{\mu\alpha}\right)_{;\nu} - \left(g^{\mu\nu}\delta\Gamma^{\alpha}{}_{\mu\nu}\right)_{;\alpha},
\end{aligned} \tag{8.20}$$

where we have used that the covariant derivative of the metric vanishes, $g^{\mu\nu}{}_{;\lambda} = 0$. Notice that the two terms above are covariant divergences of vectors. Recall then that

$$V^{\mu}{}_{;\mu} = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^{\mu}} \left(\sqrt{g} V^{\mu}\right). \tag{8.21}$$

We can then rewrite the equation in the form

$$\begin{aligned} g^{\mu\nu}\delta R_{\mu\nu} &= \frac{1}{\sqrt{g}} \left[\frac{\partial}{\partial x^{\nu}} \left(\sqrt{g} g^{\mu\nu} \delta\Gamma^{\alpha}{}_{\mu\alpha} \right) - \frac{\partial}{\partial x^{\alpha}} \left(\sqrt{g} g^{\mu\nu} \delta\Gamma^{\alpha}{}_{\mu\nu} \right) \right] \\ \Rightarrow \sqrt{g} g^{\mu\nu}\delta R_{\mu\nu} &= \frac{\partial}{\partial x^{\nu}} \left(\sqrt{g} g^{\mu\nu} \delta\Gamma^{\alpha}{}_{\mu\alpha} - \sqrt{g} g^{\mu\alpha} \delta\Gamma^{\nu}{}_{\mu\alpha} \right), \end{aligned} \tag{8.22}$$

which is a total derivative. **Remark.** Discarding this total derivative requires more than $\delta g_{\mu\nu} = 0$ on the boundary: the current contains $\delta\Gamma \sim \partial\delta g$, i.e. normal derivatives of $\delta g_{\mu\nu}$, which are not fixed in a Dirichlet variational problem. A well-posed principle adds the Gibbons–Hawking–York boundary term $\propto \oint \mathrm{d}^3y \sqrt{|h|}\, K$ (trace of the extrinsic curvature of the boundary), whose variation cancels these normal-derivative terms; it does not change the field equations. See Wald, App. E.1; Carroll, Sec. 4.3; Gibbons & Hawking, Phys. Rev. D15 (1977) 2752. Let us now return to the variation of the action, (8.10). Collecting the three contributions together, we have

$$\begin{aligned} \delta S_G &= \text{const} \int \mathrm{d}^4x \, \left[(\delta\sqrt{g}) R + \sqrt{g} (\delta g^{\mu\nu}) R_{\mu\nu} + \sqrt{g} g^{\mu\nu} \delta R_{\mu\nu} \right] \\ &= \text{const} \int \mathrm{d}^4x \, \left[\frac{1}{2} \sqrt{g} g^{\mu\nu} \delta g_{\mu\nu} R - R_{\alpha\beta} \sqrt{g} g^{\alpha\mu} g^{\beta\nu} \delta g_{\mu\nu} \right] + \cancelto{0}{\int \mathrm{d}^4x \frac{\partial}{\partial x^{\alpha}} (\cdots)} \\ &= \text{const} \int \mathrm{d}^4x \, \left[\frac{g^{\mu\nu}}{2} R - R^{\mu\nu} \right] \sqrt{g} \delta g_{\mu\nu}. \end{aligned} \tag{8.23}$$

Clearly, the variation of S_G gives the desired terms on the left-hand side of Einstein's equations. To add a cosmological constant, we write

$$S_G = \text{const} \int \mathrm{d}^4x \sqrt{g} (R - 2\Lambda). \tag{8.24}$$

Then,

$$\begin{aligned} \delta S_G &= \text{const} \int \mathrm{d}^4x \left[\left(-R^{\mu\nu} + \frac{g^{\mu\nu}}{2} R \right) \sqrt{g} \delta g_{\mu\nu} - 2\Lambda \left(\frac{1}{2} \sqrt{g} g^{\mu\nu} \delta g_{\mu\nu} \right) \right] \\ &= -\text{const} \int \mathrm{d}^4x \left[R^{\mu\nu} - \frac{g^{\mu\nu}}{2} R + \Lambda g^{\mu\nu} \right] \sqrt{g} \delta g_{\mu\nu}. \end{aligned} \tag{8.25}$$

8.2 Adding a matter field

Consider a scalar field $\phi(x)$ with a mass m. This will contribute to the action with the term

$$\begin{aligned} S_{\text{matter}} = S_M &= \frac{1}{2} \int \mathrm{d}^4x \sqrt{g} \left[(\partial_{\mu}\phi) (\partial^{\mu}\phi) - m^2\phi^2 \right] \\ &= \frac{1}{2} \int \mathrm{d}^4x \sqrt{g} \left[(\partial^{\mu}\phi) (\partial^{\nu}\phi) g_{\mu\nu} - m^2\phi^2 \right]. \end{aligned} \tag{8.26}$$

Notice the appearance of the metric, through which gravity couples to the field, as required by the principle of equivalence. Under $g_{\mu\nu} \to g_{\mu\nu} + \delta g_{\mu\nu}$ the variation of the action is

$$\begin{aligned} \delta S_M &= \int \mathrm{d}^4x \left[(\delta\sqrt{g})\mathcal{L}_M + \frac{1}{2}\sqrt{g}\left(\delta g^{\mu\nu}\right)(\partial_\mu\phi)(\partial_\nu\phi) \right] \\ &= \int \mathrm{d}^4x \left[\frac{1}{2}\sqrt{g}g^{\mu\nu}\mathcal{L}_M\delta g_{\mu\nu} - \frac{1}{2}(\partial^\mu\phi)(\partial^\nu\phi)\sqrt{g}\delta g_{\mu\nu} \right] \\ \to\ \delta S_M &= -\int \mathrm{d}^4x \frac{1}{2}\sqrt{g}\delta g_{\mu\nu}\left[(\partial^\mu\phi)(\partial^\nu\phi) - \mathcal{L}_M g^{\mu\nu} \right]. \end{aligned} \tag{8.27}$$

We define the energy-momentum tensor of the scalar field as

$$T_M^{\mu\nu} = (\partial^\mu\phi)(\partial^\nu\phi) - g^{\mu\nu}\mathcal{L}_M. \tag{8.28}$$

For $g^{\mu\nu} \to \eta^{\mu\nu}$, we recover the correct definition of the energy-momentum tensor for a scalar field in flat spacetime. Putting together the S_G and S_M terms, we have

$$S = S_G + S_M = \int \mathrm{d}^4x\sqrt{g}\left(\mathrm{const}\,(R - 2\Lambda) + \mathcal{L}_M\right). \tag{8.29}$$

The variation under $g_{\mu\nu} \to g_{\mu\nu} + \delta g_{\mu\nu}$ is

$$\delta S = -\frac{1}{2}\int \mathrm{d}^4x\sqrt{g}\delta g_{\mu\nu}\left[2\mathrm{const}\left(R^{\mu\nu} - \frac{g^{\mu\nu}}{2}R + \Lambda g^{\mu\nu} \right) + T^{\mu\nu} \right] = 0. \tag{8.30}$$

This indeed gives Einstein's equations if we normalise $\mathrm{const} = +\frac{1}{16\pi G}$. Then, the action takes the form

$$S = \int \mathrm{d}^4x\sqrt{g}\left(+\frac{1}{16\pi G}(R - 2\Lambda) + \mathcal{L}_M \right), \tag{8.31}$$

and the equations of motion obtained by varying the metric are

$$R^{\mu\nu} - \frac{g^{\mu\nu}}{2}R + \Lambda g^{\mu\nu} + 8\pi G T^{\mu\nu} = 0. \tag{8.32}$$

The action (8.31) is the so-called Einstein–Hilbert action. Before we conclude, let us also derive the Euler–Lagrange equations obtained by varying the scalar field,

$$\phi \to \phi + \delta\phi. \tag{8.33}$$

This yields

$$\begin{aligned} \delta S &= \int \mathrm{d}^4x\sqrt{g}\left[g^{\mu\nu}(\partial_\mu\phi)(\partial_\nu\delta\phi) - m^2\phi\delta\phi \right] \\ &= \int \mathrm{d}^4x\ \left[\partial_\nu\left(\sqrt{g}g^{\mu\nu}(\partial_\mu\phi)\delta\phi\right) - \partial_\nu\left(\sqrt{g}g^{\mu\nu}\partial_\mu\phi\right)\delta\phi - m^2\phi\sqrt{g}\delta\phi \right] \\ &= -\int \mathrm{d}^4x\sqrt{g}\delta\phi\left[m^2\phi + \frac{1}{\sqrt{g}}\partial_\nu\left(\sqrt{g}g^{\mu\nu}\partial_\mu\phi\right) \right], \end{aligned} \tag{8.34}$$

which yields the equations of motion

$$\begin{aligned} &\frac{1}{\sqrt{g}}\partial_\nu\left(\sqrt{g}g^{\mu\nu}\partial_\mu\phi\right) + m^2\phi = 0 \\ \Leftrightarrow\ &\frac{1}{\sqrt{g}}\partial_\nu\left(\sqrt{g}g^{\mu\nu}\phi_{;\mu}\right) + m^2\phi = 0 \\ \Leftrightarrow\ &\left(g^{\mu\nu}\phi_{;\mu}\right)_{;\nu} + m^2\phi = 0 \\ \Leftrightarrow\ &g^{\mu\nu}\phi_{;\mu;\nu} + m^2\phi = 0. \end{aligned} \tag{8.35}$$

