Physics In Curved Spacetimes

Yi-Zen Chu

Contents

1 Physical Constants and Dimensional Analysis

Throughout these notes we will set Planck's reduced constant and the speed of light to unity: $\hbar = c = 1$. (In the General Relativity literature, Newton's gravitational constant $G_{\rm N}$ is also often set to one.) What this means is, we are using $\hbar$ as our base unit for angular momentum; and c for speed.

Since $[c]$ is Length/Time, setting it to unity means

$$[\text{Length}] = [\text{Time}] \ .$$

In particular, since in SI units $c = 299,792,458$ meters/second, we have

$$1 \text{ second} = 299,792,458 \text{ meters}, \qquad (c = 1). \tag{1.0.1}$$

Einstein's $E = mc^2$, once $c = 1$, becomes the statement that

$$[\text{Energy}] = [\text{Mass}]$$

Because $[\hbar]$ is Energy $\times$ Time, setting it to unity means

$$[\text{Energy}] = [1/\text{Time}] \ .$$

In SI units, $\hbar \approx 1.0545718 \times 10^{-34}$ Joules second – hence,

$$1 \text{ second} \approx 1/(1.0545718 \times 10^{-34} \text{ Joules}) \qquad (\hbar = 1). \tag{1.0.2}$$

Altogether, with $\hbar = c = 1$, we may state

$$[\text{Mass}] = [\text{Energy}] = [1/\text{Time}] = [1/\text{Length}] \ .$$

Physically speaking, the energy-mass and time-length equivalence can be attributed to relativity (c); whereas the (energy/mass)-(time/length) equivalence can be attributed to quantum mechanics ($\hbar$).

High energy physicists prefer to work with eV (or its multiples, such as MeV or GeV); and so it is useful to know the relation $\hbar c = 197.326,98$ MeV fm. (fm = femtometer = 10^{-15} meters.)

$$10^{-15} \text{ meters} \approx 1/(197.326,98 \text{ MeV}), \qquad (\hbar c = 1). \tag{1.0.3}$$

Using these 'natural units' $\hbar = c = 1$ is a very common practice throughout the physics literature.

One key motivation behind setting to unity physical constants occurring frequently in your physics analysis, is that it allows you to focus on the quantities that are more specific (and hence more important) to the problem at hand. Carrying these physical constants around clutter your calculation, and increases the risk of mistakes due to this additional burden. For instance, in the Bose-Einstein or Fermi-Dirac statistical distribution $1/(\exp(E/(k_B T)) \pm 1)$ – where E, k_B and T are respectively the energy of the particle(s), k_B is the Boltzmann constant, and T is the temperature of the system – what's physically important is the ratio of the energy scales, E versus $k_B T$. The Boltzmann constant k_B is really a distraction, and ought to be set to one, so that temperature is now measured in units of energy: the cleaner expression now reads $1/(\exp(E/T) \pm 1)$.

Another reason why one may want to set a physical constant to unity is because, it could be such an important benchmark in the problem at hand that it should be employed as a base unit.

Most down-to-Earth engineering problems may not benefit from using the speed of light c as their basic unit for speed. In non-relativistic astrophysical systems bound by their mutual gravity, however, it turns out that General Relativistic corrections to the Newtonian law of gravity will be akin to a series in v/c, where v is the typical speed of the bodies that comprise the system. The expansion parameter then becomes $0 \leq v < 1$ if we set $c = 1$ – i.e., if we measure all speeds relative to c – which in turn means this 'post-Newtonian' expansion is a series in the gravitational potential $G_{\rm N}M/r$ through the virial theorem (kinetic energy $\sim$ potential energy) $v \sim \sqrt{G_{\rm N}M/r}$.

Newton's gravitational constant takes the form $G_{\rm N} \approx 6.7086 \times 10^{-39} \hbar c (\text{GeV}/c^2)^{-2}$. Just from this dimensional analysis alone, when $\hbar = c = 1$, one may form a mass-energy scale ('Planck mass')

$$M_{\rm pl} \equiv \frac{1}{\sqrt{32\pi G_{\rm N}}}. \tag{1.0.4}$$

(The 32π is for technical convenience.) We will provide further justification below, but this suggests – since $M_{\rm pl}$ appears to involve relativity (c), quantum mechanics ($\hbar$) and gravitation ($G_{\rm N}$) – that the energy scale required to probe quantum aspects of gravity is roughly $M_{\rm pl}$. Therefore, it may be useful to set $M_{\rm pl} = 1$ in quantum gravity calculations, so that all other energy scales in a given problem, say the quantum amplitude of scattering gravitons, are now measured relative to it.

I recommend the following resource for physical and astrophysical constants, particle physics data, etc.:

Particle Data Group: http://pdg.lbl.gov .

Problem 1.1. Let $\hbar = c = 1$.

- If angular momentum is 3.34, convert it to SI units.
- What is the mass of the Sun in MeV? What is its mass in parsec?
- If Pluto is orbiting roughly 40 astronomical units from the Sun, how many seconds is this orbital distance?
- Work out the Planck mass in eq. (1.0.4) in seconds, meters, and GeV.

□

2 Differential Geometry of & Kinematics In Curved Spacetimes

2.1 Minkowski & Constancy of c; Orthonormal Frames; Timelike, Spacelike vs. Null Vectors; Gravitational Time Dilation

[1]Cartesian coordinates play a basic but special role in interpreting physics in both flat Euclidean space δ_{ij} and flat Minkowski spacetime $\eta_{\mu\nu}$: they parametrize time durations and spatial distances in orthogonal directions – i.e., every increasing tick mark along a given Cartesian axis corresponds directly to a measurement of increasing length or time in that direction. This is generically not so, say, for coordinates in curved space(time) because the notion of what constitutes a 'straight line' is significantly more subtle there; or even spherical coordinates $(r \geq 0, 0 \leq \theta \leq \pi, 0 \leq \phi < 2\pi)$ in flat 3D space – for the latter, only the radial coordinate r corresponds to actual distance (from the origin).

Flat Spacetimes We will therefore begin in flat spacetime written in Cartesian coordinates $\{x^\mu \equiv (t, \vec{x})\}$. Flat spacetime is also otherwise known as *Minkowski* spacetime, and the 'square' of the distance between x^μ and $x^\mu + \mathrm{d}x^\mu$, is given by

$$\begin{aligned} \mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu &= (\mathrm{d}x^0)^2 - \mathrm{d}\vec{x}\cdot\mathrm{d}\vec{x} \\ &= (\mathrm{d}t)^2 - \delta_{ij}\mathrm{d}x^i\mathrm{d}x^j; \end{aligned} \tag{2.1.1}$$

where the Minkowski metric tensor reads

$$\eta_{\mu\nu} \doteq \mathrm{diag}[1, -1, \dots, -1]. \tag{2.1.2}$$

[2]**Constancy of c** One of the primary motivations that led Einstein to recognize eq. (2.1.1) as the proper geometric setting to describe physics, is the realization that the speed of light c is constant in all inertial frames. In modern physics, the latter is viewed as a consequence of spacetime translation and Lorentz symmetry, as well as the null character of the trajectories swept out by photons. That is, for transformation matrices $\{\Lambda\}$ satisfying

$$\Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu \eta_{\alpha\beta} = \eta_{\mu\nu}, \tag{2.1.3}$$

and constant vectors $\{a^\mu\}$ we have

$$\eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu = \eta_{\mu\nu}\mathrm{d}x'^\mu\mathrm{d}x'^\nu \tag{2.1.4}$$

whenever

$$x^\alpha = \Lambda^\alpha{}_\mu x'^\mu + a^\alpha. \tag{2.1.5}$$

The physical interpretation is that the frames parametrized by $\{x^\mu = (t, \vec{x})\}$ and $\{x'^\mu = (t', \vec{x}')\}$ are *inertial* frames: compact bodies with no external forces acting on them will sweep out

[1]These notes were written assuming the reader is familiar with Chapter 7 of Analytical Methods in Physics.

[2]Notice we have switched from Latin/English alphabets, say $i, j, k, \dots \in \{1, 2, 3, \dots, D\}$ to Greek ones $\mu, \nu, \dots \in \{0, 1, 2, \dots, D \equiv d-1\}$; the former run over the spatial coordinates while the latter over time (0th) and space $(1, \dots, D)$. Also note that the opposite 'mostly plus' sign convention $\eta_{\mu\nu} = \mathrm{diag}[-1, +1, \dots, +1]$ is equally valid and, in fact, more popular in the contemporary physics literature.

geodesics $\mathrm{d}^2 x^\mu/\mathrm{d}\tau^2 = 0 = \mathrm{d}^2 x'^\mu/\mathrm{d}\tau'^2$, where the proper times τ and τ' are defined through the relations $\mathrm{d}\tau = \mathrm{d}t\sqrt{1-(\mathrm{d}\vec{x}/\mathrm{d}t)^2}$ and $\mathrm{d}\tau' = \mathrm{d}t'\sqrt{1-(\mathrm{d}\vec{x}'/\mathrm{d}t')^2}$. To interpret physical phenomenon taking place in one frame from the other frame's perspective, one would first have to figure out how to translate between x and x'.

Let x^μ be the spacetime Cartesian coordinates of a single photon; in a different Lorentz frame it has Cartesian coordinates x'^μ. Invoking its null character, namely $\mathrm{d}s^2 = 0$ – which holds in any inertial frame – we have $(\mathrm{d}x^0)^2 = \mathrm{d}\vec{x}\cdot\mathrm{d}\vec{x}$ and $(\mathrm{d}x'^0)^2 = \mathrm{d}\vec{x}'\cdot\mathrm{d}\vec{x}'$. This in turn tells us the speeds in both frames is unity:

$$\frac{|\mathrm{d}\vec{x}|}{\mathrm{d}x^0} = \frac{|\mathrm{d}\vec{x}'|}{\mathrm{d}x'^0} = 1. \tag{2.1.6}$$

A deeper motivational justification would be to recognize, it is the sign difference between the 'time' part and the 'space' part of the metric in eq. (2.1.1) – together with its Lorentz invariance – that gives rise to the wave equations obeyed by the photon. Equation (2.1.6) then follows as a consequence.[3]

Curved Spacetime, Spacetime Volume & Orthonormal Basis The generalization of the distance between x^μ to $x^\mu + \mathrm{d}x^\mu$, from the Minkowski to the curved case, is the following 'line element':

$$\mathrm{d}s^2 = g_{\mu\nu}(x)\mathrm{d}x^\mu \mathrm{d}x^\nu, \tag{2.1.7}$$

where x is simply shorthand for the spacetime coordinates $\{x^\mu\}$, which we emphasize may no longer be Cartesian. We also need to demand that $g_{\mu\nu}$ be real, symmetric, and has 1 positive eigenvalue associated with the one 'time' coordinate and $(d-1)$ negative ones for the spatial coordinates. The infinitesimal spacetime volume continues to take the form

$$\mathrm{d}(\text{vol.}) = \mathrm{d}^d x\sqrt{|g(x)|}, \tag{2.1.8}$$

where $|g(x)| = |\det g_{\mu\nu}(x)|$ is now the absolute value of the determinant of the metric $g_{\mu\nu}$.

Just like the curved space case, to interpret physics in the neighborhood of some spacetime location x^μ, we introduce an orthonormal basis $\{\varepsilon^{\widehat{\mu}}{}_\alpha\}$ through the 'diagonalization' process:

$$g_{\mu\nu}(x) = \eta_{\alpha\beta}\varepsilon^{\widehat{\alpha}}{}_\mu(x)\varepsilon^{\widehat{\beta}}{}_\nu(x). \tag{2.1.9}$$

By defining $\varepsilon^{\widehat{\alpha}} \equiv \varepsilon^{\widehat{\alpha}}{}_\mu \mathrm{d}x^\mu$, the analog to achieving a Cartesian-like expression for the spacetime metric is

$$\mathrm{d}s^2 = \left(\varepsilon^{\widehat{0}}\right)^2 - \sum_{i=1}^{D}\left(\varepsilon^{\widehat{i}}\right)^2 = \eta_{\mu\nu}\varepsilon^{\widehat{\mu}}\varepsilon^{\widehat{\nu}}. \tag{2.1.10}$$

This means under a local Lorentz transformation – i.e., for all

$$\Lambda^\mu{}_\alpha(x)\Lambda^\nu{}_\beta(x)\eta_{\mu\nu} = \eta_{\alpha\beta}, \tag{2.1.11}$$

$$\varepsilon'^{\widehat{\mu}}(x) = \Lambda^\mu{}_\alpha(x)\varepsilon'^{\widehat{\alpha}}(x) \tag{2.1.12}$$

[3]We will discuss these issues in some detail below.

– the metric remains the same:

$$\mathrm{d}s^2 = \eta_{\mu\nu}\varepsilon^{\widehat{\mu}}\varepsilon^{\widehat{\nu}} = \eta_{\mu\nu}\varepsilon'^{\widehat{\mu}}\varepsilon'^{\widehat{\nu}}. \tag{2.1.13}$$

By viewing $\widehat{\varepsilon}$ as the matrix with the αth row and μth column given by $\varepsilon^{\widehat{\alpha}}{}_{\mu}$, the determinant of the metric $g_{\mu\nu}$ can be written as

$$\det g_{\mu\nu}(x) = (\det\widehat{\varepsilon})^2 \det\eta_{\mu\nu}. \tag{2.1.14}$$

The infinitesimal spacetime volume in eq. (2.1.8) now can be expressed as

$$\mathrm{d}^d x\sqrt{|g(x)|} = \mathrm{d}^d x \det\widehat{\varepsilon}. \tag{2.1.15}$$

Of course, that $g_{\mu\nu}$ may be 'diagonalized' follows from the fact that $g_{\mu\nu}$ is a real symmetric matrix:

$$g_{\mu\nu} = \sum_{\alpha,\beta} O^{\alpha}{}_{\mu}\lambda_\alpha\eta_{\alpha\beta}O^{\beta}{}_{\nu} = \sum_{\alpha,\beta}\varepsilon^{\widehat{\alpha}}{}_{\mu}\eta_{\alpha\beta}\varepsilon^{\widehat{\beta}}{}_{\nu}, \tag{2.1.16}$$

where all $\{\lambda_\alpha\}$ are positive, so we may take their positive root:

$$\varepsilon^{\widehat{\alpha}}{}_{\mu} = \sqrt{\lambda_\alpha}O^{\alpha}{}_{\mu}, \qquad \{\lambda_\alpha > 0\}, \qquad (\text{No sum over } \alpha). \tag{2.1.17}$$

That $\varepsilon^{\widehat{0}}{}_{\mu}$ acts as 'standard clock' and $\{\varepsilon^{\widehat{i}}{}_{\mu}|i = 1, 2, \dots, D\}$ act as 'standard rulers' is because they are of unit length:

$$g^{\mu\nu}\varepsilon^{\widehat{\alpha}}{}_{\mu}\varepsilon^{\widehat{\beta}}{}_{\nu} = \eta^{\alpha\beta}. \tag{2.1.18}$$

The $\widehat{\cdot}$ on the index indicates it is to be moved with the flat metric, namely

$$\varepsilon^{\widehat{\alpha}}{}_{\mu} = \eta^{\alpha\beta}\varepsilon_{\widehat{\beta}\mu} \qquad \text{and} \qquad \varepsilon_{\widehat{\alpha}\mu} = \eta_{\alpha\beta}\varepsilon^{\widehat{\beta}}{}_{\mu}; \tag{2.1.19}$$

while the spacetime index is to be moved with the spacetime metric

$$\varepsilon^{\widehat{\alpha}\mu} = g^{\mu\nu}\varepsilon^{\widehat{\alpha}}{}_{\nu} \qquad \text{and} \qquad \varepsilon^{\widehat{\alpha}}{}_{\mu} = g_{\mu\nu}\varepsilon^{\widehat{\alpha}\nu}. \tag{2.1.20}$$

In other words, we view $\varepsilon_{\widehat{\alpha}}{}^{\mu}$ as the μth spacetime component of the αth vector field in the basis set $\{\varepsilon_{\widehat{\alpha}}{}^{\mu}|\alpha = 0, 1, 2, \dots, D \equiv d-1\}$. We may elaborate on the interpretation that $\{\varepsilon^{\widehat{\alpha}}{}_{\mu}\}$ act as 'standard clock/rulers' as follows. For a test (scalar) function $f(x)$ defined throughout spacetime, the rate of change of f along $\varepsilon_{\widehat{0}}$ is

$$\langle \mathrm{d}f|\,\varepsilon_{\widehat{0}}\rangle = \varepsilon_{\widehat{0}}{}^{\mu}\partial_\mu f \equiv \frac{\mathrm{d}f}{\mathrm{d}y^0}; \tag{2.1.21}$$

whereas that along $\varepsilon_{\widehat{i}}$ is

$$\langle \mathrm{d}f|\,\varepsilon_{\widehat{i}}\rangle = \varepsilon_{\widehat{i}}{}^{\mu}\partial_\mu f \equiv \frac{\mathrm{d}f}{\mathrm{d}y^i}; \tag{2.1.22}$$

where y^0 and $\{y^i\}$ are to be viewed as 'time' and 'spatial' parameters along the integral curves of $\{\varepsilon_{\widehat{\mu}}{}^{\alpha}\}$. That these are Cartesian-like can now be expressed as

$$\left\langle \frac{\mathrm{d}}{\mathrm{d}y^\mu} \middle| \frac{\mathrm{d}}{\mathrm{d}y^\nu} \right\rangle = \varepsilon_{\widehat{\mu}}{}^{\alpha}\varepsilon_{\widehat{\nu}}{}^{\beta} \left\langle \partial_\mu \middle| \partial_\nu \right\rangle = \varepsilon_{\widehat{\mu}}{}^{\alpha}\varepsilon_{\widehat{\nu}}{}^{\beta} g_{\mu\nu} = \eta_{\mu\nu}. \tag{2.1.23}$$

Note that the first equalities of eq. (2.1.16) are really assumptions, in that the definitions of curved spaces include assuming all the eigenvalues of the metric are positive whereas that of curved spacetimes include assuming all but one eigenvalue is negative.[4]

Note that the $\{\mathrm{d}/\mathrm{d}y^\mu\}$ in eq. (2.1.23) do not, generically, commute. For instance, acting on a scalar function,

$$\left[\frac{\mathrm{d}}{\mathrm{d}y^\mu}, \frac{\mathrm{d}}{\mathrm{d}y^\nu}\right] f(x) = \left(\frac{\mathrm{d}}{\mathrm{d}y^\mu}\frac{\mathrm{d}}{\mathrm{d}y^\nu} - \frac{\mathrm{d}}{\mathrm{d}y^\nu}\frac{\mathrm{d}}{\mathrm{d}y^\mu}\right) f(x) \tag{2.1.24}$$

$$= \left(\varepsilon_{\widehat{\mu}}{}^{\alpha}\partial_\alpha\varepsilon_{\widehat{\nu}}{}^{\beta} - \varepsilon_{\widehat{\nu}}{}^{\alpha}\partial_\alpha\varepsilon_{\widehat{\mu}}{}^{\beta}\right)\partial_\beta f(x) \neq 0. \tag{2.1.25}$$

A theorem in differential geometry – see, for instance, Schutz [4] for a pedagogical discussion – tells us:

> A set of $1 < N \leq d$ vector fields $\{\mathrm{d}/\mathrm{d}\xi^\mu\}$ form a coordinate basis in the $N-$dimensional space(time) they inhabit, if and only if they commute.

When $N = d$, and if $[\mathrm{d}/\mathrm{d}y^\mu, \mathrm{d}/\mathrm{d}y^\nu] = 0$ in eq. (2.1.23), we would not only have found coordinates $\{y^\mu\}$ for our spacetime, we would have found this spacetime is a flat one.

It is important to clarify what a coordinate system is. In 2D, for instance, if we had $[\mathrm{d}/\mathrm{d}y^0, \mathrm{d}/\mathrm{d}y^1] \neq 0$. This means it is not possible to vary the 'coordinate' y^0 (i.e., along the integral curve of $\mathrm{d}/\mathrm{d}y^0$) without holding the 'coordinate' y^1 fixed; and it is not possible to hold y^0 fixed while moving along the integral curve of $\mathrm{d}/\mathrm{d}y^1$.

Problem 2.1. Example: Schutz [4] Exercise 2.1 In 2D flat space, starting from Cartesian coordinates x^i, we may convert to cylindrical coordinates

$$(x^1, x^2) = r(\cos\phi, \sin\phi). \tag{2.1.26}$$

The pair of vector fields $(\partial_r, \partial_\phi)$ do form a coordinate basis – it is possible to hold r fixed while going along the integral curve of ∂_ϕ and vice versa. However, show via a direct calculation that the following commutator involving the unit vector fields $\widehat{r}$ and $\widehat{\phi}$ is not zero:

$$\left[\widehat{r}, \widehat{\phi}\right] f(r, \phi) \neq 0; \tag{2.1.27}$$

where

$$\widehat{r} \equiv \cos(\phi)\partial_{x^1} + \sin(\phi)\partial_{x^2}, \tag{2.1.28}$$

$$\widehat{\phi} \equiv -\sin(\phi)\partial_{x^1} + \cos(\phi)\partial_{x^2}. \tag{2.1.29}$$

Therefore $\widehat{r}$ and $\widehat{\phi}$ do not form a coordinate basis. □

[4] In $d-$spacetime dimensions, if there were n 'time' directions and $(d-n)$ 'spatial' ones, then this carries with it the assumption that $g_{\mu\nu}$ has n positive eigenvalues and $(d-n)$ negative ones.

Timelike, spacelike, and null distances/vectors A fundamental difference between (curved) space and spacetime, is that the former involves strictly positive distances while the latter – because of the $\eta_{00} = +1$ for orthonormal 'time' versus $\eta_{ii} = -1$ for the ith orthonormal space component – involves positive, zero, and negative distances.

With our 'mostly minus' sign convention (cf. eq. (3.0.2)), a vector v^μ is:

- *Time-like* if $v^2 \equiv \eta_{\mu\nu} v^{\widehat{\mu}} v^{\widehat{\nu}} > 0$. We shall soon see that, if $v^2 > 0$, it is always possible to find a Lorentz transformation Λ (cf. eq. (2.1.3)) such that $\Lambda^\mu{}_\alpha v^{\widehat{\alpha}} = (v'^{\widehat{0}}, \vec{0})$. In flat spacetime, if $\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu > 0$ then this result indicates it is always possible to find an inertial frame where $\mathrm{d}s^2 = \mathrm{d}t'^2$: hence the phrase 'timelike'.

  More generally, for a timelike trajectory $z^\mu(\lambda)$ in curved spacetime – i.e., $g_{\mu\nu}(\mathrm{d}z^\mu/\mathrm{d}\lambda)(\mathrm{d}z^\nu/\mathrm{d}\lambda) > 0$, we may identify

  $$\mathrm{d}\tau \equiv \mathrm{d}\lambda \sqrt{g_{\mu\nu}(z(\lambda)) \frac{\mathrm{d}z^\mu}{\mathrm{d}\lambda} \frac{\mathrm{d}z^\nu}{\mathrm{d}\lambda}} \tag{2.1.30}$$

  as the (infinitesimal) *proper time*, the time read by the watch of an observer whose worldline is $z^\mu(\lambda)$. (As a check: when $g_{\mu\nu} = \eta_{\mu\nu}$ and the observer is at rest, namely $\mathrm{d}\vec{z} = 0$, then $\mathrm{d}\tau = \mathrm{d}t$.) Using orthonormal frame fields in eq. (2.1.16),

  $$\mathrm{d}\tau = \mathrm{d}\lambda \sqrt{\eta_{\alpha\beta} \frac{\mathrm{d}z^{\widehat{\alpha}}}{\mathrm{d}\lambda} \frac{\mathrm{d}z^{\widehat{\beta}}}{\mathrm{d}\lambda}}, \qquad \frac{\mathrm{d}z^{\widehat{\alpha}}}{\mathrm{d}\lambda} \equiv \varepsilon^{\widehat{\alpha}}{}_\mu \frac{\mathrm{d}z^\mu}{\mathrm{d}\lambda}. \tag{2.1.31}$$

  Furthermore, since $v^{\widehat{\mu}} \equiv \mathrm{d}z^{\widehat{\mu}}/\mathrm{d}\lambda$ is assumed to be timelike, it must be possible to find a Lorentz transformation $\Lambda^\mu{}_\nu$ such that $\Lambda^\mu{}_\nu v^{\widehat{\nu}} = (v'^{\widehat{0}}, \vec{0})$; assuming $\mathrm{d}\lambda > 0$,

  $$\begin{aligned} \mathrm{d}\tau &= \mathrm{d}\lambda \sqrt{\eta_{\mu\nu} \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \frac{\mathrm{d}z^{\widehat{\alpha}}}{\mathrm{d}\lambda} \frac{\mathrm{d}z^{\widehat{\beta}}}{\mathrm{d}\lambda}}, \\ &= \mathrm{d}\lambda \sqrt{\left(\frac{\mathrm{d}z'^{\widehat{0}}}{\mathrm{d}\lambda}\right)^2} = |\mathrm{d}z'^{\widehat{0}}|. \end{aligned} \tag{2.1.32}$$

- *Space-like* if $v^2 \equiv \eta_{\mu\nu} v^{\widehat{\mu}} v^{\widehat{\nu}} < 0$. We shall soon see that, if $v^2 < 0$, it is always possible to find a Lorentz transformation Λ such that $\Lambda^\mu{}_\alpha v^{\widehat{\alpha}} = (0, v'^{\widehat{i}})$. In flat spacetime, if $\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu < 0$ then this result indicates it is always possible to find an inertial frame where $\mathrm{d}s^2 = -\mathrm{d}\vec{x}'^2$: hence the phrase 'spacelike'.

  More generally, for a spacelike trajectory $z^\mu(\lambda)$ in curved spacetime – i.e., $g_{\mu\nu}(\mathrm{d}z^\mu/\mathrm{d}\lambda)(\mathrm{d}z^\nu/\mathrm{d}\lambda) < 0$, we may identify

  $$\mathrm{d}\ell \equiv \mathrm{d}\lambda \sqrt{\left| g_{\mu\nu}(z(\lambda)) \frac{\mathrm{d}z^\mu}{\mathrm{d}\lambda} \frac{\mathrm{d}z^\nu}{\mathrm{d}\lambda} \right|} \tag{2.1.33}$$

  as the (infinitesimal) *proper length*, the distance read off some measuring rod whose trajectory is $z^\mu(\lambda)$. (As a check: when $g_{\mu\nu} = \eta_{\mu\nu}$ and $\mathrm{d}t = 0$, i.e., the rod is lying on the constant$-t$ surface, then $\mathrm{d}\ell = |\mathrm{d}\vec{x} \cdot \mathrm{d}\vec{x}|^{1/2}$.) Using the orthonormal frame fields in eq. (2.1.16),

$$\mathrm{d}\ell = \mathrm{d}\lambda \sqrt{\left|\eta_{\alpha\beta} \frac{\mathrm{d}z^{\widehat{\alpha}}}{\mathrm{d}\lambda} \frac{\mathrm{d}z^{\widehat{\beta}}}{\mathrm{d}\lambda}\right|}, \qquad \frac{\mathrm{d}z^{\widehat{\alpha}}}{\mathrm{d}\lambda} \equiv \varepsilon^{\widehat{\alpha}}{}_{\mu} \frac{\mathrm{d}z^{\mu}}{\mathrm{d}\lambda}. \tag{2.1.34}$$

Furthermore, since $v^{\widehat{\mu}} \equiv \mathrm{d}z^{\widehat{\mu}}/\mathrm{d}\lambda$ is assumed to be spacelike, it must be possible to find a Lorentz transformation $\Lambda^{\mu}{}_{\nu}$ such that $\Lambda^{\mu}{}_{\nu} v^{\widehat{\nu}} = (0, v'^{\widehat{i}})$; assuming $\mathrm{d}\lambda > 0$,

$$\mathrm{d}\ell = \mathrm{d}\lambda \sqrt{\eta_{\mu\nu} \Lambda^{\mu}{}_{\alpha} \Lambda^{\nu}{}_{\beta} \frac{\mathrm{d}z^{\widehat{\alpha}}}{\mathrm{d}\lambda} \frac{\mathrm{d}z^{\widehat{\beta}}}{\mathrm{d}\lambda}} = |\mathrm{d}\vec{z}'|\,; \tag{2.1.35}$$

$$\mathrm{d}z'^{\widehat{i}} \equiv \Lambda^{i}{}_{\mu} \varepsilon^{\widehat{\mu}}{}_{\nu} \mathrm{d}z^{\nu}. \tag{2.1.36}$$

- *Null* if $v^2 \equiv \eta_{\mu\nu} v^{\widehat{\mu}} v^{\widehat{\nu}} = 0$. We have already seen, in flat spacetime, if $\mathrm{d}s^2 = \eta_{\mu\nu} \mathrm{d}x^{\mu} \mathrm{d}x^{\nu} = 0$ then $|\mathrm{d}\vec{x}|/\mathrm{d}x^0 = |\mathrm{d}\vec{x}'|/\mathrm{d}x'^0 = 1$ in all inertial frames.