For $g_{\mu\nu} = \eta_{\mu\nu}$ and $\Gamma^{\mu}{}_{\nu\rho} = 0$ we obtain the correct equation of motion for a scalar field in special relativity,

$$\eta^{\mu\nu}\partial_\mu\partial_\nu\phi + m^2\phi = 0 \;\Leftrightarrow\; (\Box + m^2)\phi = 0. \tag{8.36}$$

Analogously, we can couple gravity to the electromagnetic field (or to any other field; spinor fields, however, require some extra care and the vierbein, or tetrad, formalism). Including the electromagnetic field, the action reads

$$S = \int \mathrm{d}^4x\,\sqrt{g}\Bigg[+ \frac{1}{16\pi G}(R - 2\Lambda) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{m^2}{2}\phi^2 + \underbrace{\cdots}_{\text{other fields}}\Bigg]. \tag{8.37}$$

8.3 Adding the electromagnetic field

Let us now consider the action of the electromagnetic field coupled to gravity,

$$\begin{aligned} S &= \int \mathrm{d}^4x\sqrt{g}\left(\frac{R - 2\Lambda}{16\pi G} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\right) \\ &= \int \mathrm{d}^4x\sqrt{g}\left(\frac{R - 2\Lambda}{16\pi G} - \frac{1}{4}F_{\rho\sigma}F_{\mu\nu}g^{\mu\rho}g^{\nu\sigma}\right), \end{aligned} \tag{8.38}$$

where

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = A_{\nu;\mu} - A_{\mu;\nu}. \tag{8.39}$$

Remark. With the convention $E_i = F^{0i}$ of Eq. (4.15), this definition requires $A^\mu = -(\phi_{\mathrm{el}}, \vec{A})$ in terms of the usual scalar and vector potentials ($\vec{E} = -\vec{\nabla}\phi_{\mathrm{el}} - \partial_t\vec{A}$); with $A^\mu = (\phi_{\mathrm{el}}, \vec{A})$ one would have $F^{0i} = -E_i$. The energy-momentum tensor below is quadratic in F and does not depend on this choice. Coupling to a current through $\sqrt{g}\,J^\mu A_\mu$ in the Lagrangian, the variation with respect to A_μ gives $F^{\mu\nu}{}_{;\mu} = -J^\nu$, Eq. (4.21). Under the variation $g_{\mu\nu} \to g_{\mu\nu} + \delta g_{\mu\nu}$ of the metric, we find

$$\delta S = \int \mathrm{d}^4x\,\left[-\sqrt{g}\delta g_{\alpha\beta}\left(R^{\alpha\beta} - \frac{g^{\alpha\beta}}{2}R + \Lambda g^{\alpha\beta}\right)\frac{1}{16\pi G} - \frac{1}{4}F_{\rho\sigma}F_{\mu\nu}\delta(\sqrt{g}g^{\rho\mu}g^{\nu\sigma})\right], \tag{8.40}$$

where, using $\delta g^{\rho\mu} = -g^{\rho\alpha}g^{\mu\beta}\delta g_{\alpha\beta}$,

$$\begin{aligned} \delta(\sqrt{g}g^{\rho\mu}g^{\sigma\nu}) &= (\delta\sqrt{g})g^{\rho\mu}g^{\sigma\nu} + \sqrt{g}(\delta g^{\rho\mu})g^{\sigma\nu} + \sqrt{g}g^{\rho\mu}\delta g^{\sigma\nu} \\ &= \sqrt{g}\left(\frac{1}{2}g^{\alpha\beta}g^{\rho\mu}g^{\sigma\nu} - g^{\rho\alpha}g^{\mu\beta}g^{\sigma\nu} - g^{\rho\mu}g^{\sigma\alpha}g^{\nu\beta}\right)\delta g_{\alpha\beta}. \end{aligned} \tag{8.41}$$

Thus,

$$\begin{aligned} F_{\rho\sigma}F_{\mu\nu}\,\delta\left(\sqrt{g}g^{\rho\mu}g^{\sigma\nu}\right) &= \sqrt{g}\,\delta g_{\alpha\beta}\left(\frac{1}{2}g^{\alpha\beta}F^{\mu\nu}F_{\mu\nu} - F^{\alpha}{}_{\nu}F^{\beta\nu} - F_{\mu}{}^{\alpha}F^{\mu\beta}\right) \\ &= \sqrt{g}\,\delta g_{\alpha\beta}\left(\frac{1}{2}g^{\alpha\beta}F^{\mu\nu}F_{\mu\nu} - 2F^{\alpha\sigma}F^{\beta}{}_{\sigma}\right) \\ &= \frac{1}{2}\sqrt{g}\,\delta g_{\mu\nu}\left(g^{\mu\nu}F^{\alpha\beta}F_{\alpha\beta} + 4F^{\mu}{}_{\rho}F^{\rho\nu}\right), \end{aligned} \tag{8.42}$$

and the variation of the action is

$$\delta S = -\frac{1}{2}\int \mathrm{d}^4x\sqrt{g}\delta g_{\mu\nu}\left[\left(R^{\mu\nu} - \frac{g^{\mu\nu}}{2}R + \Lambda g^{\mu\nu}\right)\frac{1}{8\pi G} + T^{\mu\nu}_{\mathrm{EM}}\right], \tag{8.43}$$

with

$$T^{\mu\nu}_{\mathrm{EM}} = F^{\mu}{}_{\rho} F^{\rho\nu} + \frac{g^{\mu\nu}}{4} F^{\alpha\beta} F_{\alpha\beta}. \tag{8.44}$$

For $g_{\mu\nu} = \eta_{\mu\nu}$ this is indeed the energy-momentum tensor of the electromagnetic field (Exercise). Once again, we find Einstein's equations,

$$R^{\mu\nu} - \frac{g^{\mu\nu}}{2} R + \Lambda g^{\mu\nu} + 8\pi G T^{\mu\nu}_{\mathrm{EM}} = 0. \tag{8.45}$$

Exercise: Derive Maxwell's equations in the presence of the gravitational field by the variation of the action under $A_\mu \to A_\mu + \delta A_\mu$.

9

Symmetries

The gravitational action

$$S = \int \mathrm{d}^4x\sqrt{g}\left(+\frac{R}{16\pi G} + \mathcal{L}_M\right), \tag{9.1}$$

is invariant under general coordinate transformations $x \to x' = x'(x)$. Often there are coordinate transformations which preserve not only the action but also the form of the metric. Such symmetries are called isometries. They are defined as the coordinate transformations

$$x \to x' = x'(x) \quad \text{with} \quad g'_{\mu\nu}(x') = g_{\mu\nu}(x'). \tag{9.2}$$

From

$$\begin{aligned} g'_{\mu\nu}(x') &= \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x^\sigma}{\partial x'^\nu}g_{\rho\sigma}(x) \\ \Leftrightarrow\ g_{\mu\nu}(x) &= \frac{\partial x'^\rho}{\partial x^\mu}\frac{\partial x'^\sigma}{\partial x^\nu}g'_{\rho\sigma}(x'), \end{aligned} \tag{9.3}$$

we find for an isometry

$$g_{\mu\nu}(x) = \frac{\partial x'^\rho}{\partial x^\mu}\frac{\partial x'^\sigma}{\partial x^\nu}g_{\rho\sigma}(x'). \tag{9.4}$$

Let us now focus on infinitesimal isometries,

$$x^\mu \to x'^\mu = x^\mu + \varepsilon\xi^\mu, \quad \varepsilon \text{ small}. \tag{9.5}$$

This is sufficient for continuous symmetries, as finite transformations can be built from many successive small ones. Substituting the above into (9.4) and expanding to first order in ε, we have

$$\begin{aligned} g_{\mu\nu}(x) &= \left(\delta^\rho{}_\mu + \varepsilon\frac{\partial\xi^\rho}{\partial x^\mu}\right)\left(\delta^\sigma{}_\nu + \varepsilon\frac{\partial\xi^\sigma}{\partial x^\nu}\right)g_{\rho\sigma}(x+\varepsilon\xi) \\ &= \left(\delta^\rho{}_\mu + \varepsilon\frac{\partial\xi^\rho}{\partial x^\mu}\right)\left(\delta^\sigma{}_\nu + \varepsilon\frac{\partial\xi^\sigma}{\partial x^\nu}\right)\left(g_{\rho\sigma}(x) + \varepsilon\xi^\alpha\frac{\partial g_{\rho\sigma}(x)}{\partial x^\alpha} + \cdots\right) \\ &= g_{\mu\nu}(x) + \varepsilon\left(g_{\mu\sigma}(x)\frac{\partial\xi^\sigma}{\partial x^\nu} + g_{\rho\nu}(x)\frac{\partial\xi^\rho}{\partial x^\mu} + \xi^\alpha\frac{\partial g_{\mu\nu}}{\partial x^\alpha}\right) + O(\varepsilon^2). \end{aligned} \tag{9.6}$$