It is physically important to reiterate: one of the reasons why it is important to make such a distinction between vectors, is because it is *not possible* to find a Lorentz transformation that would linearly transform one of the above three types of vectors into another different type – for e.g., it is not possible to Lorentz transform a null vector into a time-like one (a photon has no 'rest frame'); or a time-like vector into a space-like one; etc. This is because their Lorentzian 'norm'

$$v^2 \equiv \eta_{\mu\nu} v^{\widehat{\mu}} v^{\widehat{\nu}} = \eta_{\mu\nu} \Lambda^{\mu}{}_{\alpha} \Lambda^{\nu}{}_{\beta} v'^{\widehat{\alpha}} v'^{\widehat{\beta}} \tag{2.1.37}$$

has to be invariant under all Lorentz transformations $v^{\widehat{\mu}} \equiv \Lambda^{\mu}{}_{\alpha} v'^{\widehat{\alpha}}$. This in turn teaches us: if v^2 were positive, it has to remain so; likewise, if it were zero or negative, a Lorentz transformation cannot alter this attribute.

Problem 2.2. Orthonormal Frames in Kerr-Schild Spacetimes A special class of geometries, known as *Kerr-Schild* spacetimes, take the following form.

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + H k_{\mu} k_{\nu} \tag{2.1.38}$$

Many of the known black hole spacetimes can be put in this form; and in such a context, $\bar{g}_{\mu\nu}$ usually refers to flat or de Sitter spacetime.[5] The k_{μ} is null with respect to $\bar{g}_{\mu\nu}$, and we shall move its indices with $\bar{g}_{\mu\nu}$.

Verify that the inverse metric is

$$g^{\mu\nu} = \bar{g}^{\mu\nu} - H k^{\mu} k^{\nu}, \tag{2.1.39}$$

where $\bar{g}^{\mu\sigma}$ is the inverse of $\bar{g}_{\mu\sigma}$, namely $\bar{g}^{\mu\sigma} \bar{g}_{\sigma\nu} \equiv \delta^{\mu}{}_{\nu}$. Then, verify that the orthonormal frame fields are

$$\varepsilon^{\widehat{\alpha}}{}_{\mu} = \delta^{\alpha}{}_{\mu} + \frac{1}{2} H k^{\alpha} k_{\mu}. \tag{2.1.40}$$

Can you explain why k^{μ} is also null with respect to the full metric $g_{\mu\nu}$? □

[5] See Gibbons et al. [6] arXiv: hep-th/0404008. The special property of Kerr-Schild coordinates is that Einstein's equations become *linear* in these coordinates.

Proper times and Gravitational Time Dilation Consider two observers sweeping out their respective timelike worldlines in spacetime, $y^\mu(\lambda)$ and $z^\mu(\lambda)$. If we use the time coordinate of the geometry to parameterize their trajectories, their proper times – i.e., the time read by their watches – are given by

$$\mathrm{d}\tau_y \equiv \mathrm{d}t\sqrt{g_{\mu\nu}(y(t))\dot{y}^\mu \dot{y}^\nu}, \qquad \dot{y}^\mu \equiv \frac{\mathrm{d}y^\mu}{\mathrm{d}t}; \tag{2.1.41}$$

$$\mathrm{d}\tau_z \equiv \mathrm{d}t\sqrt{g_{\mu\nu}(z(t))\dot{z}^\mu \dot{z}^\nu}, \qquad \dot{z}^\mu \equiv \frac{\mathrm{d}z^\mu}{\mathrm{d}t}. \tag{2.1.42}$$

In flat spacetime, clocks that are synchronized in one frame are no longer synchronized in a different frame – chronology is not a Lorentz invariant. We see that, in curved spacetime, the infinitesimal *passage* of proper time measured by observers at the same 'coordinate time' t depends on their spacetime locations:

$$\frac{\mathrm{d}\tau_y}{\mathrm{d}\tau_z} = \sqrt{\frac{g_{\mu\nu}(y(t))\dot{y}^\mu \dot{y}^\nu}{g_{\alpha\beta}(z(t))\dot{y}^\alpha \dot{y}^\beta}}. \tag{2.1.43}$$

More generally, to compare the passage of time on one observer's trajectory $y^\mu(\lambda_1 \leq \lambda \leq \lambda_2)$ against another $z^\mu(\lambda'_1 \leq \lambda' \leq \lambda'_2)$, the integrals $\int_{\lambda_1}^{\lambda_2} \mathrm{d}\lambda\sqrt{g_{\mu\nu}\dot{y}^\mu\dot{y}^\nu}$ and $\int_{\lambda'_1}^{\lambda'_2} \mathrm{d}\lambda'\sqrt{g_{\mu\nu}\dot{z}^\mu\dot{z}^\nu}$ have to be evaluated.

Problem 2.3. Example We will have more to say about this later, but the spacetime geometry around the Earth itself can be approximated by the line element

$$\mathrm{d}s^2 = \left(1 - \frac{r_{s,E}}{r}\right)\mathrm{d}t^2 - \frac{\mathrm{d}r^2}{1 - r_{s,E}/r} - r^2\left(\mathrm{d}\theta^2 + \sin(\theta)^2\mathrm{d}\phi^2\right), \tag{2.1.44}$$

where t is the time coordinate and (r, θ, ϕ) are analogs of the spherical coordinates. Whereas $r_{s,E}$ is known as the Schwarzschild radius of the Earth, and depends on the Earth's mass M_{E} through the expression

$$r_{s,E} \equiv 2G_{\mathrm{N}}M_{\mathrm{E}}. \tag{2.1.45}$$

Find the 4−beins of the geometry in eq. (2.1.44). Then find the numerical value of $r_{s,E}$ in eq. (2.1.45) and take the ratio $r_{s,E}/R_{\mathrm{E}}$, where R_{E} is the radius of the Earth. Explain why this means we may – for practical purposes – expand the metric in eq. (2.1.45) as

$$\begin{aligned}\mathrm{d}s^2 = \left(1 - \frac{r_{s,E}}{r}\right)\mathrm{d}t^2 - \mathrm{d}r^2\left(1 + \frac{r_{s,E}}{r} + \left(\frac{r_{s,E}}{r}\right)^2 + \left(\frac{r_{s,E}}{r}\right)^3 + \dots\right) \\ - r^2\left(\mathrm{d}\theta^2 + \sin(\theta)^2\mathrm{d}\phi^2\right).\end{aligned} \tag{2.1.46}$$

Since we are not in flat spacetime, the (t, r, θ, ϕ) are no longer subject to the same interpretation. However, use your computation of $r_{s,E}/R_{\mathrm{E}}$ to *estimate* the error incurred if we do continue to interpret t and r as though they measured time and radial distances, with respect to a frame centered at the Earth's core.

Consider placing one clock at the base of the Taipei 101 tower and another at its tip. Denoting the time elapsed at the base of the tower as $\Delta\tau_{\rm B}$; that at the tip as $\Delta\tau_{\rm T}$; and assuming for simplicity the Earth is a perfect sphere – show that

$$\frac{\Delta\tau_{\rm B}}{\Delta\tau_{\rm T}} = \sqrt{\frac{g_{00}(R_{\rm E})}{g_{00}(R_{\rm E}+h_{101})}} \approx 1 + \frac{1}{2}\left(\frac{r_{s,E}}{R_{\rm E}+h_{101}} - \frac{r_{s,E}}{R_{\rm E}}\right). \tag{2.1.47}$$

Here, $R_{\rm E}$ is the radius of the Earth and h_{101} is the height of the Taipei 101 tower.

In actuality, both clocks are in motion, since the Earth is rotating. Can you estimate what is the error incurred from assuming they are at rest? First arrive at eq. (2.1.47) analytically, then plug in the relevant numbers to compute the numerical value of $\Delta\tau_{\rm B}/\Delta\tau_{\rm T}$. Does the clock at the base of Taipei 101 or that on its tip tick more slowly?

This gravitational time dilation is an effect that needs to be accounted for when setting up a network of Global Positioning Satellites (GPS); for details, see here. See also the Wikipedia article on the Pound-Rebka experiment, the first verification of the gravitational time dilation effect. □

2.2 Connections, Curvature, Geodesics, Isometries, Weak & Einstein Equivalence Principles, Tidal Forces

Connections, Christoffel Symbols & Riemann curvature tensor Because the partial derivative itself cannot yield a tensor once it acts on tensor, we need to introduce a connection $\Gamma^{\mu}{}_{\alpha\beta}$, i.e.,

$$\nabla_\sigma V^\mu = \partial_\sigma V^\mu + \Gamma^\mu{}_{\sigma\rho} V^\rho. \tag{2.2.1}$$

Under a coordinate transformation of the partial derivatives and V^μ, say going from x to x',

$$\partial_\sigma V^\mu + \Gamma^\mu{}_{\sigma\rho} V^\rho = \frac{\partial x'^\lambda}{\partial x^\sigma}\frac{\partial x^\mu}{\partial x'^\nu}\partial_{\lambda'} V^{\nu'} + \left(\frac{\partial x'^\lambda}{\partial x^\sigma}\frac{\partial^2 x^\mu}{\partial x'^\lambda x'^\nu} + \Gamma^\mu{}_{\sigma\rho}\frac{\partial x^\rho}{\partial x'^\nu}\right) V^{\nu'}. \tag{2.2.2}$$

On the other hand, if $\nabla_\sigma V^\mu$ were to transform as a tensor,

$$\partial_\sigma V^\mu + \Gamma^\mu{}_{\sigma\rho} V^\rho = \frac{\partial x'^\lambda}{\partial x^\sigma}\frac{\partial x^\mu}{\partial x'^\nu}\partial_{\lambda'} V^{\nu'} + \frac{\partial x'^\lambda}{\partial x^\sigma}\frac{\partial x^\mu}{\partial x'^\tau}\Gamma^{\tau'}{}_{\lambda'\nu'} V^{\nu'}. \tag{2.2.3}$$

Since $V^{\nu'}$ is an arbitrary vector, we may read off its coefficient on the right hand sides of equations (2.2.2) and (2.2.3), and deduce the connection has to transform as

$$\frac{\partial x'^\lambda}{\partial x^\sigma}\frac{\partial^2 x^\mu}{\partial x'^\lambda x'^\nu} + \Gamma^\mu{}_{\sigma\rho}\frac{\partial x^\rho}{\partial x'^\nu} = \frac{\partial x'^\lambda}{\partial x^\sigma}\frac{\partial x^\mu}{\partial x'^\tau}\Gamma^{\tau'}{}_{\lambda'\nu'}. \tag{2.2.4}$$

Moving all the Jacobians onto the connection written in the $\{x^\mu\}$ frame,

$$\Gamma^{\tau'}{}_{\kappa'\nu'}(x') = \frac{\partial x'^\tau}{\partial x^\mu}\frac{\partial^2 x^\mu}{\partial x'^\kappa x'^\nu} + \frac{\partial x'^\tau}{\partial x^\mu}\Gamma^\mu{}_{\sigma\rho}(x)\frac{\partial x^\sigma}{\partial x'^\kappa}\frac{\partial x^\rho}{\partial x'^\nu}. \tag{2.2.5}$$

All connections have to satisfy this non-tensorial transformation law. On the other hand, if we found an object that transforms according to eq. (2.2.5), and if one employs it in eq. (2.2.1), then the resulting $\nabla_\alpha V^\mu$ would transform as a tensor.

Problem 2.4. Let us take the derivative of a 1-form instead:

$$\nabla_\alpha V_\mu = \partial_\alpha V_\mu + \Gamma'^\sigma{}_{\alpha\mu} V_\sigma. \tag{2.2.6}$$

Can you prove that

$$\Gamma'^\sigma{}_{\alpha\mu} = -\Gamma^\sigma{}_{\alpha\mu}, \tag{2.2.7}$$

where $\Gamma^\sigma{}_{\alpha\mu}$ is the connection in eq. (2.2.1) – if we define the covariant derivative of a scalar to be simply the partial derivative acting on the same, i.e.,

$$\nabla_\alpha \left(V^\mu W_\mu\right) = \partial_\alpha \left(V^\mu W_\mu\right)? \tag{2.2.8}$$

You should assume the product rule holds, namely $\nabla_\alpha (V^\mu W_\mu) = (\nabla_\alpha V^\mu) W_\mu + V^\mu (\nabla_\alpha W_\mu)$. Expand these covariant derivatives in terms of the connections and argue why this leads to eq. (2.2.7). □

Suppose we found two such connections, ${}_{(1)}\Gamma^\tau{}_{\kappa\nu}(x)$ and ${}_{(2)}\Gamma^\tau{}_{\kappa\nu}(x)$. Notice their difference does transform as a tensor because the first term on the right hand side involving the Hessian $\partial^2 x/\partial x' \partial x'$ cancels out:

$$ {}_{(1)}\Gamma^{\tau'}{}_{\kappa'\nu'}(x') - {}_{(2)}\Gamma^{\tau'}{}_{\kappa'\nu'}(x') = \frac{\partial x'^\tau}{\partial x^\mu} \left({}_{(1)}\Gamma^\mu{}_{\sigma\rho}(x) - {}_{(2)}\Gamma^\mu{}_{\sigma\rho}(x) \right) \frac{\partial x^\sigma}{\partial x'^\kappa} \frac{\partial x^\rho}{\partial x'^\nu}. \tag{2.2.9}$$

Now, any connection can be decomposed into its symmetric and antisymmetric parts in the following sense:

$$\Gamma^\mu{}_{\alpha\beta} = \frac{1}{2}\Gamma^\mu{}_{\{\alpha\beta\}} + \frac{1}{2}\Gamma^\mu{}_{[\alpha\beta]}. \tag{2.2.10}$$

This is, of course, mere tautology. However, let us denote

$$ {}_{(1)}\Gamma^\mu{}_{\alpha\beta} \equiv \frac{1}{2}\Gamma^\mu{}_{\alpha\beta}, \tag{2.2.11}$$

$$ {}_{(2)}\Gamma^\mu{}_{\alpha\beta} \equiv \frac{1}{2}\Gamma^\mu{}_{\beta\alpha}; \tag{2.2.12}$$

so that

$$\frac{1}{2}\Gamma^\mu{}_{[\alpha\beta]} = {}_{(1)}\Gamma^\mu{}_{\alpha\beta} - {}_{(2)}\Gamma^\mu{}_{\alpha\beta} \equiv T^\mu{}_{\alpha\beta}. \tag{2.2.13}$$

We then see that this anti-symmetric part of the connection is in fact a tensor. It is the symmetric part $(1/2)\Gamma^\mu{}_{\{\alpha\beta\}}$ that does not transform as a tensor. *For the rest of these notes, by $\Gamma^\mu{}_{\alpha\beta}$ we shall always mean a symmetric connection.* This means our covariant derivative would now read

$$\nabla_\alpha V^\mu = \partial_\alpha V^\mu + \Gamma^\mu{}_{\alpha\beta} V^\beta + T^\mu{}_{\alpha\beta} V^\beta. \tag{2.2.14}$$