The above is satisfied trivially at $O(\varepsilon^0)$. Requiring that it also holds at $O(\varepsilon)$, we obtain the following condition for infinitesimal isometries,

$$\begin{aligned} g_{\mu\sigma}\frac{\partial \xi^\sigma}{\partial x^\nu} + g_{\rho\nu}\frac{\partial \xi^\rho}{\partial x^\mu} + \xi^\alpha \frac{\partial g_{\mu\nu}}{\partial x^\alpha} &= 0 \\ \Rightarrow \quad \frac{\partial}{\partial x^\nu}\left(g_{\mu\sigma}\xi^\sigma\right) + \frac{\partial}{\partial x^\mu}\left(g_{\rho\nu}\xi^\rho\right) - \xi^\sigma\frac{\partial g_{\mu\sigma}}{\partial x^\nu} - \xi^\rho\frac{\partial g_{\nu\rho}}{\partial x^\mu} + \xi^\alpha\frac{\partial g_{\mu\nu}}{\partial x^\alpha} &= 0 \\ \frac{\partial \xi_\mu}{\partial x^\nu} + \frac{\partial \xi_\nu}{\partial x^\mu} - \xi^\alpha\left(\frac{\partial g_{\mu\alpha}}{\partial x^\nu} + \frac{\partial g_{\nu\alpha}}{\partial x^\mu} - \frac{\partial g_{\mu\nu}}{\partial x^\alpha}\right) &= 0 \\ \Rightarrow \quad \frac{\partial \xi_\mu}{\partial x^\nu} + \frac{\partial \xi_\nu}{\partial x^\mu} - 2\xi_\beta \Gamma^\beta{}_{\mu\nu} &= 0 \\ \left(\frac{\partial \xi_\mu}{\partial x^\nu} - \xi_\beta\Gamma^\beta{}_{\mu\nu}\right) + \left(\frac{\partial \xi_\nu}{\partial x^\mu} - \xi_\beta\Gamma^\beta{}_{\mu\nu}\right) &= 0. \end{aligned} \tag{9.7}$$

This leads to the condition

$$\boxed{\xi_{\mu;\nu} + \xi_{\nu;\mu} = 0} \tag{9.8}$$

for the vector ξ_μ to generate an isometry. Vectors which satisfy this equation are called Killing vectors.

Let us now look at the second derivative of the Killing vector. We start with the commutator

$$\xi_{\mu;\alpha;\beta} - \xi_{\mu;\beta;\alpha} = -R^\rho{}_{\mu\alpha\beta}\xi_\rho. \tag{9.9}$$

Relabelling indices, we also have

$$\begin{aligned} \xi_{\alpha;\beta;\mu} - \xi_{\alpha;\mu;\beta} &= -R^\rho{}_{\alpha\beta\mu}\xi_\rho \\ \xi_{\beta;\mu;\alpha} - \xi_{\beta;\alpha;\mu} &= -R^\rho{}_{\beta\mu\alpha}\xi_\rho. \end{aligned} \tag{9.10}$$

Adding the above three equations, we have

$$\xi_{\mu;\alpha;\beta} - \xi_{\mu;\beta;\alpha} + \xi_{\alpha;\beta;\mu} - \xi_{\alpha;\mu;\beta} + \xi_{\beta;\mu;\alpha} - \xi_{\beta;\alpha;\mu} = -\xi_\rho \underbrace{\left(R^\rho{}_{\mu\alpha\beta} + R^\rho{}_{\alpha\beta\mu} + R^\rho{}_{\beta\mu\alpha}\right)}_{=0,\ \text{cyclicity}} = 0. \tag{9.11}$$

This equation is valid for any vector. For a Killing vector satisfying (9.8) it simplifies to

$$2\xi_{\mu;\alpha;\beta} - 2\xi_{\mu;\beta;\alpha} - 2\xi_{\beta;\alpha;\mu} = 0, \tag{9.12}$$

and thus

$$\xi_{\beta;\alpha;\mu} = \xi_{\mu;\alpha;\beta} - \xi_{\mu;\beta;\alpha}. \tag{9.13}$$

Substituting the right-hand side of the above into (9.9), we have

$$\xi_{\beta;\alpha;\mu} = -R^\rho{}_{\mu\alpha\beta}\xi_\rho. \tag{9.14}$$

This expresses the second derivative of the Killing vector linearly in terms of the vector itself; hence all higher derivatives at a point are fixed by ξ_μ and $\xi_{\mu;\nu}$ there. To see the consequences of this, consider a function $f(x)$ with

$$f''(x) = \alpha(x)f'(x) + \beta(x)f(x), \tag{9.15}$$

which is the structure (stripped of the indices etc.) of (9.14). If $f(x)$ is a smooth, well-behaved function, we can expand it around $x = X$ as an infinite Taylor series. We have

$$\begin{aligned} f(x) &= f(X) + f'(X)(x-X) + f''(X)\frac{(x-X)^2}{2!} + \big(f''(X)\big)'\frac{(x-X)^3}{3!} + \cdots \\ &= f(X) + f'(X)(x-X) + \big(\alpha(X)f'(X) + \beta(X)f(X)\big)\frac{(x-X)^2}{2!} \\ &\qquad + \frac{\mathrm{d}}{\mathrm{d}x}\big(\alpha(x)f'(x) + \beta(x)f(x)\big)\Big|_{x=X}\frac{(x-X)^3}{3!} + \cdots . \end{aligned} \tag{9.16}$$

From the above we see that the only input needed to determine $f(x)$ at any point x is the value of the function and of its first derivative at one point X. The complete function must therefore be of the form

$$f(x) = \underbrace{f(X)}_{\text{const}}\, w(x) + \underbrace{f'(X)}_{\text{const}}\, y(x). \tag{9.17}$$

We can now ask how many linearly independent functions $f^{(n)}(x)$ can satisfy the differential equation (9.15). This number is at most 2, the number of coefficient functions $w(x)$, $y(x)$. Indeed, assume that there is a third function satisfying (9.15) and (9.17). For the first two functions we can write

$$\begin{aligned} \begin{pmatrix} f^{(1)}(x) \\ f^{(2)}(x) \end{pmatrix} &= \underbrace{\begin{pmatrix} f^{(1)}(X) & f^{(1)\prime}(X) \\ f^{(2)}(X) & f^{(2)\prime}(X) \end{pmatrix}}_{=\mathrm{C},\ \text{constants}} \begin{pmatrix} w(x) \\ y(x) \end{pmatrix} \\ \Rightarrow \begin{pmatrix} w(x) \\ y(x) \end{pmatrix} &= C^{-1} \begin{pmatrix} f^{(1)}(x) \\ f^{(2)}(x) \end{pmatrix}. \end{aligned} \tag{9.18}$$

Substituting, (9.17) can be written as

$$f(x) = c_1 f^{(1)}(x) + c_2 f^{(2)}(x). \tag{9.19}$$

Thus, any third solution $f^{(3)}(x)$ is a linear combination of the other two. Let us now apply the same arguments to Killing vectors, which satisfy (9.14). A Killing vector can be expressed in the form

$$\xi_\mu(x) = \alpha^\rho{}_\mu(x)\xi_\rho(X) + \beta^{\nu\rho}{}_\mu(x)\xi_{\nu;\rho}(X). \tag{9.20}$$

Due to the antisymmetry $\xi_{\nu;\rho} = -\xi_{\rho;\nu}$ the number of independent $\beta^{\nu\rho}{}_\mu(x)$ functions in D dimensions is $\frac{D(D-1)}{2}$. In addition, we require D independent functions $\alpha^\rho{}_\mu$. In total, the functional form of $\xi_\mu(x)$ depends on $\frac{D(D-1)}{2} + D = \frac{D(D+1)}{2}$ coefficient functions. Therefore, in $D = 4$ dimensions one can have at most $\frac{D(D+1)}{2} = 10$ linearly independent Killing vectors.

9.1 Homogeneity and isotropy

1: A metric space is homogeneous if there are isometries which take any given point X to any other point in its vicinity,

$$X_\mu \to X'_\mu = X_\mu + \varepsilon_\mu. \tag{9.21}$$

The generic form of the transformation is given through the Killing vectors as

$$x_\mu \to x'_\mu = x_\mu + \varepsilon_{(n)} \xi^{(n)}_\mu(x, X). \tag{9.22}$$

To satisfy (9.21) we can choose D Killing vectors with

$$\xi^{(n)}_\mu(X, X) = \delta^n{}_\mu. \tag{9.23}$$

These Killing vectors are linearly independent. If they were not, there would exist a non-trivial relation

$$\begin{aligned} & c_n \xi^{(n)}_\mu(x, X) = 0 \\ \Rightarrow\ & c_n \xi^{(n)}_\mu(X, X) = 0 \\ & c_n \delta^n{}_\mu = 0 \ \Rightarrow\ c_\mu = 0. \end{aligned} \tag{9.24}$$

The above shows that such a relation cannot exist.

2: A metric space is isotropic about a point X if an isometry transformation

$$x_\mu \to x'_\mu = x_\mu + \varepsilon \xi_\mu(x, X), \tag{9.25}$$

keeps X fixed, $X'_\mu = X_\mu$, while it moves the other points in its vicinity. In general,

$$\xi_\mu(x, X) = \alpha^\rho{}_\mu(x, X) \xi_\rho(X, X) + \beta^{\nu\rho}{}_\mu(x, X) \xi_{\nu;\rho}(X, X), \tag{9.26}$$

with

$$\alpha^\rho{}_\mu(X, X) = \delta^\rho{}_\mu, \quad \beta^{\nu\rho}{}_\mu(X, X) = 0. \tag{9.27}$$

From $\xi_\mu(X, X) = 0$ it follows that

$$\xi_\mu(x, X) = \underbrace{\beta^{\nu\rho}{}_\mu(x, X)}_{\text{antisymmetric}} \xi_{\nu;\rho}(X, X). \tag{9.28}$$

We can choose $\frac{D(D-1)}{2} = 6$ Killing vectors such that

$$\xi^{(n,m)}_\mu(X, X) = 0, \tag{9.29}$$

and

$$\xi^{(n,m)}_{\nu;\rho}(X, X) = \delta^n{}_\nu \delta^m{}_\rho - \delta^n{}_\rho \delta^m{}_\nu. \tag{9.30}$$

Then, indeed, X is not shifted. These Killing vectors are antisymmetric in (n, m),

$$\xi^{(n,m)}_\mu(x, X) = \beta^{\nu\rho}{}_\mu(x, X) \xi^{(n,m)}_{\nu;\rho}(X, X) = \beta^{nm}{}_\mu - \beta^{mn}{}_\mu = -\xi^{(m,n)}_\mu(x, X). \tag{9.31}$$

They are also linearly independent: from

$$\begin{aligned} & c_{nm} \xi^{(n,m)}_\mu(x, X) = 0 \quad (c_{nm} = -c_{mn}) \\ \Rightarrow\ & c_{nm} \xi^{(n,m)}_{\mu;\nu}(x, X) = 0 \\ & c_{nm} \left(\delta^n{}_\mu \delta^m{}_\nu - \delta^n{}_\nu \delta^m{}_\mu \right) = 0 \\ \Rightarrow\ & c_{\mu\nu} - c_{\nu\mu} = 2 c_{\mu\nu} = 0 \ \Rightarrow\ c_{\mu\nu} = 0. \end{aligned} \tag{9.32}$$

3: A metric space which is isotropic about every point.

Let us choose two points X and $X + \mathrm{d}X$. There are Killing vectors $\xi_\lambda^{(\mu\nu)}(x, X)$ and $\xi_\lambda^{(\mu\nu)}(x, X + \mathrm{d}X)$ which satisfy

$$\begin{aligned}
\xi_\lambda^{(\mu\nu)}(X, X) &= \xi_\lambda^{(\mu\nu)}(X + \mathrm{d}X, X + \mathrm{d}X) = 0 \\
\xi_\lambda^{(\mu\nu)}(x, X) &= -\xi_\lambda^{(\nu\mu)}(x, X) \\
\xi_{\lambda;\rho}^{(\mu\nu)} &= \left[\frac{\partial}{\partial x^\rho}\xi_\lambda^{(\mu\nu)}(x, X) + \Gamma\cancelto{0}{\xi(x, X)}\right]_{x=X} \\
&= \left[\frac{\partial}{\partial x^\rho}\xi_\lambda^{(\mu\nu)}(x, X)\right]_{x=X} = \delta^\mu{}_\lambda\delta^\nu{}_\rho - \delta^\mu{}_\rho\delta^\nu{}_\lambda,
\end{aligned} \tag{9.33}$$

and exactly the same result for $X + \mathrm{d}X$. The difference of two Killing vectors is also a Killing vector. Thus,

$$\xi_\lambda^{(\mu\nu)}(x, X + \mathrm{d}X) - \xi_\lambda^{(\mu\nu)}(x, X) = \frac{\partial\xi_\lambda^{(\mu\nu)}(x, X)}{\partial X^\rho}\mathrm{d}X^\rho, \tag{9.34}$$

is a Killing vector. As $\mathrm{d}X^\rho$ is arbitrary, $\frac{\partial\xi_\lambda^{(\mu\nu)}(x,X)}{\partial X^\rho}$ is a Killing vector of the metric. We now need to evaluate it at $x = X$. To this end, we note that

$$\begin{aligned}
\xi_\lambda^{(\mu\nu)}(X, X) = 0 \\
\Rightarrow \frac{\partial}{\partial X^\rho}\xi_\lambda^{(\mu\nu)}(X, X) = 0 \\
\Rightarrow \left[\frac{\partial}{\partial x^\rho}\xi_\lambda^{(\mu\nu)}(x, X)\right]_{x=X} + \left[\frac{\partial}{\partial X^\rho}\xi_\lambda^{(\mu\nu)}(x, X)\right]_{x=X} = 0 \\
\left[\frac{\partial}{\partial X^\rho}\xi_\lambda^{(\mu\nu)}(x, X)\right]_{x=X} = -\left.\xi_{\lambda;\rho}^{(\mu\nu)}(x, X)\right|_{x=X} = -(\delta^\mu{}_\lambda\delta^\nu{}_\rho - \delta^\mu{}_\rho\delta^\nu{}_\lambda).
\end{aligned} \tag{9.35}$$

We now have a few too many Killing vectors at our disposal, so not all of them can be linearly independent. Let us look at the subset obtained by setting $\mu = \rho$ and summing:

$$\begin{aligned}
\left[\frac{\partial}{\partial X^\rho}\xi_\lambda^{(\rho\nu)}(x, X)\right]_{x=X} &= -(\delta^\rho{}_\lambda\delta^\nu{}_\rho - \delta^\rho{}_\rho\delta^\nu{}_\lambda) = (D-1)\delta^\nu{}_\lambda \\
\Rightarrow \frac{1}{D-1}\frac{\partial}{\partial X^\rho}\xi_\lambda^{(\rho\nu)}(x, X)\Big|_{x=X} &= \delta^\nu{}_\lambda.
\end{aligned} \tag{9.36}$$

At $x = X$, this has the properties of the Killing vectors of a homogeneous space. Our set of Killing vectors can thus be chosen as

$$\xi_\lambda^{(\mu\nu)}(x, X), \tag{9.37}$$

with $\xi_{\lambda;\rho}^{(\mu\nu)}(X, X) = -\xi_{\lambda;\rho}^{(\nu\mu)}(X, X) = \delta^\mu{}_\lambda\delta^\nu{}_\rho - \delta^\mu{}_\rho\delta^\nu{}_\lambda$ and

$$\xi_\lambda^{(\mu)}(x, X) = \frac{1}{D-1}\frac{\partial}{\partial X^\rho}\xi_\lambda^{(\rho\mu)}(x, X), \tag{9.38}$$

with $\xi_\lambda^{(\mu)}(X, X) = \delta^\mu{}_\lambda$.

This set saturates the maximum number of independent Killing vectors. Is it truly independent? Yes. Consider a linear combination

$$c_\mu \xi^{(\mu)}_\lambda(x,X) + c_{\mu\nu}\xi^{(\mu\nu)}_\lambda(x,X) = 0. \tag{9.39}$$

For $x = X$ we have $\xi^{(\mu\nu)}_\lambda(X,X) = 0$ and then

$$c_\mu \xi^{(\mu)}_\lambda(X,X) = 0 \;\Rightarrow\; c_\mu \delta^\mu{}_\lambda = 0 \;\Rightarrow\; c_\lambda = 0. \tag{9.40}$$

Differentiating with respect to x, we have

$$\underbrace{c_\mu}_{=0}\; \xi^{(\mu)}_{\lambda;\rho}(x,X)\Big|_{x=X} + c_{\mu\nu}\xi^{(\mu\nu)}_{\lambda;\rho}(X,X) = 0 \;\Rightarrow\; \cdots \;\Rightarrow\; c_{\mu\nu} = 0. \tag{9.41}$$

A metric space with $\frac{D(D+1)}{2}$ independent Killing vectors is called maximally symmetric. We have just shown that a space which is isotropic about every point, and thus also homogeneous, is maximally symmetric. We now show the converse: a maximally symmetric space is necessarily homogeneous and isotropic about any point.