As is common within the physics literature, we proceed to set to zero the torsion term $T^\mu{}_{\alpha\beta}$. If we further impose the metric compatibility condition,

$$\nabla_\mu g_{\alpha\beta} = 0, \tag{2.2.15}$$

then this implies

$$\Gamma^{\mu}{}_{\alpha\beta} = \frac{1}{2} g^{\mu\sigma} \left(\partial_\alpha g_{\beta\sigma} + \partial_\beta g_{\alpha\sigma} - \partial_\sigma g_{\alpha\beta}\right). \tag{2.2.16}$$

The measure of local spacetime curvature that is equivalent to parallel transporting a vector V^μ around an infinitesimal closed loop formed by the tangent vectors A^μ and B^μ, is the following commutator. With the notation $\nabla_A \equiv A^\sigma \nabla_\sigma$, denoting the covariant derivative along A, etc.,

$$\begin{aligned}[\nabla_A, \nabla_B] V^\mu &\equiv A^\sigma \nabla_\sigma \left(B^\rho \nabla_\rho V^\mu\right) - B^\sigma \nabla_\sigma \left(A^\rho \nabla_\rho V^\mu\right) \\ &= \left(A^\sigma \nabla_\sigma B^\rho \nabla_\rho - B^\sigma \nabla_\sigma A^\rho \nabla_\rho\right) V^\mu + A^\sigma B^\rho [\nabla_\sigma, \nabla_\rho] V^\mu. \end{aligned} \tag{2.2.17}$$

Let us compute the two groups separately. Firstly,

$$\begin{aligned}[A, B]^\rho \nabla_\rho V^\mu &\equiv \left(A^\sigma \nabla_\sigma B^\rho \nabla_\rho - B^\sigma \nabla_\sigma A^\rho \nabla_\rho\right) V^\mu \\ &= \left(A^\sigma \partial_\sigma B^\rho + \Gamma^\rho{}_{\sigma\lambda} A^\sigma B^\lambda - A^\sigma \partial_\sigma B^\rho - \Gamma^\rho{}_{\sigma\lambda} A^\sigma B^\lambda\right) \nabla_\rho V^\mu \\ &= \left(A^\sigma \partial_\sigma B^\rho - A^\sigma \partial_\sigma B^\rho\right) \nabla_\rho V^\mu. \end{aligned} \tag{2.2.18}$$

Next, we need $A^\sigma B^\rho [\nabla_\sigma, \nabla_\rho] V^\mu = A^\sigma B^\rho (\nabla_\sigma \nabla_\rho - \nabla_\rho \nabla_\sigma) V^\mu$. The first term is

$$\begin{aligned} A^\sigma B^\rho \nabla_\sigma \nabla_\rho V^\mu &= A^\sigma B^\rho \left(\partial_\sigma \nabla_\rho V^\mu - \Gamma^\lambda{}_{\sigma\rho} \nabla_\lambda V^\mu + \Gamma^\mu{}_{\sigma\lambda} \nabla_\rho V^\lambda\right) \\ &= A^\sigma B^\rho \left(\partial_\sigma \left(\partial_\rho V^\mu + \Gamma^\mu{}_{\rho\lambda} V^\lambda\right) - \Gamma^\lambda{}_{\sigma\rho} \left(\partial_\lambda V^\mu + \Gamma^\mu{}_{\lambda\omega} V^\omega\right) + \Gamma^\mu{}_{\sigma\lambda} \left(\partial_\rho V^\lambda + \Gamma^\lambda{}_{\rho\omega} V^\omega\right)\right) \\ &= A^\sigma B^\rho \Big\{ \partial_\sigma \partial_\rho V^\mu + \partial_\sigma \Gamma^\mu{}_{\rho\lambda} V^\lambda + \Gamma^\mu{}_{\rho\lambda} \partial_\sigma V^\lambda - \Gamma^\lambda{}_{\sigma\rho} \left(\partial_\lambda V^\mu + \Gamma^\mu{}_{\lambda\omega} V^\omega\right) \\ &\qquad + \Gamma^\mu{}_{\sigma\lambda} \left(\partial_\rho V^\lambda + \Gamma^\lambda{}_{\rho\omega} V^\omega\right) \Big\}. \end{aligned} \tag{2.2.19}$$

Swapping $(\sigma \leftrightarrow \rho)$ within the parenthesis $\{\dots\}$ and subtract the two results, we gather

$$\begin{aligned} A^\sigma B^\rho [\nabla_\sigma, \nabla_\rho] V^\mu &= A^\sigma B^\rho \Big\{ \partial_{[\sigma} \Gamma^\mu{}_{\rho]\lambda} V^\lambda + \Gamma^\mu{}_{\lambda[\rho} \partial_{\sigma]} V^\lambda - \Gamma^\lambda{}_{[\sigma\rho]} \left(\partial_\lambda V^\mu + \Gamma^\mu{}_{\lambda\omega} V^\omega\right) \\ &\qquad + \Gamma^\mu{}_{\lambda[\sigma} \partial_{\rho]} V^\lambda + \Gamma^\mu{}_{\lambda[\sigma} \Gamma^\lambda{}_{\rho]\omega} V^\omega \Big\} \end{aligned} \tag{2.2.20}$$

$$= A^\sigma B^\rho \left\{ \partial_{[\sigma} \Gamma^\mu{}_{\rho]\lambda} V^\lambda + \Gamma^\mu{}_{\lambda[\sigma} \Gamma^\lambda{}_{\rho]\omega} V^\omega \right\} \tag{2.2.21}$$

Notice we have used the symmetry of the Christoffel symbols $\Gamma^\mu{}_{\alpha\beta} = \Gamma^\mu{}_{\beta\alpha}$ to arrive at this result. Since A and B are arbitrary, let us observe that the commutator of covariant derivatives acting on a vector field is not a different operator, but rather an algebraic operation:

$$[\nabla_\alpha, \nabla_\beta] V^\mu = R^\mu{}_{\nu\alpha\beta} V^\nu, \tag{2.2.22}$$

$$R^\alpha{}_{\beta\mu\nu} = \partial_{[\mu} \Gamma^\alpha{}_{\nu]\beta} + \Gamma^\alpha{}_{\sigma[\mu} \Gamma^\sigma{}_{\nu]\beta} \tag{2.2.23}$$

$$= \partial_\mu \Gamma^\alpha{}_{\nu\beta} - \partial_\nu \Gamma^\alpha{}_{\mu\beta} + \Gamma^\alpha{}_{\sigma\mu} \Gamma^\sigma{}_{\nu\beta} - \Gamma^\alpha{}_{\sigma\nu} \Gamma^\sigma{}_{\mu\beta}. \tag{2.2.24}$$

Inserting the results in equations (2.2.18) and (2.2.21) into eq. (2.2.17) – we gather, for arbitrary vector fields A and B:

$$\left([\nabla_A, \nabla_B] - \nabla_{[A,B]}\right) V^\mu = R^\mu{}_{\nu\alpha\beta} V^\nu A^\alpha B^\beta. \tag{2.2.25}$$

Problem 2.5. Symmetries of the Riemann tensor Explain why, if a tensor $\Sigma_{\alpha\beta}$ is anti-symmetric in one coordinate system, it has to be anti-symmetric in any other coordinate system. Similarly, explain why, if $\Sigma_{\alpha\beta}$ is symmetric in one coordinate system, it has to be symmetric in any other coordinate system. Compute the Riemann tensor in a locally flat coordinate system and show that

$$R_{\alpha\beta\mu\nu} = \frac{1}{2}\left(\partial_\beta \partial_{[\mu} g_{\nu]\alpha} - \partial_\alpha \partial_{[\mu} g_{\nu]\beta}\right). \tag{2.2.26}$$

From this result, argue that Riemann has the following symmetries:

$$R_{\mu\nu\alpha\beta} = R_{\alpha\beta\mu\nu}, \qquad R_{\mu\nu\alpha\beta} = -R_{\nu\mu\alpha\beta}, \qquad R_{\mu\nu\alpha\beta} = -R_{\mu\nu\beta\alpha}. \tag{2.2.27}$$

This indicates the components of the Riemann tensor are not all independent. Below, we shall see there are additional differential relations (aka "Bianchi identities") between various components of the Riemann tensor.

Finally, use these symmetries to show that

$$[\nabla_\alpha, \nabla_\beta] V_\nu = -R^\mu{}_{\nu\alpha\beta} V_\mu. \tag{2.2.28}$$

Hint: Start with $[\nabla_\alpha, \nabla_\beta](g_{\nu\sigma} V^\sigma)$. □

Ricci tensor and scalar Because of the symmetries of Riemann in eq. (2.2.27), we have $g^{\alpha\beta} R_{\alpha\beta\mu\nu} = -g^{\alpha\beta} R_{\beta\alpha\mu\nu} = -g^{\beta\alpha} R_{\beta\alpha\mu\nu} = 0$; and likewise, $R_{\alpha\beta\mu}{}^{\mu} = 0$. In fact, the Ricci tensor is defined as the sole distinct and non-zero contraction of Riemann:

$$R_{\mu\nu} \equiv R^\sigma{}_{\mu\sigma\nu}. \tag{2.2.29}$$

This is a symmetric tensor, $R_{\mu\nu} = R_{\nu\mu}$, because of eq. (2.2.27); for,

$$R_{\mu\nu} = g^{\sigma\rho} R_{\sigma\mu\rho\nu} = g^{\rho\sigma} R_{\rho\nu\sigma\mu} = R_{\nu\mu}. \tag{2.2.30}$$

Its contraction yields the Ricci scalar

$$\mathcal{R} \equiv g^{\mu\nu} R_{\mu\nu}. \tag{2.2.31}$$

Problem 2.6. Commutator of covariant derivatives on higher rank tensor Prove that

$$\begin{aligned}
&[\nabla_\mu, \nabla_\nu] T^{\alpha_1 \dots \alpha_N}{}_{\beta_1 \dots \beta_M} \\
&= R^{\alpha_1}{}_{\sigma\mu\nu} T^{\sigma\alpha_2 \dots \alpha_N}{}_{\beta_1 \dots \beta_M} + R^{\alpha_2}{}_{\sigma\mu\nu} T^{\alpha_1 \sigma \alpha_3 \dots \alpha_N}{}_{\beta_1 \dots \beta_M} + \dots + R^{\alpha_N}{}_{\sigma\mu\nu} T^{\alpha_1 \dots \alpha_{N-1}\sigma}{}_{\beta_1 \dots \beta_M} \\
&- R^\sigma{}_{\beta_1\mu\nu} T^{\alpha_1 \dots \alpha_N}{}_{\sigma\beta_2 \dots \beta_M} - R^\sigma{}_{\beta_2\mu\nu} T^{\alpha_1 \dots \alpha_N}{}_{\beta_1 \sigma \beta_3 \dots \beta_M} - \dots - R^\sigma{}_{\beta_M\mu\nu} T^{\alpha_1 \dots \alpha_N}{}_{\beta_1 \dots \beta_{M-1}\sigma}.
\end{aligned} \tag{2.2.32}$$

What is $[\nabla_\alpha, \nabla_\beta]\varphi$, where φ is a scalar? □

Problem 2.7. Differential Bianchi identities I Let ω_μ be an arbitrary 1-form. Show that $\nabla_{[\alpha} \nabla_\beta \omega_{\delta]} = 0$. Why does that imply that

$$R^\mu{}_{[\alpha\beta\delta]} = 0? \tag{2.2.33}$$

□

Problem 2.8. Differential Bianchi identities II If $[A, B] \equiv AB - BA$, can you show that the differential operator

$$[\nabla_\alpha, [\nabla_\beta, \nabla_\delta]] + [\nabla_\beta, [\nabla_\delta, \nabla_\alpha]] + [\nabla_\delta, [\nabla_\alpha, \nabla_\beta]] \tag{2.2.34}$$

is actually zero? Why does that imply

$$\nabla_{[\alpha} R^{\mu\nu}{}_{\beta\delta]} = 0? \tag{2.2.35}$$

Using this result, show that

$$\nabla_\sigma R^{\sigma\beta}{}_{\mu\nu} = \nabla_{[\mu} R^{\beta}{}_{\nu]}. \tag{2.2.36}$$

The *Einstein tensor* is defined as

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \mathcal{R}. \tag{2.2.37}$$

From eq. (2.2.36) can you show the divergence-less property of the Einstein tensor, i.e.,

$$\nabla^\mu G_{\mu\nu} = \nabla^\mu \left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \mathcal{R} \right) = 0? \tag{2.2.38}$$

This will be an important property when discussing Einstein's equations for General Relativity.
□

Geodesics As already noted, even in flat spacetime, $\mathrm{d}s^2$ is not positive-definite (cf. (2.1.1)), unlike its purely spatial counterpart. Therefore, when computing the distance along a line in spacetime $z^\mu(\lambda)$, with boundary values $z(\lambda_1) \equiv x'$ and $z(\lambda_2) \equiv x$, we need to take the square root of its absolute value:

$$s = \int_{\lambda_1}^{\lambda_2} \left| g_{\mu\nu}\left(z(\lambda)\right) \frac{\mathrm{d}z^\mu(\lambda)}{\mathrm{d}\lambda} \frac{\mathrm{d}z^\nu(\lambda)}{\mathrm{d}\lambda} \right|^{1/2} \mathrm{d}\lambda. \tag{2.2.39}$$

A geodesic in curved spacetime that joins two points x and x' is a path that extremizes the distance between them. Using an affine parameter to describe the geodesic, i.e., using a λ such that $\sqrt{|g_{\mu\nu}\dot{z}^\mu \dot{z}^\nu|}$ = constant, this amounts to imposing the principle of stationary action on Synge's world function:

$$\sigma(x, x') \equiv \frac{1}{2}(\lambda_2 - \lambda_1) \int_{\lambda_1}^{\lambda_2} g_{\alpha\beta}\left(z(\lambda)\right) \frac{\mathrm{d}z^\alpha}{\mathrm{d}\lambda} \frac{\mathrm{d}z^\beta}{\mathrm{d}\lambda} \mathrm{d}\lambda, \tag{2.2.40}$$

$$z^\mu(\lambda_1) = x'^\mu, \qquad z^\mu(\lambda_2) = x^\mu. \tag{2.2.41}$$

When evaluated on geodesics, eq. (2.2.40) is half the square of the geodesic distance between x and x'. The curved spactime geodesic equation in affine-parameter form which follows from eq. (2.2.40), is

$$\frac{D^2 z^\mu}{\mathrm{d}\lambda^2} \equiv \frac{\mathrm{d}^2 z^\mu}{\mathrm{d}\lambda^2} + \Gamma^\mu{}_{\alpha\beta} \frac{\mathrm{d}z^\alpha}{\mathrm{d}\lambda} \frac{\mathrm{d}z^\beta}{\mathrm{d}\lambda} = 0. \tag{2.2.42}$$

The Lagragian associated with eq. (2.2.40),

$$L_{\rm g} \equiv \frac{1}{2} g_{\mu\nu}(z(\lambda)) \dot{z}^\mu \dot{z}^\nu, \qquad \dot{z}^\mu \equiv \frac{{\rm d}z^\mu}{{\rm d}\lambda}, \tag{2.2.43}$$

not only oftentimes provides a more efficient means of computing the Christoffel symbols, it is a constant of motion. Unlike the curved space case, however, this Lagrangian $L_{\rm g}$ can now be positive, zero, or negative.