Proof: For concreteness, let us take $D = 2$ dimensions and a metric space with $n = 1, \cdots, \frac{D(D+1)}{2} = 3$ independent Killing vectors, $\xi^1_\mu(x)$, $\xi^2_\mu(x)$, $\xi^3_\mu(x)$. They of course take the form

$$\xi^{(n)}_\mu(x) = \alpha^\rho{}_\mu(x)\xi^{(n)}_\rho(X) + \beta^{\rho\nu}{}_\mu(x)\xi^{(n)}_{\rho;\nu}(X), \tag{9.42}$$

where $\xi^{(n)}_{\rho;\nu}(X) = -\xi^{(n)}_{\nu;\rho}(X)$. Let us organise the constants $\xi^{(n)}_\rho(X)$ and $\xi^{(n)}_{\rho;\nu}(X)$ into a matrix

$$M = \begin{pmatrix} \xi^{(1)}_1(X) & \xi^{(1)}_2(X) & \xi^{(1)}_{1;2}(X) \\ \xi^{(2)}_1(X) & \xi^{(2)}_2(X) & \xi^{(2)}_{1;2}(X) \\ \xi^{(3)}_1(X) & \xi^{(3)}_2(X) & \xi^{(3)}_{1;2}(X) \end{pmatrix}. \tag{9.43}$$

We must have $\det M \neq 0$. If $\det M = 0$, then the system of equations

$$\vec{c}\cdot M = 0, \tag{9.44}$$

has a solution with

$$\vec{c} = (c_1, c_2, c_3) \neq 0. \tag{9.45}$$

In other words, the equations

$$\begin{aligned} c_1\xi^{(1)}_1(X) + c_2\xi^{(2)}_1(X) + c_3\xi^{(3)}_1(X) &= 0 \\ c_1\xi^{(1)}_2(X) + c_2\xi^{(2)}_2(X) + c_3\xi^{(3)}_2(X) &= 0 \\ c_1\xi^{(1)}_{1;2}(X) + c_2\xi^{(2)}_{1;2}(X) + c_3\xi^{(3)}_{1;2}(X) &= 0, \end{aligned} \tag{9.46}$$

or, more compactly,

$$\sum_n c_n \xi^{(n)}_\rho(X) = 0, \quad \sum_n c_n \xi^{(n)}_{\rho;\nu}(X) = 0, \tag{9.47}$$

are satisfied for some $\vec{c} = \{c_n\} \neq 0$. But then, for the same values of $\{c_n\}$ we will have

$$\sum_n c_n \xi^{(n)}_\mu(x) = \alpha^\rho{}_\mu(x)\underbrace{\left(\sum_n c_n \xi^{(n)}_\rho(X)\right)}_{=0} + \beta^{\rho\nu}{}_\mu(x)\underbrace{\left(\sum_n c_n \xi^{(n)}_{\rho;\nu}(X)\right)}_{=0} = 0, \tag{9.48}$$

and the Killing vectors $\xi_\mu^{(n)}(x)$ would be linearly dependent, contrary to our assumption. Given that $\det M \neq 0$, the matrix M has an inverse and the equation

$$\vec{D} \cdot M = \vec{A}, \tag{9.49}$$

with $\vec{D} = (d_1, d_2, d_3)$, $\vec{A} = (q_1, q_2, b_{12})$, has a solution $\vec{D} = \vec{A} \cdot M^{-1}$ for any $\vec{A}$. Then, the superposition

$$\begin{aligned}\tilde{\xi}_\mu(x) &= \sum_n d_n \xi_\mu^{(n)}(x) = \alpha^\rho{}_\mu(x)\left(\sum_n d_n \xi_\rho^{(n)}(X)\right) + \beta^{\rho\nu}{}_\mu(x)\left(\sum_n d_n \xi_{\rho;\nu}^{(n)}(X)\right) \\ &= \alpha^\rho{}_\mu(x) q_\rho + \beta^{\rho\nu}{}_\mu(x) b_{\rho\nu},\end{aligned} \tag{9.50}$$

has

$$\begin{aligned}\tilde{\xi}_\rho(X) &= q_\rho \quad \text{arbitrary (homogeneous space)} \\ \tilde{\xi}_{\rho;\nu}(X) &= b_{\rho\nu} = -b_{\nu\rho} \quad \text{arbitrary (isotropic space).}\end{aligned}$$

Therefore, a maximally symmetric space is both homogeneous and isotropic about any point X.

Example: Consider a flat spacetime, for which the Riemann tensor vanishes. In Minkowski coordinates, a Killing vector then satisfies

$$\begin{aligned}\xi_{\mu;\alpha;\beta} &= -\underbrace{R^\rho{}_{\beta\alpha\mu}}_{=0}\xi_\rho \\ \Rightarrow \frac{\partial^2 \xi_\mu}{\partial x^\alpha \partial x^\beta} &= 0 \Rightarrow \xi_\mu(x) = a_\mu + b_{\mu\rho}x^\rho \quad \text{(linear in the coordinates).}\end{aligned} \tag{9.51}$$

From $\xi_{\mu;\nu} = -\xi_{\nu;\mu}$ we have $b_{\mu\nu} = -b_{\nu\mu}$. We can then define $\frac{D(D+1)}{2}$ independent Killing vectors as

$$\xi_\mu^{(\nu)}(x) = \delta^\nu{}_\mu, \tag{9.52}$$

and

$$\xi_\mu^{(\nu\rho)}(x) = \frac{1}{2}\left(\delta^\nu{}_\mu x^\rho - \delta^\rho{}_\mu x^\nu\right). \tag{9.53}$$

A generic Killing vector is decomposed as

$$\begin{aligned}\xi_\mu(x) &= a_\nu \xi_\mu^{(\nu)} + b_{\nu\rho}\xi_\mu^{(\nu\rho)}(x) \\ &= a_\nu \delta^\nu{}_\mu + \frac{b_{\nu\rho}}{2}\left(\delta^\nu{}_\mu x^\rho - \delta^\rho{}_\mu x^\nu\right) \\ &= a_\mu + \frac{1}{2}\left(b_{\mu\rho}x^\rho - b_{\nu\mu}x^\nu\right) \\ &= a_\mu + \frac{b_{\mu\rho} - b_{\rho\mu}}{2}x^\rho.\end{aligned} \tag{9.54}$$

The space is homogeneous and isotropic about every point.

Returning to our general discussion, we remark that the number of Killing vectors does not depend on the choice of coordinates. If we have Killing vectors $\xi_\mu^{(n)}$, $n = 1, \ldots, M$ in one coordinate system, we will have M Killing vectors

$$\xi_\mu'^{(n)} \equiv \frac{\partial x^\rho}{\partial x'^\mu}\xi_\rho^{(n)}, \tag{9.55}$$

under a transformation $x \to x'$.

9.2 Curvature of maximally symmetric spaces

Let us consider the covariant derivative of a Killing vector, $\xi_{\mu;\nu}$, which, like every second-rank tensor, satisfies

$$\left(\xi_{\mu;\nu}\right)_{;\rho;\sigma} - \left(\xi_{\mu;\nu}\right)_{;\sigma;\rho} = -R^{\lambda}{}_{\mu\rho\sigma}\xi_{\lambda;\nu} - R^{\lambda}{}_{\nu\rho\sigma}\xi_{\mu;\lambda}. \tag{9.56}$$

Recall also that

$$\xi_{\alpha;\beta;\gamma} = -R^{\lambda}{}_{\gamma\beta\alpha}\xi_{\lambda}. \tag{9.57}$$

Then (9.56) becomes

$$\begin{aligned}
&\left(-R^{\lambda}{}_{\rho\nu\mu}\xi_{\lambda}\right)_{;\sigma} - \left(-R^{\lambda}{}_{\sigma\nu\mu}\xi_{\lambda}\right)_{;\rho} = -R^{\lambda}{}_{\mu\rho\sigma}\xi_{\lambda;\nu} - R^{\lambda}{}_{\nu\rho\sigma}\xi_{\mu;\lambda} \\
&\Rightarrow \left(R^{\lambda}{}_{\sigma\nu\mu;\rho} - R^{\lambda}{}_{\rho\nu\mu;\sigma}\right)\xi_{\lambda} = R^{\lambda}{}_{\rho\nu\mu}\xi_{\lambda;\sigma} - R^{\lambda}{}_{\sigma\nu\mu}\xi_{\lambda;\rho} - R^{\lambda}{}_{\mu\rho\sigma}\xi_{\lambda;\nu} + R^{\lambda}{}_{\nu\rho\sigma}\xi_{\lambda;\mu} \\
&\left(R^{\lambda}{}_{\sigma\nu\mu;\rho} - R^{\lambda}{}_{\rho\nu\mu;\sigma}\right)\xi_{\lambda} = Q^{\lambda}{}_{\mu\nu\rho\sigma}{}^{\alpha}\xi_{\lambda;\alpha}(x),
\end{aligned} \tag{9.58}$$

with

$$Q^{\lambda}{}_{\mu\nu\rho\sigma}{}^{\alpha} \equiv +R^{\lambda}{}_{\nu\rho\sigma}\delta^{\alpha}{}_{\mu} + R^{\lambda}{}_{\rho\nu\mu}\delta^{\alpha}{}_{\sigma} - R^{\lambda}{}_{\sigma\nu\mu}\delta^{\alpha}{}_{\rho} - R^{\lambda}{}_{\mu\rho\sigma}\delta^{\alpha}{}_{\nu}. \tag{9.59}$$