- If $\dot{z}^\mu$ is timelike, then by choosing the affine parameter to be proper time ${\rm d}\lambda\sqrt{g_{\mu\nu}\dot{z}^\mu\dot{z}^\nu} = {\rm d}\tau$, we see that the Lagrangian is then set to $L_{\rm g} = 1/2$.
- If $\dot{z}^\mu$ is spacelike, then by choosing the affine parameter to be proper length ${\rm d}\lambda\sqrt{|g_{\mu\nu}\dot{z}^\mu\dot{z}^\nu|} = {\rm d}\ell$, we see that the Lagrangian is then set to $L_{\rm g} = -1/2$.
- If $\dot{z}^\mu$ is null, then the Lagrangian is zero: $L_{\rm g} = 0$.

Formal solution to geodesic equation We may re-write eq. (2.2.42) into an integral equation by simply integrating both sides with respect to the affine parameter λ:

$$v^\mu(\lambda) = v^\mu(\lambda_1) - \int_{z(\lambda_1)}^{z(\lambda)} \Gamma^\mu{}_{\alpha\beta} v^\alpha {\rm d}z^\beta; \tag{2.2.44}$$

where $v^\mu \equiv {\rm d}z^\mu/{\rm d}\lambda$; the lower limit is $\lambda = \lambda_1$; and we have left the upper limit indefinite. The integral on the right hand side can be viewed as an integral operator acting on the tangent vector at $v^\alpha(z(\lambda))$. By iterating this equation infinite number of times – akin to the Born series expansion in quantum mechanics – it is possible to arrive at a formal (as opposed to explicit) solution to the geodesic equation.

Problem 2.9. Synge's World Function In Minkowski Verify that Synge's world function (cf. (2.2.40)) in Minkowski spacetime is

$$\bar{\sigma}(x, x') = \frac{1}{2}(x - x')^2 \equiv \frac{1}{2}\eta_{\mu\nu}(x - x')^\mu (x - x')^\nu, \tag{2.2.45}$$

$$(x - x')^\mu \equiv x^\mu - x'^\mu. \tag{2.2.46}$$

Hint: If we denote the geodesic $z^\mu(0 \leq \lambda \leq 1)$ joining x' to x in Minkowski spacetime, verify that the solution is

$$z^\mu(0 \leq \lambda \leq 1) = x'^\mu + \lambda(x - x')^\mu. \tag{2.2.47}$$

□

Problem 2.10. Show that eq. (2.2.42) takes the same form under re-scaling and constant shifts of the parameter λ. That is, if

$$\lambda = a\lambda' + b, \tag{2.2.48}$$

for constants a and b, then eq. (2.2.42) becomes

$$\frac{D^2 z^\mu}{{\rm d}\lambda'^2} \equiv \frac{{\rm d}^2 z^\mu}{{\rm d}\lambda'^2} + \Gamma^\mu{}_{\alpha\beta}\frac{{\rm d}z^\alpha}{{\rm d}\lambda'}\frac{{\rm d}z^\beta}{{\rm d}\lambda'} = 0. \tag{2.2.49}$$

For the timelike and spacelike cases, this is telling us that proper time and proper length are respectively only defined up to an overall re-scaling and an additive shift. In other words, both the base units and its 'zero' may be altered at will. □

Problem 2.11. Let $v^\mu(x)$ be a vector field defined throughout a given spacetime. Show that the geodesic equation (2.2.42) follows from

$$v^\sigma \nabla_\sigma v^\mu = 0, \tag{2.2.50}$$

i.e., v^μ is parallel transported along itself – provided we recall the 'velocity flow' interpretation of a vector field:

$$v^\mu\left(z(s)\right) = \frac{\mathrm{d}z^\mu}{\mathrm{d}s}. \tag{2.2.51}$$

Parallel transport preserves norm-squared The metric compatibility condition (eq. (2.2.15)) obeyed by the covariant derivative ∇_α can be thought of as the requirement that the norm-squared $v^2 \equiv g_{\mu\nu}v^\mu v^\nu$ of a geodesic vector (v^μ subject to eq. (2.2.50)) be preserved under parallel transport. Can you explain this statement using the appropriate equations?

Non-affine form of geodesic equation Suppose instead

$$v^\sigma \nabla_\sigma v^\mu = \kappa v^\mu. \tag{2.2.52}$$

This is the more general form of the geodesic equation, where the parameter λ is not an affine one. Nonetheless, by considering the quantity $v^\sigma\nabla_\sigma(v^\mu/(v_\nu v^\nu)^p)$, for some real number p, show how eq. (2.2.52) can be transformed into the form in eq. (2.2.50); that is, identify an appropriate v'^μ such that

$$v'^\sigma \nabla_\sigma v'^\mu = 0. \tag{2.2.53}$$

You should comment on how this re-scaling fails when v^μ is null.

Starting from the finite distance integral

$$s \equiv \int_{\lambda_1}^{\lambda_2} \mathrm{d}\lambda \sqrt{|g_{\mu\nu}(z(\lambda))\dot{z}^\mu \dot{z}^\nu|}, \qquad \dot{z}^\mu \equiv \frac{\mathrm{d}z^\mu}{\mathrm{d}\lambda}, \tag{2.2.54}$$

$$z^\mu(\lambda_1) = x', \qquad z^\mu(\lambda_2) = x \tag{2.2.55}$$

show that demanding s be extremized leads to the non-affine geodesic equation

$$\ddot{z}^\mu + \Gamma^\mu{}_{\alpha\beta}\dot{z}^\alpha \dot{z}^\beta = \dot{z}^\mu \frac{\mathrm{d}}{\mathrm{d}\lambda} \ln \sqrt{g_{\alpha\beta}\dot{z}^\alpha \dot{z}^\beta}. \tag{2.2.56}$$

□

Problem 2.12. Conserved quantities along geodesics If p_μ denotes the 'momentum' variable of a geodesic

$$p_\mu \equiv \frac{\partial L_{\mathrm{g}}}{\partial \dot{z}^\mu}, \tag{2.2.57}$$

where L_{g} is defined in eq. (2.2.43), and if ξ^μ is a Killing vector of the same geometry $\nabla_{\{\alpha}\xi_{\beta\}} = 0$, show that

$$\xi^\mu(z(\lambda))p_\mu(\lambda) \tag{2.2.58}$$

is a constant along the geodesic $z^\mu(\lambda)$.

The vector field version of this result goes as follows.

If the geodesic equation $v^\sigma \nabla_\sigma v^\mu = 0$ holds, and if ξ^μ is a Killing vector, then $\xi_\nu v^\nu$ is conserved along the integral curve of v^μ.

Can you demonstrate the validity of this statement? $\square$

Weak Equivalence Principle, "Free-Fall" & Gravity as a Non-Force The universal nature of gravitation – how it appears to act in the same way upon all material bodies independent of their internal composition – is known as the Weak Equivalence Principle. Within non-relativistic physics, the acceleration of some mass M_1 located at $\vec{x}_1$, due to the Newtonian gravitational 'force' exerted by some other mass M_2 at $\vec{x}_2$, is given by

$$M_1 \frac{\mathrm{d}^2 \vec{x}_1}{\mathrm{d}t^2} = -\widehat{n} \frac{G_\mathrm{N} M_1 M_2}{|\vec{x}_1 - \vec{x}_2|^2}, \qquad \widehat{n} \equiv \frac{\vec{x}_1 - \vec{x}_2}{|\vec{x}_1 - \vec{x}_2|}. \tag{2.2.59}$$

Strictly speaking the M_1 on the left hand side is the 'inertial mass', a characterization of the resistance – so to speak – of any material body to being accelerated by an external force. While the M_1 on the right hand side is the 'gravitational mass', describing the strength to which the material body interacts with the gravitational 'force'. Viewed from this perspective, the equivalence principle is the assertion that the inertial and gravitational masses are the same, so that the resulting motion does not depend on them:

$$\frac{\mathrm{d}^2 \vec{x}_1}{\mathrm{d}t^2} = -\widehat{n} \frac{G_\mathrm{N} M_2}{|\vec{x}_1 - \vec{x}_2|^2}. \tag{2.2.60}$$

This Weak Equivalence Principle[6] is one of the primary motivations that led Einstein to recognize gravitation as the manifestation of curved spacetime. The reason why inertial mass equals gravitational mass, for instance, is because material bodies now follow (timelike) geodesics $z^\mu(\tau)$ in curved spacetimes:

$$\frac{D^2 z^\mu}{\mathrm{d}\tau^2} \equiv \frac{\mathrm{d}^2 z^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu{}_{\alpha\beta} \frac{\mathrm{d}z^\alpha}{\mathrm{d}\tau} \frac{\mathrm{d}z^\beta}{\mathrm{d}\tau} = 0; \qquad g_{\mu\nu}(z(\lambda)) \frac{\mathrm{d}z^\mu}{\mathrm{d}\tau} \frac{\mathrm{d}z^\nu}{\mathrm{d}\tau} > 0; \tag{2.2.61}$$

so that their motion only depends on the curved geometry itself and does not depend on their own mass.[7] From this point of view, gravity is no longer a force, and this is why the title of this section reads "...Kinematics in Curved Spacetime".

Note that, strictly speaking, this "gravity-induced-dynamics-as-geodesics" is actually an idealization that applies for material bodies with no internal structure and whose proper sizes are very small compared to the length scale(s) associated with the geometric curvature itself. In reality, all physical systems have internal structure – non-trivial quadrupole moments, spin/rotation, etc. – and may furthermore be large enough that their full dynamics require detailed analysis to understand properly.

Newton vs. Einstein Observe that the Newtonian gravity of eq. (2.2.59) in an instantaneous force, in that the force on body 1 due to body 2 (or, vice versa) changes immediately

[6]See Will [5] arXiv: 1403.7377 for a review on experimental tests of various versions of the Equivalence Principle and other aspects of General Relativity.

[7]If there *were* an external non-gravitational force f^μ, then the covariant Newton's second law for a system of mass M would read: $M D^2 z^\mu / \mathrm{d}\tau^2 = f^\mu$.

when body 2 starts changing its position $\vec{x}_2$ – even though it is located at a finite distance away. However, Special Relativity tells us there ought to be an ultimate speed limit in Nature, i.e., nothing can travel faster than c. This apparent inconsistency between Newtonian gravity and Einstein's Special Relativity is of course a driving motivation that led Einstein to General Relativity. As we shall see shortly, by postulating that the effects of gravitation are in fact the result of residing in a curved spacetime, the Lorentz symmetry responsible for Special Relativity is recovered in any local "free-falling" frame.

Massless particles Finally, this dynamics-as-geodesics also led Einstein to realize – if gravitation does indeed apply universally – that massless particles such as photons, i.e., electromagnetic waves, must also be influenced by the gravitational field too. This is a significant departure from Newton's law of gravity in eq. (2.2.59), which may lead one to suspect otherwise, since $M_{\text{photon}} = 0$. We shall justify this statement more quantitatively below, but to leading order in the JWKB approximation, photons in fact sweep out *null* geodesics $z^\mu(\lambda)$ in curved spacetimes:

$$\frac{D^2 z^\mu}{\mathrm{d}\lambda^2} = 0, \qquad g_{\mu\nu}\left(z(\lambda)\right)\frac{\mathrm{d}z^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}z^\nu}{\mathrm{d}\lambda} = 0. \tag{2.2.62}$$

Locally flat coordinates, Einstein Equivalence Principle & Symmetries We now come to one of the most important features of curved spacetimes. In the neighborhood of a timelike geodesic $y^\mu = (s, \vec{y})$, one may choose *Fermi normal coordinates* $x^\mu \equiv (s, \vec{x})$ such that spacetime appears flat up to distances of $\mathcal{O}(1/|\max R_{\mu\nu\alpha\beta}(y(s))|^{1/2})$:

$$g_{00}(x) = 1 - R_{0a0b}(s)\cdot(x^a - y^a)(x^b - y^b) + \mathcal{O}\left((x-y)^3\right), \tag{2.2.63}$$

$$g_{0i}(x) = -\frac{2}{3}R_{0aib}(s)\cdot(x^a - y^a)(x^b - y^b) + \mathcal{O}\left((x-y)^3\right), \tag{2.2.64}$$

$$g_{ij}(x) = \eta_{ij} - \frac{1}{3}R_{iajb}(s)\cdot(x^a - y^a)(x^b - y^b) + \mathcal{O}\left((x-y)^3\right). \tag{2.2.65}$$

Here $x^0 = s$ is the time coordinate, and is also the proper time of the observer with the trajectory $y^\mu(s)$. Suppose you were placed inside a closed box, so you cannot tell what's outside. Then provided the box is small enough, you will not be able to distinguish between being in "free-fall" in a gravitational field versus being in a completely empty Minkowski spacetime.

As already alluded to in the "Newton vs. Einstein" discussion above, just as the rotation and translation symmetries of flat Euclidean space carried over to a small enough region of curved spaces – the FNC expansion of equations (2.2.63) through (2.2.65) indicates that, within the spacetime neighborhood of a freely-falling observer, any curved spacetime is Lorentz and spacetime-translation symmetric. To sum:

> Physically speaking, in a freely falling frame $\{x^\mu\}$ – i.e., centered along a timelike geodesic at $x = y$ – physics in a curved spacetime is the same as that in flat Minkowski spacetime up to corrections that go at least as
>
> $$\epsilon_{\mathrm{E}} \equiv \frac{\text{Length or inverse mass scale of system}}{\text{Length scale of the spacetime geometric curvature}}. \tag{2.2.66}$$

This is the essence of the equivalence principle that lead Einstein to recognize curved spacetime to be the setting to formulate his General Theory of Relativity. As a simple example, the geodesic y^μ itself obeys the free-particle version of Newton's 2nd law: $\mathrm{d}^2y^i/\mathrm{d}s^2 = 0$.

Problem 2.13. Verify that the coefficients in front of the Riemann tensor in equations (2.2.63), (2.2.64) and (2.2.65) are independent of the spacetime dimension. That is, starting with

$$g_{00}(x) = 1 - A \cdot R_{0a0b}(s) \cdot (x-y)^a (x-y)^b + \mathcal{O}\left((x-y)^3\right), \tag{2.2.67}$$

$$g_{0i}(x) = -B \cdot R_{0aib}(s) \cdot (x-y)^a (x-y)^b + \mathcal{O}\left((x-y)^3\right), \tag{2.2.68}$$

$$g_{ij}(x) = \eta_{ij} - C \cdot R_{iajb}(s) \cdot (x-y)^a (x-y)^b + \mathcal{O}\left((x-y)^3\right), \tag{2.2.69}$$

where A, B, C are unknown constants, compute the Riemann tensor at $x = y$. ☐

Problem 2.14. Gravitational force in a weak gravitational field Consider the following metric:

$$g_{\mu\nu}(t, \vec{x}) = \eta_{\mu\nu} + 2\Phi(\vec{x})\delta_{\mu\nu}, \tag{2.2.70}$$

where $\Phi(\vec{x})$ is time-independent. Assume this is a weak gravitational field, in that $|\Phi| \ll 1$ everywhere in spacetime, and there are no non-gravitational forces. Below, you will show that the linearized Einstein's equations reduce to the familiar Poisson equation

$$\vec{\nabla}^2 \Phi = 4\pi G_{\mathrm{N}} \rho, \tag{2.2.71}$$

where $\rho(\vec{x})$ is the mass/energy density of matter.

Starting from the non-affine form of the action principle

$$\begin{aligned} -Ms &= -M \int_{t_1}^{t_2} \mathrm{d}t \sqrt{g_{\mu\nu} \dot{z}^\mu \dot{z}^\nu}, & \dot{z}^\mu &\equiv \frac{\mathrm{d}z^\mu}{\mathrm{d}t} \\ &= -M \int_{t_1}^{t_2} \mathrm{d}t \sqrt{1 - \vec{v}^2 + 2\Phi(1+\vec{v}^2)}, & \vec{v}^2 &\equiv \delta_{ij} \dot{z}^i \dot{z}^j; \end{aligned} \tag{2.2.72}$$

expand this action to lowest order in $\vec{v}^2$ and Φ and work out the geodesic equation of a 'test mass' M sweeping out some worldline z^μ in such a spacetime. (You should find something very familiar from Classical Mechanics.) Show that, in this non-relativistic limit, Newton's law of gravitation is recovered:

$$\frac{\mathrm{d}^2 z^i}{\mathrm{d}t^2} = -\partial_i \Phi. \tag{2.2.73}$$

We see that, in the weakly curved spacetime of eq. (2.2.70), Φ may be identified as the Newtonian potential. ☐

Geodesic deviation & Tidal Forces We now turn to the derivation of the *geodesic deviation* equation. Consider two geodesics that are infinitesimally close-by. Let both of them be parametrized by λ, so that we may connect one geodesic to the other at the same λ via an infinitesimal vector ξ^μ. We will denote the tangent vector to one of geodesics to be U^μ, such that

$$U^\sigma \nabla_\sigma U^\mu = 0. \tag{2.2.74}$$

Furthermore, we will assume that $[U,\xi]=0$, i.e., U and ξ may be integrated to form a 2D coordinate system in the neighborhood of this pair of geodesics. Then

$$U^\alpha U^\beta \nabla_\alpha \nabla_\beta \xi^\mu = \nabla_U \nabla_U \xi^\mu = -R^\mu{}_{\nu\alpha\beta} U^\nu \xi^\alpha U^\beta. \tag{2.2.75}$$

As its name suggests, this equation tells us how the deviation vector ξ^μ joining two infinitesimally displaced geodesics is accelerated by the presence of spacetime curvature through the Riemann tensor. If spacetime were flat, the acceleration will be zero: two initially parallel geodesics will remain so.

For a macroscopic system, if U^μ is a timelike vector tangent to, say, the geodesic trajectory of its center-of-mass, the geodesic deviation equation (2.2.75) then describes *tidal forces* acting on it. In other words, the relative acceleration between the 'particles' that comprise the system – induced by spacetime curvature – would compete with the system's internal forces.[8]

Derivation of eq. (2.2.75) Starting with the geodesic equation $U^\sigma \nabla_\sigma U^\mu = 0$, we may take its derivative along ξ.

$$\begin{aligned}
\xi^\alpha \nabla_\alpha \left(U^\beta \nabla_\beta U^\mu\right) &= 0, \\
\left(\xi^\alpha \nabla_\alpha U^\beta - U^\alpha \nabla_\alpha \xi^\beta\right) \nabla_\beta U^\mu + U^\beta \nabla_\beta \xi^\alpha \nabla_\alpha U^\mu + \xi^\alpha U^\beta \nabla_\alpha \nabla_\beta U^\mu &= 0 \\
[\xi, U]^\beta \nabla_\beta U^\mu + U^\beta \nabla_\beta (\xi^\alpha \nabla_\alpha U^\mu) - U^\beta \xi^\alpha \nabla_\beta \nabla_\alpha U^\mu + \xi^\alpha U^\beta \nabla_\alpha \nabla_\beta U^\mu &= 0 \\
U^\beta \nabla_\beta (U^\alpha \nabla_\alpha \xi^\mu) &= -\xi^\alpha U^\beta [\nabla_\alpha, \nabla_\beta] U^\mu \\
U^\beta \nabla_\beta (U^\alpha \nabla_\alpha \xi^\mu) &= -\xi^\alpha U^\beta R^\mu{}_{\nu\alpha\beta} U^\nu. \qquad (2.2.76)
\end{aligned}$$

We have repeatedly used $[\xi,U]=0$ to state, for example, $\nabla_U \xi^\rho = U^\sigma \nabla_\sigma \xi^\rho = \xi^\sigma \nabla_\sigma U^\rho = \nabla_\xi U^\rho$. It is also possible to use a more elegant notation to arrive at eq. (2.2.75).

$$\nabla_U U^\mu = 0 \tag{2.2.77}$$

$$\nabla_\xi \nabla_U U^\mu = 0 \tag{2.2.78}$$

$$\nabla_U \underbrace{\nabla_\xi U^\mu}_{=\nabla_U \xi^\mu} + [\nabla_\xi, \nabla_U] U^\mu = 0 \tag{2.2.79}$$

$$\nabla_U \nabla_U \xi^\mu = -R^\mu{}_{\nu\alpha\beta} U^\nu \xi^\alpha U^\beta \tag{2.2.80}$$

On the last line, we have exploited the assumption that $[U,\xi]=0$ to say $[\nabla_\xi, \nabla_U] U^\mu = ([\nabla_\xi, \nabla_U] - \nabla_{[\xi,U]}) U^\mu$ – recall eq. (2.2.25).

Problem 2.15. Geodesic deviation & FNC Argue that all the Christoffel symbols $\Gamma^\alpha{}_{\mu\nu}$ evaluated along the free-falling geodesic in equations (2.2.63)-(2.2.65), namely when $x=y$, vanish. Then argue that all the time derivatives of the Christoffel symbols vanish along y too: $\partial_s^{n\geq 1} \Gamma^\alpha{}_{\mu\nu} = 0$. Why does this imply, denoting $U^\mu \equiv \mathrm{d}y^\mu/\mathrm{d}s$, the geodesic equation

$$U^\nu \nabla_\nu U^\mu = \frac{\mathrm{d}U^\mu}{\mathrm{d}s} = 0? \tag{2.2.81}$$

[8] The first gravitational wave detectors were in fact based on measuring the tidal squeezing and stretching of solid bars of aluminum. They are known as "Weber bars", named after their inventor Joseph Weber.

Next, evaluate the geodesic deviation equation in these Fermi Normal Coordinates (FNC) system. Specifically, show that

$$U^\alpha U^\beta \nabla_\alpha \nabla_\beta \xi^\mu = \frac{\mathrm{d}^2 \xi^\mu}{\mathrm{d}s^2} = -R^\mu{}_{0\nu 0} \xi^\nu. \tag{2.2.82}$$

Why does this imply, *if* the deviation vector is purely spatial at a given $s = s_0$, specifically $\xi^0(s_0) = \mathrm{d}\xi^0/\mathrm{d}s_0 = 0$, then it remains so for all time? □

Let us pause to summarize the physics we have revealed thus far.

> In a curved spacetime, the collective motion of a system of mass M sweeps out a timelike geodesic, whose dynamics is actually independent of M as long as its internal structure can be neglected. In the co-moving frame of an observer situated within this same system, physical laws appear to be the same as that in Minkowski spacetime up to distances of order $1/|\max R_{\widehat{\alpha}\widehat{\beta}\widehat{\mu}\widehat{\nu}}|^{1/2}$. However, once the finite size of the physical system is taken into account, one would find tidal forces exerted upon it due to spacetime curvature itself – this is described by the geodesic deviation eq. (2.2.82).

Killing vectors A geometry is said to enjoy an isometry – or, symmetry – when we perform the following infinitesimal coordinate transformation

$$x^\mu \to x^\mu + \xi^\mu(x) \tag{2.2.83}$$

and find that the geometry is unchanged

$$g_{\mu\nu}(x) \to g_{\mu\nu}(x) + \mathcal{O}\left(\xi^2\right). \tag{2.2.84}$$

Generically, under the infinitesimal coordinate transformation of eq. (2.2.83),

$$g_{\mu\nu}(x) \to g_{\mu\nu}(x) + \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu. \tag{2.2.85}$$

where

$$\nabla_{\{\mu}\xi_{\nu\}} = \xi^\sigma \partial_\sigma g_{\mu\nu} + g_{\sigma\{\mu}\partial_{\nu\}}\xi^\sigma. \tag{2.2.86}$$

If an isometry exists along the integral curve of ξ^μ, it has to obey Killing's equation

$$\nabla_{\{\mu}\xi_{\nu\}} = 0. \tag{2.2.87}$$

In fact, by exponentiating the infinitesimal coordinate transformation, it is possible to show that – if ξ^μ is a Killing vector (i.e., it satisfies eq. (2.2.87)), then an isometry exists along its integral curve. In other words,

> A spacetime geometry enjoys an isometry (aka symmetry) along the integral curve of ξ^μ iff it obeys $\nabla_{\{\alpha}\xi_{\beta\}} = 0$.

In a $d-$dimensional spacetime, there are at most $d(d+1)/2$ Killing vectors. A spacetime that has $d(d+1)/2$ Killing vectors is called *maximally symmetric*. (See Weinberg [3] for a discussion.)

Now let us also consider the second derivatives of ξ^μ. Consider

$$0 = \nabla_\delta \nabla_{\{\alpha}\xi_{\beta\}} \tag{2.2.88}$$

$$= [\nabla_\delta, \nabla_\alpha]\xi_\beta + \nabla_\alpha\nabla_\delta\xi_\beta + [\nabla_\delta, \nabla_\beta]\xi_\alpha + \nabla_\beta\nabla_\delta\xi_\alpha \tag{2.2.89}$$

$$= -R^\lambda{}_{\beta\delta\alpha}\xi_\lambda - \nabla_\alpha\nabla_\beta\xi_\delta - R^\lambda{}_{\alpha\delta\beta}\xi_\lambda - \nabla_\beta\nabla_\alpha\xi_\delta \tag{2.2.90}$$

Because Bianchi says $0 = R^\lambda{}_{[\alpha\beta\delta]} \Rightarrow R^\lambda{}_{\alpha\beta\delta} = R^\lambda{}_{\beta\alpha\delta} + R^\lambda{}_{\delta\beta\alpha}$.

$$0 = -R^\lambda{}_{\beta\delta\alpha}\xi_\lambda - \nabla_\alpha\nabla_\beta\xi_\delta + \left(R^\lambda{}_{\beta\alpha\delta} + R^\lambda{}_{\delta\beta\alpha}\right)\xi_\lambda - \nabla_\beta\nabla_\alpha\xi_\delta \tag{2.2.91}$$

$$0 = -2R^\lambda{}_{\beta\delta\alpha}\xi_\lambda - \nabla_{\{\beta}\nabla_{\alpha\}}\xi_\delta - [\nabla_\beta, \nabla_\alpha]\xi_\delta \tag{2.2.92}$$

$$0 = -2R^\lambda{}_{\beta\delta\alpha}\xi_\lambda - 2\nabla_\beta\nabla_\alpha\xi_\delta \tag{2.2.93}$$

We have arrived at

$$\nabla_\alpha\nabla_\beta\xi_\delta = R^\lambda{}_{\alpha\beta\delta}\xi_\lambda. \tag{2.2.94}$$

Next, we will show that

The commutator of 2 Killing vectors is also a Killing vector.

Let U and V be Killing vectors. Let us compute

$$\begin{aligned}
&\nabla_\alpha(U^\mu\nabla_\mu V_\beta - V^\mu\nabla_\mu U_\beta) + (\alpha \leftrightarrow \beta) \\
&= \nabla_\alpha U^\mu\nabla_\mu V_\beta - \nabla_\alpha V^\mu\nabla_\mu U_\beta + U^\mu\nabla_\alpha\nabla_\mu V_\beta - V^\mu\nabla_\alpha\nabla_\mu U_\beta + (\alpha \leftrightarrow \beta) \\
&= -\nabla_\mu U_\alpha\nabla^\mu V_\beta + \nabla_\mu V_\alpha\nabla^\mu U_\beta + U^\mu\nabla_{[\alpha}\nabla_{\mu]}V_\beta + U^\mu\nabla_\mu\nabla_\alpha V_\beta - V^\mu\nabla_{[\alpha}\nabla_{\mu]}U_\beta - V^\mu\nabla_\mu\nabla_\alpha U_\beta + (\alpha \leftrightarrow \beta) \\
&= -U^\mu R^\sigma{}_{\beta\alpha\mu}V_\sigma + V^\mu R^\sigma{}_{\beta\alpha\mu}U_\sigma + (\alpha \leftrightarrow \beta) \\
&= -U^{[\mu}V^{\sigma]}R_{\sigma\{\beta\alpha\}\mu} = 0
\end{aligned}$$

The $(\alpha \leftrightarrow \beta)$ means we are taking all the terms preceding it and swapping $\alpha \leftrightarrow \beta$. Moreover, we have repeatedly used the fact that $\nabla_\alpha U_\beta = -\nabla_\beta U_\alpha$ and $\nabla_\alpha V_\beta = -\nabla_\beta V_\alpha$.