Equation (9.58) holds for every Killing vector at any point x of our space. At any point x, there exist Killing vectors with $\xi_{\lambda}(x) = 0$ and $\xi_{\lambda;\mu}(x) = -\xi_{\mu;\lambda}(x)$. For them only the right-hand side of (9.58) survives. As $\xi_{\lambda;\mu}(x)$ is arbitrary and antisymmetric, the part of the coefficient which is antisymmetric in $\lambda \leftrightarrow \alpha$ must vanish,

$$Q^{\lambda}{}_{\mu\nu\rho\sigma}{}^{\alpha} = Q^{\alpha}{}_{\mu\nu\rho\sigma}{}^{\lambda}. \tag{9.60}$$

Also, at every point x, we can find Killing vectors with $\xi_{\lambda}(x)$ arbitrary and $\xi_{\lambda;\mu}(x) = 0$ (homogeneity). For these, we must also have

$$R^{\lambda}{}_{\sigma\nu\mu;\rho} - R^{\lambda}{}_{\rho\nu\mu;\sigma} = 0. \tag{9.61}$$

Contracting α with σ in the symmetry condition for Q gives

$$Q^{\lambda}{}_{\mu\nu\rho\sigma}{}^{\sigma} = Q^{\sigma}{}_{\mu\nu\rho\sigma}{}^{\lambda}. \tag{9.62}$$

Performing the contractions and exploiting the antisymmetry properties of the Riemann tensor, we arrive at

$$(D-1)R_{\lambda\rho\sigma\nu} = R_{\rho\nu}g_{\lambda\sigma} - R_{\sigma\rho}g_{\lambda\nu}. \tag{9.63}$$

But,

$$\begin{aligned}
R_{\lambda\rho\sigma\nu} &= -R_{\rho\lambda\sigma\nu} \\
R_{\rho\nu}g_{\lambda\sigma} - R_{\sigma\rho}g_{\lambda\nu} &= -R_{\lambda\nu}g_{\rho\sigma} + R_{\lambda\sigma}g_{\rho\nu} \\
\overset{\times g^{\lambda\sigma}}{\Rightarrow}\quad R_{\rho\nu}D - R_{\nu\rho} &= -R_{\rho\nu} + g_{\rho\nu}R \\
\Rightarrow\quad R_{\rho\nu} &= \frac{g_{\rho\nu}}{D}R.
\end{aligned} \tag{9.64}$$

Then (9.63) becomes

$$\begin{aligned}
(D-1)R_{\lambda\rho\sigma\nu} &= \frac{R}{D}\left[g_{\rho\nu}g_{\lambda\sigma} - g_{\rho\sigma}g_{\lambda\nu}\right] \\
\Rightarrow\quad R_{\mu\nu\rho\sigma} &= \frac{g_{\mu\rho}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\rho}}{D(D-1)}R.
\end{aligned} \tag{9.65}$$

From the Bianchi identity, we find

$$\begin{aligned} \left(R^{\mu\nu} - \frac{g^{\mu\nu}}{2}R\right)_{;\mu} &= 0 \\ \left(\frac{g^{\mu\nu}}{D} - \frac{g^{\mu\nu}}{2}\right)R_{;\mu} &= 0 \\ \Rightarrow\ (D-2)\frac{\partial R}{\partial x^\mu} &= 0. \end{aligned} \tag{9.66}$$

For $D \neq 2$ it follows that $\frac{\partial R}{\partial x^\alpha} = 0$ and thus

$$R = \text{const} \equiv -D(D-1)\kappa. \tag{9.67}$$

The constant κ is known as the curvature constant. In terms of it, the Riemann tensor is

$$R_{\alpha\beta\gamma\delta} = +\kappa\left(g_{\alpha\delta}g_{\beta\gamma} - g_{\alpha\gamma}g_{\beta\delta}\right), \tag{9.68}$$

and

$$R_{\alpha\beta} = -(D-1)\kappa g_{\alpha\beta}. \tag{9.69}$$

Theorem (without proof; see Weinberg, Ch. 13)

If two metrics, $g_{\mu\nu}(x)$ and $g'_{\mu\nu}(x')$, satisfy the above equations with the same curvature constant κ and have the same number of positive and negative eigenvalues, then they are related by a coordinate transformation

$$g'_{\mu\nu}(x') = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu}g_{\alpha\beta}(x), \tag{9.70}$$

and, thus, they are physically equivalent.

9.3 Construction of maximally symmetric spaces

Given the uniqueness of a maximally symmetric space of a given curvature κ, we may construct its metric in any way we like, even by guessing it. One possibility is to start from a flat $(D+1)$-dimensional space with the metric

$$\mathrm{d}\tau^2 = g_{AB}\mathrm{d}x^A\mathrm{d}x^B = c_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu + \frac{\mathrm{d}z^2}{\kappa}, \tag{9.71}$$

where κ and $c_{\mu\nu}$ are constants, and restrict it to the "sphere"

$$z^2 + \kappa c_{\alpha\beta}x^\alpha x^\beta = 1. \tag{9.72}$$

Notice that both of the above equations are invariant under the rigid "rotations

$$\begin{aligned} x^\mu \to x'^\mu &= R^\mu{}_\nu x^\nu + R^\mu{}_z z \\ z \to z' &= R^z{}_\mu x^\mu + R^z{}_z z, \end{aligned} \tag{9.73}$$

with

$$\begin{aligned} c_{\mu\nu}R^\mu{}_\rho R^\nu{}_\sigma + \kappa^{-1}R^z{}_\rho R^z{}_\sigma &= c_{\rho\sigma} \\ c_{\mu\nu}R^\mu{}_\rho R^\nu{}_z + \kappa^{-1}R^z{}_\rho R^z{}_z &= 0 \\ c_{\mu\nu}R^\mu{}_z R^\nu{}_z + \kappa^{-1}\left(R^z{}_z\right)^2 &= \kappa^{-1}, \end{aligned} \tag{9.74}$$

These transformations have $\frac{D(D+1)}{2}$ generators. The resulting space inherits these symmetries and thus has the maximum number of Killing vectors in D dimensions. The metric in the D-dimensional space is easily computed. We find

$$\mathrm{d}z^2 = \frac{\kappa^2 \left(c_{\mu\nu} x^\mu \mathrm{d}x^\nu\right)^2}{\left(1 - \kappa c_{\alpha\beta} x^\alpha x^\beta\right)}, \tag{9.75}$$

and thus

$$\mathrm{d}\tau^2 = c_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu + \frac{\kappa^2}{\kappa} \frac{\left(c_{\mu\alpha} c_{\nu\beta} x^\alpha x^\beta\right) \mathrm{d}x^\mu \mathrm{d}x^\nu}{\left(1 - \kappa c_{\alpha\beta} x^\alpha x^\beta\right)}, \tag{9.76}$$

which yields $\mathrm{d}\tau^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$ with

$$g_{\mu\nu}(x) = c_{\mu\nu} + \kappa \frac{c_{\mu\alpha} c_{\nu\beta} x^\alpha x^\beta}{1 - \kappa c_{\alpha\beta} x^\alpha x^\beta}. \tag{9.77}$$

Exercise: Show that the affine connection is

$$\Gamma^\mu{}_{\nu\rho} = +\kappa x^\mu g_{\nu\rho}. \tag{9.78}$$

Then, we can easily compute the Riemann tensor

$$\begin{aligned} R^\mu{}_{\nu\rho\sigma} &= \frac{\partial \Gamma^\mu{}_{\nu\rho}}{\partial x^\sigma} + \Gamma^\alpha{}_{\nu\rho}\Gamma^\mu{}_{\alpha\sigma} - \frac{\partial \Gamma^\mu{}_{\nu\sigma}}{\partial x^\rho} - \Gamma^\alpha{}_{\nu\sigma}\Gamma^\mu{}_{\alpha\rho} \\ \Rightarrow\ R_{\mu\nu\rho\sigma} &= \kappa \left(g_{\mu\sigma} g_{\rho\nu} - g_{\mu\rho} g_{\nu\sigma}\right). \end{aligned} \tag{9.79}$$

Thus, κ is indeed the curvature constant.

Killing vectors: We now identify independent Killing vectors which generate the isometries of our maximally symmetric space. The symmetry transformations include the following two classes.

A:

$$\begin{aligned} x^\mu &\to x'^\mu = R^\mu{}_\nu x^\nu \\ z &\to z' = z, \end{aligned} \tag{9.80}$$

with $R^z{}_\mu = R^\mu{}_z = 0$ and $R^z{}_z = 1$. Then,

$$R^\mu{}_\alpha R^\nu{}_\beta c_{\mu\nu} = c_{\alpha\beta}. \tag{9.81}$$

For infinitesimal transformations, we have

$$\begin{aligned} \left(\delta^\mu{}_\alpha + \underline{O}^\mu{}_\alpha\right)\left(\delta^\nu{}_\beta + \underline{O}^\nu{}_\beta\right) c_{\mu\nu} &= c_{\alpha\beta} \\ \Rightarrow\ \underline{O}^\mu{}_\alpha c_{\mu\beta} + \underline{O}^\nu{}_\beta c_{\alpha\nu} &= 0. \end{aligned} \tag{9.82}$$

We also have

$$x'^\mu = x^\mu + \underline{O}^\mu{}_\alpha x^\alpha = x^\mu + \xi^\mu_{\underline{O}}(x). \tag{9.83}$$

One can check that (9.82) is equivalent to the condition

$$\xi_{\mu;\alpha} + \xi_{\alpha;\mu} = 0, \tag{9.84}$$

of Killing vectors.