Problem 2.16. Killing vectors in Minkowski In Minkowski spacetime $g_{\mu\nu} = \eta_{\mu\nu}$, with Cartesian coordinates $\{x^\mu\}$, use eq. (2.2.94) to argue that the most general Killing vector takes the form

$$\xi_\mu = \ell_\mu + \omega_{\mu\nu}x^\nu, \tag{2.2.95}$$

for constant ℓ_μ and $\omega_{\mu\nu}$. (Hint: Think about Taylor expansions.) Then use the Killing equation (2.2.87) to infer that

$$\omega_{\mu\nu} = -\omega_{\nu\mu}. \tag{2.2.96}$$

The ℓ_μ corresponds to infinitesimal spacetime translation and the $\omega_{\mu\nu}$ to infinitesimal Lorentz boosts and rotations. Explain why this implies the following are the Killing vectors of flat spacetime:

$$\partial_\mu \qquad \text{(Generators of spacetime translations)} \tag{2.2.97}$$

and

$$x^{[\mu}\partial^{\nu]} \qquad \text{(Generators of Lorentz boosts or rotations).} \tag{2.2.98}$$

There are d distinct ∂_μ's and (due to their antisymmetry) $(1/2)(d^2-d)$ distinct $x^{[\mu}\partial^{\nu]}$'s. Therefore there are a total of $d(d+1)/2$ Killing vectors in Minkowski – i.e., it is maximally symmetric. □

It might be instructive to check our understanding of rotation and boosts against the 2D case we will work out below via different means. Up to first order in the rotation angle θ, the 2D rotation matrix in eq. (3.0.44) reads

$$\widehat{R}^i{}_j(\theta) = \begin{bmatrix} 1 & -\theta \\ \theta & 1 \end{bmatrix} + \mathcal{O}\left(\theta^2\right). \tag{2.2.99}$$

In other words, $\widehat{R}^i{}_j(\theta) = \delta_{ij} - \theta\epsilon_{ij}$, where ϵ_{ij} is the Levi-Civita symbol in 2D with $\epsilon_{12} \equiv 1$. Applying a rotation of the 2D Cartesian coordinates x^i upon a test (scalar) function f,

$$f(x^i) \to f\left(\widehat{R}^i{}_j x^j\right) = f\left(x^i - \theta\epsilon_{ij}x^j + \mathcal{O}\left(\theta^2\right)\right) \tag{2.2.100}$$

$$= f(\vec{x}) - \theta\epsilon_{ij}x^j\partial_i f(\vec{x}) + \mathcal{O}\left(\theta^2\right). \tag{2.2.101}$$

Since θ is arbitrary, the basic differential operator that implements an infinitesimal rotation of the coordinate system on any Minkowski scalar is

$$-\epsilon_{ij}x^j\partial_i = x^1\partial_2 - x^2\partial_1. \tag{2.2.102}$$

This is the 2D version of eq. (2.2.98) for rotations. As for 2D Lorentz boosts, eq. (3.0.43) below tells us

$$\Lambda^\mu{}_\nu(\xi) = \begin{bmatrix} 1 & \xi \\ \xi & 1 \end{bmatrix} + \mathcal{O}\left(\xi^2\right). \tag{2.2.103}$$

(This ξ is known as *rapidity*.) Here, we have $\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \xi \cdot \epsilon^\mu{}_\nu$, where $\epsilon_{\mu\nu}$ is the Levi-Civita tensor in 2D Minkowski with $\epsilon_{01} \equiv 1$. Therefore, to implement an infinitesimal Lorentz boost on the Cartesian coordinates within a test (scalar) function $f(x^\mu)$, we do

$$f(x^\mu) \to f\left(\Lambda^\mu{}_\nu x^\nu\right) = f\left(x^\mu + \xi\epsilon^\mu{}_\nu x^\nu + \mathcal{O}\left(\xi^2\right)\right) \tag{2.2.104}$$

$$= f(x) - \xi\epsilon_{\nu\mu}x^\nu\partial^\mu f(x) + \mathcal{O}\left(\xi^2\right). \tag{2.2.105}$$

Since ξ is arbitrary, to implement a Lorentz boost of the coordinate system on any Minkowski scalar, the appropriate differential operator is

$$\epsilon_{\mu\nu}x^\mu\partial^\nu = x^0\partial^1 - x^1\partial^0; \tag{2.2.106}$$

which again is encoded within eq. (2.2.98).

Problem 2.17. Co-moving Observers & Rulers In Cosmology We live in a universe that, at the very largest length scales, is described by the following spatially flat Friedmann-Lemaître-Robertson-Walker (FLRW) metric

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a(t)^2\mathrm{d}\vec{x}\cdot\mathrm{d}\vec{x}; \tag{2.2.107}$$

where $a(t)$ describes the relative size of the universe. Enumerate as many constants-of-motion as possible of this geometry. (Hint: Focus on the spatial part of the metric and try to draw a connection with the previous problem.)

In this cosmological context, a co-moving observer is one that does not move spatially, i.e., $\mathrm{d}\vec{x} = 0$. Solve the geodesic swept out by such an observer.

Galaxies A and B are respectively located at $\vec{x}$ and $\vec{x}'$ at a fixed cosmic time t. What is their spatial distance on this constant t slice of spacetime? □

Problem 2.18. In d spacetime dimensions, show that

$$\partial_{[\alpha_1} J^{\mu} \widetilde{\epsilon}_{\alpha_2 \dots \alpha_d]\mu} \tag{2.2.108}$$

is proportional to $\nabla_\sigma J^\sigma$. What is the proportionality factor? (This discussion provides a differential forms based language to write $\mathrm{d}^d x \sqrt{|g|} \nabla_\sigma J^\sigma$.) If $\nabla_\sigma J^\sigma = 0$, what does the Poincaré lemma tell us about eq. (2.2.108)? Find the dual of your result and argue there must an antisymmetric tensor $\Sigma^{\mu\nu}$ such that

$$J^\mu = \nabla_\nu \Sigma^{\mu\nu}. \tag{2.2.109}$$

□

Problem 2.19. Killing identities involving Ricci Prove the following results. If ξ^μ is a Killing vector and $R_{\alpha\beta}$ and $\mathcal{R}$ are the Ricci tensor and scalar respectively, then

$$\xi^\alpha \nabla^\beta R_{\alpha\beta} = 0 \qquad \text{and} \qquad \xi^\alpha \nabla_\alpha \mathcal{R} = 0. \tag{2.2.110}$$

Hints: First use eq. (2.2.94) to show that

$$\nabla^\delta \nabla_\beta \xi_\delta = R^\lambda{}_\beta \xi_\lambda. \tag{2.2.111}$$

Argue why $\xi^\alpha \nabla^\beta R_{\alpha\beta} = \nabla^\beta (\xi^\alpha R_{\alpha\beta})$ and use the above result to show that $\xi^\alpha \nabla^\beta R_{\alpha\beta} = 0$. Then employ the Einstein tensor Bianchi identity $\nabla^\mu G_{\mu\nu} = 0$ to infer that $\xi^\alpha \nabla_\alpha \mathcal{R} = 0$. □

3 Poincaré and Lorentz symmetry

Poincaré and Lorentz symmetries play fundamental roles in our understanding of both classical relativistic physics and quantum theories of elementary particle interactions. In this section, we shall study it in some detail.

The metric of flat spacetime is, in Cartesian coordinates $\{x^\mu\}$,

$$\mathrm{d}s^2 \equiv \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu, \tag{3.0.1}$$

$$\eta_{\mu\nu} \equiv \mathrm{diag}[1, -1, \dots, -1]. \tag{3.0.2}$$

We shall define Poincaré transformations[9] $x(x')$ to be the set of all coordinate transformations that leave the flat spacetime metric invariant:

$$\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = \eta_{\alpha'\beta'}\mathrm{d}x'^\alpha \mathrm{d}x'^\beta. \tag{3.0.3}$$

As we will now proceed to demonstrate, the most general invertible Poincaré transformation is

$$x^\mu = a^\mu + \Lambda^\mu{}_\nu x'^\nu, \tag{3.0.4}$$

where a^μ is a constant vector describing a spacetime translation; and $\Lambda^\mu{}_\nu$ is an arbitrary (spacetime-constant) Lorentz transformation, which in turn is defined as one that leaves $\eta_{\mu\nu}$ invariant in the following manner:

$$\Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \eta_{\mu\nu} = \eta_{\alpha\beta}. \tag{3.0.5}$$

Derivation of eq. (3.0.3)[10] Now, under a coordinate transformation, eq. (3.0.3) reads

$$\eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = \eta_{\mu\nu}\frac{\partial x^\mu}{\partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta}\mathrm{d}x'^\alpha \mathrm{d}x'^\beta = \eta_{\alpha'\beta'}\mathrm{d}x'^\alpha \mathrm{d}x'^\beta. \tag{3.0.6}$$

Let us differentiate both sides of eq. (3.0.6) with respect to x'^σ.

$$\eta_{\mu\nu}\frac{\partial^2 x^\mu}{\partial x'^\sigma \partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta} + \eta_{\mu\nu}\frac{\partial x^\mu}{\partial x'^\alpha}\frac{\partial^2 x^\nu}{\partial x'^\sigma \partial x'^\beta} = 0. \tag{3.0.7}$$

Next, consider symmetrizing $\sigma\alpha$ and anti-symmetrizing $\sigma\beta$.

$$2\eta_{\mu\nu}\frac{\partial^2 x^\mu}{\partial x'^\sigma \partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta} + \eta_{\mu\nu}\frac{\partial x^\mu}{\partial x'^\alpha}\frac{\partial^2 x^\nu}{\partial x'^\sigma \partial x'^\beta} + \eta_{\mu\nu}\frac{\partial x^\mu}{\partial x'^\sigma}\frac{\partial^2 x^\nu}{\partial x'^\alpha \partial x'^\beta} = 0 \tag{3.0.8}$$

$$\eta_{\mu\nu}\frac{\partial^2 x^\mu}{\partial x'^\sigma \partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta} - \eta_{\mu\nu}\frac{\partial^2 x^\mu}{\partial x'^\beta \partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\sigma} = 0 \tag{3.0.9}$$

Since partial derivatives commute, the second term from the left of eq. (3.0.7) vanishes upon anti-symmetrization of $\sigma\beta$. Adding equations (3.0.8) and (3.0.9) hands us

$$3\eta_{\mu\nu}\frac{\partial^2 x^\mu}{\partial x'^\sigma \partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta} + \eta_{\mu\nu}\frac{\partial x^\mu}{\partial x'^\alpha}\frac{\partial^2 x^\nu}{\partial x'^\sigma \partial x'^\beta} = 0. \tag{3.0.10}$$

[9] Poincaré transformations are also sometimes known as inhomogeneous Lorentz transformations.
[10] This argument can be found in Weinberg [3].

Finally, subtracting eq. (3.0.7) from eq. (3.0.10) produces

$$2\eta_{\mu\nu}\frac{\partial^2 x^\mu}{\partial x'^\sigma \partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta} = 0. \tag{3.0.11}$$

Because we have assumed Poincaré transformations are invertible, we may contract both sides with $\partial x'^\beta/\partial x^\kappa$.

$$\eta_{\mu\nu}\frac{\partial^2 x^\mu}{\partial x'^\sigma \partial x'^\alpha}\frac{\partial x^\nu}{\partial x'^\beta}\frac{\partial x'^\beta}{\partial x^\kappa} = \eta_{\mu\nu}\frac{\partial^2 x^\mu}{\partial x'^\sigma \partial x'^\alpha}\delta^\nu_\kappa = 0. \tag{3.0.12}$$

Finally, we contract both sides with $\eta^{\kappa\rho}$:

$$\eta_{\mu'\kappa'}\eta^{\kappa'\rho}\frac{\partial^2 x^\mu}{\partial x'^\sigma \partial x'^\alpha} = \frac{\partial^2 x^\rho}{\partial x'^\sigma \partial x'^\alpha} = 0. \tag{3.0.13}$$

In words: the transformation from x to x' can at most go linearly as x'; it cannot involve higher powers of x'. This implies the form in eq. (3.0.4); plugging the latter into eq. (3.0.6), we recover the necessary definition of the Lorentz transformation in eq. (3.0.5).

> The most general invertible coordinate transformations that leave the Cartesian Minkowski metric invariant involve the (spacetime-constant) Lorentz transformations $\{\Lambda^\mu{}_\alpha\}$ of eq (3.0.5) plus constant spacetime translations.

(Homogeneous) Lorentz Transformations form a Group If $\Lambda^\mu{}_\alpha$ and $\Lambda'^\mu{}_\alpha$ denotes different Lorentz transformations, then notice the composition

$$\Lambda''^\mu{}_\alpha \equiv \Lambda^\mu{}_\sigma \Lambda'^\sigma{}_\alpha \tag{3.0.14}$$

is also a Lorentz transformation. For, keeping in mind the fundamental definition in eq. (3.0.5), we may directly compute

$$\begin{aligned}\Lambda''^\mu{}_\alpha \Lambda''^\nu{}_\beta \eta_{\mu\nu} &= \Lambda^\mu{}_\sigma \Lambda'^\sigma{}_\alpha \Lambda^\nu{}_\rho \Lambda'^\rho{}_\beta \eta_{\mu\nu} \\ &= \Lambda'^\sigma{}_\alpha \Lambda'^\rho{}_\beta \eta_{\sigma\rho} = \eta_{\alpha\beta}.\end{aligned} \tag{3.0.15}$$

To summarize:

> The set of all Lorentz transformations $\{\Lambda^\mu{}_\alpha\}$ satisfying eq. (3.0.5), together with the composition law in eq. (3.0.14) for defining successive Lorentz transformations, form a *Group*.

Proof Let $\Lambda^\mu{}_\alpha$, $\Lambda'^\mu{}_\alpha$ and $\Lambda''^\mu{}_\alpha$ denote distinct Lorentz transformations.