B: There are also infinitesimal "translations", with Killing vectors

$$\xi^{\mu}_{(a)} = a^{\mu}(1 - \kappa c_{\rho\sigma} x^{\rho} x^{\sigma})^{1/2}. \tag{9.85}$$

Exercise: Check that

$$\xi_{(a)\mu;\rho} + \xi_{(a)\rho;\mu} = 0. \tag{9.86}$$

We can simplify the metric (9.77) further by a linear transformation,

$$x^{\mu} \to x'^{\mu} = A^{\mu}{}_{\nu} x^{\nu}. \tag{9.87}$$

Then,

$$c_{\mu\nu} x^{\mu} x^{\nu} \to c_{\mu\nu} A^{\mu}{}_{\rho} A^{\nu}{}_{\sigma} x^{\rho} x^{\sigma} \to c'{}_{\rho\sigma} x^{\rho} x^{\sigma}, \tag{9.88}$$

with $c'{}_{\rho\sigma} = c_{\mu\nu} A^{\mu}{}_{\rho} A^{\nu}{}_{\sigma}$. We can choose the transformation matrix $A^{\mu}{}_{\rho}$ such that $c'{}_{\rho\sigma}$ is brought to a diagonal form. For a Euclidean space, all eigenvalues of the metric have the same sign, and we can write

$$c_{\mu\nu} = \lambda \delta_{\mu\nu}. \tag{9.89}$$

Then,

$$\mathrm{d}s^2 = g_{\mu\nu} \mathrm{d}x^{\mu} \mathrm{d}x^{\nu} = \lambda \mathrm{d}\vec{x}^2 + \frac{\kappa\lambda^2 \left(\vec{x} \cdot \mathrm{d}\vec{x}\right)^2}{1 - \kappa\lambda \vec{x}^2}. \tag{9.90}$$

If $\kappa \neq 0$, we can pick $\lambda = 1/|\kappa|$ and then

$$\mathrm{d}s^2 = \frac{1}{|\kappa|} \left[\mathrm{d}\vec{x}^2 + \mathrm{sgn}(\kappa) \frac{\left(\vec{x} \cdot \mathrm{d}\vec{x}\right)^2}{1 - \mathrm{sgn}(\kappa) \vec{x}^2} \right]. \tag{9.91}$$

For a space whose metric has eigenvalues with signs $(+, -, -, -)$, we can diagonalise to $c_{\mu\nu} = \eta_{\mu\nu}$, which yields

$$\mathrm{d}\tau^2 = \mathrm{d}t^2 - \mathrm{d}\vec{x}^2 + \frac{\kappa \left(t \mathrm{d}t - \vec{x} \cdot \mathrm{d}\vec{x}\right)^2}{1 - \kappa \left(t^2 - \vec{x}^2\right)}. \tag{9.92}$$

9.4 Tensors in maximally symmetric spaces

A generic tensor $T_{\mu\nu\ldots}$ is called form-invariant if

$$T_{\mu\nu\ldots}(x') = T'_{\mu\nu\ldots}(x'), \tag{9.93}$$

under a coordinate transformation $x \to x' = x'(x)$. Since $T_{\mu\nu\ldots}$ is a tensor, we have

$$T_{\mu\nu\ldots}(x) = \frac{\partial x'^{\rho}}{\partial x^{\mu}} \frac{\partial x'^{\sigma}}{\partial x^{\nu}} \cdots T'_{\rho\sigma\ldots}(x'), \tag{9.94}$$

and from form invariance we have

$$T_{\mu\nu}(x) = \frac{\partial x'^{\rho}}{\partial x^{\mu}} \frac{\partial x'^{\sigma}}{\partial x^{\nu}} T_{\rho\sigma}(x'), \tag{9.95}$$

where, for illustration, we have limited the rank to two. For infinitesimal transformations,

$$x'^{\mu} = x^{\mu} + \varepsilon \xi^{\mu}(x), \tag{9.96}$$

we obtain

$$\xi^\alpha \frac{\partial}{\partial x^\alpha} T_{\mu\nu}(x) + \frac{\partial \xi^\alpha}{\partial x^\mu} T_{\alpha\nu}(x) + \frac{\partial \xi^\alpha}{\partial x^\nu} T_{\mu\alpha}(x) = 0. \tag{9.97}$$

Specialising to a scalar $S(x)$, the above gives

$$\xi^\alpha(x) \frac{\partial}{\partial x^\alpha} S(x) = 0. \tag{9.98}$$

In a maximally symmetric space there are Killing vectors which take arbitrary values at any point x. Then,

$$\frac{\partial}{\partial x^\alpha} S(x) = 0 \;\Rightarrow\; S(x) = \text{const.} \tag{9.99}$$

In a maximally symmetric space, we can also choose Killing vectors which vanish at any given point X, $\xi^\alpha(X) = 0$, while their covariant derivatives do not vanish and form an arbitrary antisymmetric tensor,

$$\xi_{\mu;\rho}(X) = -\xi_{\rho;\mu}(X). \tag{9.100}$$

We have

$$\begin{aligned} \xi_{\sigma;\mu}(X) &= g_{\sigma\rho}(X) \left[\frac{\partial \xi^\rho}{\partial x^\mu} + \Gamma^\rho{}_{\alpha\mu} \underbrace{\xi^\alpha(x)}_{=0} \right]_{x=X} \\ \Rightarrow\; \xi_{\sigma;\mu}(X) &= g_{\sigma\rho}(X) \left(\frac{\partial \xi^\rho(x)}{\partial x^\mu} \right)_{x=X}. \end{aligned} \tag{9.101}$$

Then,

$$\begin{aligned} &\xi^\alpha(x) \frac{\partial T_{\mu\nu}}{\partial x^\alpha} + \frac{\partial \xi^\alpha}{\partial x^\mu} T_{\alpha\nu} + \frac{\partial \xi^\alpha}{\partial x^\nu} T_{\mu\alpha} = 0 \\ \overset{x=X}{\Rightarrow}\; &\frac{\partial \xi^\alpha}{\partial x^\mu} g_{\alpha\rho} T^\rho{}_\nu + \frac{\partial \xi^\alpha}{\partial x^\nu} g_{\alpha\rho} T_\mu{}^\rho = 0 \\ &\Rightarrow\; \xi_{\rho;\mu} T^\rho{}_\nu + \xi_{\rho;\nu} T_\mu{}^\rho = 0 \\ \Rightarrow\; &\underbrace{\xi_{\rho;\sigma}}_{\text{antisymm.}} \left(\delta^\sigma{}_\mu T^\rho{}_\nu + \delta^\sigma{}_\nu T_\mu{}^\rho \right) = 0. \end{aligned} \tag{9.102}$$

Because $\xi_{\rho;\sigma}$ is antisymmetric (and otherwise arbitrary), its coefficient at X must be symmetric in $\rho \leftrightarrow \sigma$. This gives

$$\delta^\sigma{}_\mu T^\rho{}_\nu + \delta^\sigma{}_\nu T_\mu{}^\rho = \delta^\rho{}_\mu T^\sigma{}_\nu + \delta^\rho{}_\nu T_\mu{}^\sigma. \tag{9.103}$$

Similarly, for a vector we have

$$\delta^\sigma{}_\mu A^\rho = \delta^\rho{}_\mu A^\sigma. \tag{9.104}$$

Contracting the above equation by setting $\rho = \mu$, we find

$$\begin{aligned} &\delta^\sigma{}_\mu A^\mu = \delta^\mu{}_\mu A^\sigma \\ &\Rightarrow\; A^\sigma = D A^\sigma \;\Rightarrow\; (D-1) A^\sigma = 0 \\ &\Rightarrow\; A^\sigma = 0 \quad \text{for } D > 1. \end{aligned} \tag{9.105}$$

Similarly, setting $\rho = \mu$ in (9.103) gives

$$\begin{aligned} & T^{\sigma}{}_{\nu} + \delta^{\sigma}{}_{\nu} T_{\mu}{}^{\mu} = D T^{\sigma}{}_{\nu} + T_{\nu}{}^{\sigma} \\ \overset{\times g_{\mu\sigma}}{\Rightarrow} \; & T_{\mu\nu} + g_{\mu\nu} T_{\alpha}{}^{\alpha} = D T_{\mu\nu} + T_{\nu\mu} \\ \Rightarrow \; & (D-1) T_{\mu\nu} + T_{\nu\mu} = g_{\mu\nu} T_{\alpha}{}^{\alpha}. \end{aligned} \tag{9.106}$$

Also,

$$(D-1) T_{\nu\mu} + T_{\mu\nu} = g_{\nu\mu} T_{\alpha}{}^{\alpha} \quad (\nu \leftrightarrow \mu). \tag{9.107}$$

Subtracting the two equations gives

$$\begin{aligned} (D-2)\left(T_{\mu\nu} - T_{\nu\mu}\right) &= 0 \\ \Rightarrow \; T_{\mu\nu} &= T_{\nu\mu} \quad \text{for } D \geq 3. \end{aligned} \tag{9.108}$$

Then,

$$\begin{aligned} D T_{\mu\nu} &= g_{\mu\nu} T^{\alpha}{}_{\alpha} \\ \Rightarrow \; T_{\mu\nu} &= \frac{g_{\mu\nu}}{D} T^{\alpha}{}_{\alpha}. \end{aligned} \tag{9.109}$$

As $T^{\alpha}{}_{\alpha}$ is a scalar, it is a constant. Therefore, the only form-invariant rank-2 tensor in a maximally symmetric space is the metric, up to a constant factor,

$$T_{\mu\nu} = \frac{g_{\mu\nu}}{D} \times \text{const.} \tag{9.110}$$

Exercise: Repeat this type of analysis for a rank-3 or rank-4 tensor.

9.5 Spaces with maximally symmetric subspaces

Often a space is not maximally symmetric, but it has subspaces which are. For example, the space of the universe is believed to be (approximately) isotropic and homogeneous at any given time t; the full spacetime, however, is not. Consider a space of dimension D with a maximally symmetric subspace of dimension M. We parametrise the subspace with M variables u^i, $i = 1, \ldots, M$, and the rest of the space with $D - M$ variables v^α, $\alpha = 1, \ldots, D - M$.

Maximal symmetry of the M-dimensional subspace means that there exist $\frac{M(M+1)}{2}$ Killing vectors which generate the transformations

$$\begin{aligned} u^i &\to u'^i = u^i + \xi^i(u, v) \\ v^\alpha &\to v'^\alpha = v^\alpha, \end{aligned} \tag{9.111}$$

which leave the metric of the full space form-invariant. Killing vectors satisfy the condition

$$\xi_{\mu;\nu} + \xi_{\nu;\mu} = 0, \tag{9.112}$$

or, in its original form,

$$g_{\mu\rho} \frac{\partial \xi^\rho}{\partial x^\nu} + g_{\rho\nu} \frac{\partial \xi^\rho}{\partial x^\mu} + \frac{\partial g_{\mu\nu}}{\partial x^\rho} \xi^\rho = 0. \tag{9.113}$$

For $\mu, \nu = i, j$ in the M-dimensional subspace we have

$$g_{ik}\frac{\partial \xi^k}{\partial u^j} + g_{jk}\frac{\partial \xi^k}{\partial u^i} + \frac{\partial g_{ij}}{\partial u^k}\xi^k = 0. \tag{9.114}$$

For $\mu = i$ in the M-dimensional subspace and $\nu = \alpha$ spanning the v-variables we have

$$g_{ik}\frac{\partial \xi^k}{\partial v^\alpha} + g_{\alpha k}\frac{\partial \xi^k}{\partial u^i} + \frac{\partial g_{\alpha i}}{\partial u^k}\xi^k = 0. \tag{9.115}$$

For $\mu = \alpha$, $\nu = \beta$ spanning the v-variables we have

$$g_{\alpha k}\frac{\partial \xi^k}{\partial v^\beta} + g_{\beta k}\frac{\partial \xi^k}{\partial v^\alpha} + \frac{\partial g_{\alpha\beta}}{\partial u^k}\xi^k = 0. \tag{9.116}$$

We can now simplify the metric by choosing new coordinates u'^i in which

$$g'_{i\alpha} = 0. \tag{9.117}$$

Indeed, under a coordinate transformation

$$\begin{aligned} u^i &\to u'^i = u'^i(u, v) \\ v^\alpha &\to v'^\alpha = v^\alpha, \end{aligned} \tag{9.118}$$

the metric transforms as

$$\begin{aligned} g'_{i\alpha}(u, v) &= \frac{\partial u^j}{\partial u'^i}\frac{\partial u^k}{\partial v'^\alpha}g_{jk} + \frac{\partial u^j}{\partial u'^i}\underbrace{\frac{\partial v^\beta}{\partial v'^\alpha}}_{=\delta^\beta{}_\alpha}g_{j\beta} + \cancelto{0}{\frac{\partial v^j}{\partial u'^i}}\frac{\partial v^\beta}{\partial v'^\alpha}g_{i\beta} \\ \Rightarrow g'_{i\alpha}(u, v) &= \frac{\partial u^j}{\partial u'^i}\left(\frac{\partial u^k}{\partial v'^\alpha}g_{jk} + g_{j\alpha}\right). \end{aligned} \tag{9.119}$$

It can be proven (see Weinberg, Sec. 13.5) that we can always find a solution of

$$\frac{\partial u^k}{\partial v'^\alpha}g_{jk} + g_{j\alpha} = 0, \tag{9.120}$$

and thus make the cross terms $g_{i\alpha}$ vanish. In these coordinates the Killing vector conditions become

$$\begin{aligned} (9.115) &\to g_{ik}\frac{\partial \xi^k}{\partial v^\alpha} = 0 \\ (9.116) &\to \frac{\partial g_{\alpha\beta}}{\partial u^k}\xi^k = 0 \\ (9.114) &\to g_{ik}\frac{\partial \xi^k}{\partial u^j} + g_{jk}\frac{\partial \xi^k}{\partial u^i} + \frac{\partial g_{ij}}{\partial u^k}\xi^k = 0. \end{aligned} \tag{9.121}$$

From the first equation we get $\frac{\partial \xi^k}{\partial v^\alpha}(u, v) = 0$ and thus $\xi^k = \xi^k(u)$ is a function of the subspace variables only.

From the second equation we get $g_{\alpha\beta} = g_{\alpha\beta}(v)$, i.e. it is independent of the subspace variables u.

Let us now consider the metric $g_{ij}(u, v_0)$ at a fixed value $v = v_0$. This is the metric of the maximally symmetric M-dimensional subspace. Since $\xi^k = \xi^k(u)$ does not depend on v, Eq. (9.114) states that, for every fixed v, $g_{ij}(u, v)$ is a form-invariant tensor under the $M(M+1)/2$ isometries of the maximally symmetric metric $g_{ij}(u, v_0)$ (Weinberg, Sec. 13.5). We have shown above that a form-invariant rank-2 tensor in a maximally symmetric space is proportional to its metric, with a constant coefficient, which here may depend on the parameter v:

$$g_{ij}(u, v) = f(v)\, g_{ij}(u, v_0). \tag{9.122}$$

Writing $\tilde{g}_{ij}(u) \equiv g_{ij}(u, v_0)$, we have

$$g_{ij}(u, v) = f(v)\tilde{g}_{ij}(u). \tag{9.123}$$

Thus, we have found the following general structure for a metric with a maximally symmetric subspace

$$\mathrm{d}\tau^2 = g_{\alpha\beta}(v)\mathrm{d}v^\alpha \mathrm{d}v^\beta + f(v)\tilde{g}_{ij}(u)\mathrm{d}u^i \mathrm{d}u^j, \tag{9.124}$$

where $\tilde{g}_{ij}(u)$ is the metric of a maximally symmetric space. If this subspace is purely spatial, we have the more specific form

$$\mathrm{d}\tau^2 = g_{\alpha\beta}(v)\mathrm{d}v^\alpha \mathrm{d}v^\beta - f(v)\left[\mathrm{d}\vec{u}^2 + \frac{\kappa\left(\vec{u}\cdot\mathrm{d}\vec{u}\right)^2}{1-\kappa\vec{u}^2}\right], \tag{9.125}$$

with $\kappa = +1, -1, 0$ after a rescaling of $\vec{u}$ and f. For a three-dimensional isotropic and homogeneous subspace of a four-dimensional spacetime, we introduce spherical coordinates,

$$\begin{aligned} u^1 &= r\cos\theta \\ u^2 &= r\sin\theta\cos\varphi \\ u^3 &= r\sin\theta\sin\varphi. \end{aligned} \tag{9.126}$$

We also set

$$\mathrm{d}t = \sqrt{g_{vv}(v)}\mathrm{d}v, \quad R^2(t) \equiv f(v). \tag{9.127}$$

The metric takes the form

$$\mathrm{d}\tau^2 = \mathrm{d}t^2 - R^2(t)\left[\frac{\mathrm{d}r^2}{1-\kappa r^2} + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2\right)\right]. \tag{9.128}$$

This is the so-called Robertson–Walker metric. According to the cosmological principle, space in the universe is homogeneous and isotropic (up to small perturbations). The Robertson–Walker metric is then the leading-order approximation to the gravitational field of the universe. The scale factor $R(t)$ is related, through the Einstein field equations, to the energy and matter content of the universe.