- *Closure* Above, we have just verified that applying successive Lorentz transformations yields another Lorentz transformation; for e.g., $\Lambda^\mu{}_\sigma \Lambda'^\sigma{}_\nu$ and $\Lambda^\mu{}_\sigma \Lambda'^\sigma{}_\rho \Lambda''^\rho{}_\nu$ are Lorentz transformations.
- *Associativity* Because applying successive Lorentz transformations amount to matrix multiplication, and since the latter is associative, that means Lorentz transformations are associative:

$$\Lambda \cdot \Lambda' \cdot \Lambda'' = \Lambda \cdot (\Lambda' \cdot \Lambda'') = (\Lambda \cdot \Lambda') \cdot \Lambda''. \tag{3.0.16}$$

- *Identity* $\delta^\mu{}_\alpha$ is the identity Lorentz transformation:

$$\delta^\mu{}_\sigma \Lambda^\sigma{}_\nu = \Lambda^\mu{}_\sigma \delta^\sigma{}_\nu = \Lambda^\mu{}_\nu, \tag{3.0.17}$$

  and

$$\delta^\mu{}_\alpha \delta^\nu{}_\beta \eta_{\mu\nu} = \eta_{\alpha\beta}. \tag{3.0.18}$$

- *Inverse* Let us take the determinant of both sides of eq. (3.0.5) – by viewing the latter as matrix multiplication, we have $\Lambda^T \cdot \eta \cdot \Lambda = \eta$, which in turn means

$$(\det \Lambda)^2 = 1 \qquad \Rightarrow \qquad \det \Lambda = \pm 1. \tag{3.0.19}$$

  Here, we have recalled $\det A^T = \det A$ for any square matrix A. Since the determinant of Λ is strictly non-zero, what eq. (3.0.19) teaches us is that Λ is always invertible: Λ^{-1} is guaranteed to exist. What remains is to check that, if Λ is a Lorentz transformation, so is Λ^{-1}. Starting with the matrix form of eq. (3.0.5), and utilizing $(\Lambda^{-1})^T = (\Lambda^T)^{-1}$,

$$\Lambda^T \eta \Lambda = \eta \tag{3.0.20}$$

$$(\Lambda^T)^{-1} \Lambda^T \eta \Lambda \Lambda^{-1} = (\Lambda^T)^{-1} \cdot \eta \cdot \Lambda^{-1} \tag{3.0.21}$$

$$\eta = (\Lambda^{-1})^T \cdot \eta \cdot \Lambda^{-1}. \tag{3.0.22}$$

Lorentzian 'inner product' is preserved That Λ is a Lorentz transformation means it is a linear operator that preserves the Lorentzian inner product. For suppose v and w are arbitrary vectors, the inner product of $v' \equiv \Lambda v$ and $w' \equiv \Lambda w$ is that between v and w.

$$v' \cdot w' \equiv \eta_{\alpha\beta} v'^\alpha w'^\beta = \eta_{\alpha\beta} \Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu v^\mu w^\nu \tag{3.0.23}$$

$$= \eta_{\mu\nu} v^\mu w^\nu = v \cdot w. \tag{3.0.24}$$

This is very much analogous to rotations in $\mathbb{R}^D$ being the linear transformations that preserve the Euclidean inner product between spatial vectors: $\vec{v} \cdot \vec{w} = \vec{v}' \cdot \vec{w}'$ for all $\widehat{R}^T \widehat{R} = \mathbb{I}_{D\times D}$, where $\vec{v}' \equiv \widehat{R}\vec{v}$ and $\vec{w}' \equiv \widehat{R}\vec{w}$.

Problem 3.1. 4D Lorentz Group and $\mathbf{SL}_{2,\mathbb{C}}$ Define $\{\sigma^\mu\}$ to be the basis set of 2×2 complex matrices formed by the 2×2 identity matrix together with the Pauli matrices, namely

$$\sigma^0 \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \sigma^1 \equiv \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \sigma^2 \equiv \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \qquad \sigma^3 \equiv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{3.0.25}$$

Now let $p_\mu \equiv (p_0, p_1, p_2, p_3)$ be a 4-component collection of real numbers, and verify that

$$\det p_\mu \sigma^\mu = \eta^{\mu\nu} p_\mu p_\nu \equiv p^2. \tag{3.0.26}$$

Next, consider the following transformation,

$$p_\mu \sigma^\mu \to L^\dagger \cdot p_\mu \sigma^\mu \cdot L, \tag{3.0.27}$$

where L is some arbitrary 2×2 complex matrix. (This transformation preserves the Hermitian nature of $p_\mu \sigma^\mu$ for real p_μ.) Then consider taking their determinant:

$$\det[p_\mu \sigma^\mu] \to \det\left[L^\dagger \cdot p_\mu \sigma^\mu \cdot L\right] \tag{3.0.28}$$

What property must L obey in order that this leaves the determinant invariant, i.e.,

$$\det[p_\mu \sigma^\mu] = \det\left[L^\dagger \cdot p_\mu \sigma^\mu \cdot L\right] = p^2? \tag{3.0.29}$$

Argue that the set of all L's obeying eq. (3.0.29), with $\det L = 1$ (this is the 'S'≡'special' in $\mathrm{SL}_{2,\mathbb{C}}$), forms a group. □

We wish to study in some detail what the most general form $\Lambda^\mu{}_\alpha$ may take. To this end, we shall do so by examining how it acts on some arbitrary vector field v^μ; even though this section deals with Minkowski spacetime, this v^μ may also be viewed as a vector in a curved spacetime written in an orthonormal basis.

Rotations Let us recall that any spatial vector v^i may be rotated to point along the 1−axis while preserving its Euclidean length. That is, there is always a $\widehat{R}$, obeying $\widehat{R}^T\widehat{R} = \mathbb{I}$ such that

$$\widehat{R}^i{}_j v^j \doteq \pm\,|\vec{v}|(1,0,\dots,0)^T, \qquad |\vec{v}| \equiv \sqrt{\delta_{ij}v^i v^j}. \tag{3.0.30}$$

[11]Conversely, since $\widehat{R}$ is necessarily invertible, any spatial vector v^i can be obtained by rotating it from $|\vec{v}|(1,\vec{0}^T)$. Moreover, in $D+1$ notation, these rotation matrices can be written as

$$\widehat{R}^\mu{}_\nu \doteq \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & \widehat{R}^i{}_j \end{bmatrix} \tag{3.0.31}$$

$$\widehat{R}^0{}_\nu v^\nu = v^0, \tag{3.0.32}$$

$$\widehat{R}^i{}_\nu v^\nu = \widehat{R}^i{}_j v^j = (v^0, \pm|\vec{v}|, 0, \dots, 0)^T. \tag{3.0.33}$$

These considerations tell us, if we wish to study Lorentz transformations that are *not* rotations, we may reduce their study to the $(1+1)$D case. To see this, we first observe that

$$\Lambda \begin{bmatrix} v^0 \\ v^1 \\ \vdots \\ v^D \end{bmatrix} = \Lambda \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & \widehat{R} \end{bmatrix} \begin{bmatrix} v^0 \\ \pm|\vec{v}| \\ \vec{0} \end{bmatrix}. \tag{3.0.34}$$

And if the result of this matrix multiplication yields non-zero spatial components, namely $(v'^0, v'^1, \dots, v'^D)^T$, we may again find a rotation matrix $\widehat{R}'$ such that

$$\Lambda \begin{bmatrix} v^0 \\ v^1 \\ \vdots \\ v^D \end{bmatrix} = \begin{bmatrix} v'^0 \\ v'^1 \\ \vdots \\ v'^D \end{bmatrix} = \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & \widehat{R}' \end{bmatrix} \begin{bmatrix} v'^0 \\ \pm|\vec{v}'| \\ \vec{0} \end{bmatrix}. \tag{3.0.35}$$

[11]This $\widehat{R}$ is not unique: for example, by choosing another rotation matrix $\widehat{R}''$ that only rotates the space orthogonal to v^i, $\widehat{R}\widehat{R}''\vec{v}$ and $\widehat{R}\vec{v}$ both yield the same result.

At this point, we have reduced our study of Lorentz transformations to

$$\begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & \widehat{R}'^T \end{bmatrix} \Lambda \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & \widehat{R} \end{bmatrix} \begin{bmatrix} v^0 \\ v^1 \\ \vec{0} \end{bmatrix} \equiv \Lambda' \begin{bmatrix} v^0 \\ v^1 \\ \vec{0} \end{bmatrix} = \begin{bmatrix} v'^0 \\ v'^1 \\ \vec{0} \end{bmatrix}. \tag{3.0.36}$$

Because Λ was arbitrary so is Λ', since one can be gotten from another via rotations.

Time Reversal & Parity Flips Suppose the time component of the vector v^μ were negative ($v^0 < 0$), we may write it as

$$\begin{bmatrix} -|v^0| \\ \vec{v} \end{bmatrix} = \widehat{T} \begin{bmatrix} |v^0| \\ \vec{v} \end{bmatrix}, \qquad \widehat{T} \equiv \begin{bmatrix} -1 & \vec{0}^T \\ \vec{0} & \mathbb{I}_{D\times D} \end{bmatrix}; \tag{3.0.37}$$

where $\widehat{T}$ is the time reversal matrix since it reverses the sign of the time component of the vector. You may readily check that $\widehat{T}$ itself is a Lorentz transformation in that it satisfies $\widehat{T}^T \eta \widehat{T} = \eta$.

Problem 3.2. Parity flip of the ith axis Suppose we wish to flip the sign of the ith spatial component of the vector, namely $v^i \to -v^i$. You can probably guess, this may be implemented via the diagonal matrix with all entries set to unity, except the ith component – which is set instead to -1.

$$_i\widehat{P}^\mu{}_\nu v^\nu = v^\mu, \qquad \mu \neq i, \tag{3.0.38}$$

$$_i\widehat{P}^i{}_\nu v^\nu = -v^i, \tag{3.0.39}$$

$$_i\widehat{P} \equiv \text{diag}[1, 1, \dots, 1, \underbrace{-1}_{(i+1)\text{th component}}, 1, \dots, 1]. \tag{3.0.40}$$

If $\widehat{R}^\mu{}_\nu$ is the rotation matrix that does $\widehat{R}^\mu{}_\nu v^\nu = (v^0, \pm|\vec{v}|, 0, \dots, 0)$, argue that

$$_i\widehat{P} = \widehat{R}^T \cdot {}_1\widehat{P} \cdot \widehat{R}. \tag{3.0.41}$$

Is $_i\widehat{P}$ a Lorentz transformation? □

Lorentz Boosts Focusing on the 2D case, we now turn to the transformations that would mix time and space components, and yet leave the metric of spacetime $\eta_{\mu\nu} = \text{diag}[1, -1]$ invariant. (Neither time reversal, parity flips, nor spatial rotations mix time and space.) This is what revolutionized humanity's understanding of spacetime at the beginning of the 1900's: inspired by the fact that the speed of light is the same in all inertial frames, Einstein discovered *Special Relativity*, that the space and time coordinates of one frame have to become intertwined when being translated to those in another frame. We will turn this around later when discussing Maxwell's equations: the constancy of the speed of light in all inertial frames is in fact a consequence of the Lorentz covariance of the latter.

Problem 3.3. We wish to find a 2×2 matrix Λ that obeys $\Lambda^T \cdot \eta \cdot \Lambda = \eta$, where $\eta_{\mu\nu} = \text{diag}[1, -1]$. By examining the diagonal terms of $\Lambda^T \cdot \eta \cdot \Lambda = \eta$, show that

$$\Lambda \doteq \begin{bmatrix} \sigma_1 \cosh(\xi_1) & \sigma_2 \sinh(\xi_2) \\ \sigma_3 \sinh(\xi_1) & \sigma_4 \cosh(\xi_2) \end{bmatrix}, \tag{3.0.42}$$

where the $\sigma_{1,2,3,4}$ are either $+1$ or -1; altogether, there are 8 choices of signs. (Hint: $x^2-y^2=c^2$, for constant c, describes a hyperbola on the (x,y) plane.) From the off diagonal terms of $\Lambda^T \cdot \eta \cdot \Lambda = \eta$, argue that either $\xi_1 = \xi_2 \equiv \xi$ or $\xi_1 = -\xi_2 = \xi$. Up to multiplications by parity flips and/or time reversal matrices, prove that the most general 2D Lorentz boost is therefore

$$\Lambda^\mu{}_\nu(\xi) = \begin{bmatrix} \cosh(\xi) & \sinh(\xi) \\ \sinh(\xi) & \cosh(\xi) \end{bmatrix}. \tag{3.0.43}$$

This ξ is known as *rapidity*. In 2D, the rotation matrix is

$$\widehat{R}^i{}_j(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}; \tag{3.0.44}$$

and therefore rapidity ξ is to the Lorentz boost in eq. (3.0.43) what the angle θ is to rotation $\widehat{R}^i{}_j(\theta)$. □

To understand the meaning of the rapidity ξ, let us consider applying it to an arbitrary 2D vector U^μ.

$$U' \equiv \Lambda \cdot U = \begin{bmatrix} U^0\cosh(\xi) + U^1\sinh(\xi) \\ U^1\cosh(\xi) + U^0\sinh(\xi) \end{bmatrix}. \tag{3.0.45}$$

Timelike case Suppose U were timelike, $(U^0)^2 > (U^1)^2 \Rightarrow |U^0/U^1| > 1$. Then it is not possible to find a ξ such that $U'^0 = 0$, because that would amount to solving $\tanh(\xi) = -U^0/U^1$ but tanh lies between -1 and $+1$ while $-U^0/U^1$ is either less than -1 or greater than $+1$. On the other hand, it does mean we may solve for ξ that would set the spatial component to zero: $\tanh(\xi) = -U^1/U^0$. Recall that tangent vectors may be interpreted as the derivative of the spacetime coordinates with respect to some parameter λ, namely $U^\mu \equiv \mathrm{d}x^\mu/\mathrm{d}\lambda$. Therefore

$$\frac{U^1}{U^0} = \frac{\mathrm{d}x^1}{\mathrm{d}x^0} \equiv v \tag{3.0.46}$$

is the velocity in the frame $\{x^\mu\}$. Starting from $\tanh(\xi) = -v$, some algebra would then hand us (cf. eq. (3.0.43))

$$\cosh(\xi) = \gamma \equiv \frac{1}{\sqrt{1-v^2}}, \tag{3.0.47}$$

$$\sinh(\xi) = -\gamma \cdot v = -\frac{v}{\sqrt{1-v^2}}, \tag{3.0.48}$$

$$\Lambda^\mu{}_\nu = \begin{bmatrix} \gamma & -\gamma \cdot v \\ -\gamma \cdot v & \gamma \end{bmatrix}. \tag{3.0.49}$$

This in turn yields

$$U' = \left(\mathrm{sgn}(U^0)\sqrt{\eta_{\mu\nu}U^\mu U^\nu}, 0\right)^T; \tag{3.0.50}$$

leading us to interpret the $\Lambda^\mu{}_\nu$ we have found in eq. (3.0.49) as the boost that bring observers to the frame where the flow associated with U^μ is 'at rest'. (Note that, if $U^\mu = \mathrm{d}x^\mu/\mathrm{d}\tau$, where τ is proper time, then $\eta_{\mu\nu}U^\mu U^\nu = 1$.)

As an important aside, we may generalize the two-dimensional Lorentz boost in eq. (3.0.49) to $D-$dimensions. One way to do it, is to simply append to the 2D Lorentz-boost matrix a $(D-2)\times(D-2)$ identity matrix (that leaves the $2-$ through $D-$spatial components unaltered) in a block diagonal form:

$$\Lambda^\mu{}_\nu \stackrel{?}{=} \left[\begin{array}{ccc} \gamma & -\gamma\cdot v & 0 \\ -\gamma\cdot v & \gamma & 0 \\ 0 & 0 & \mathbb{I}_{(D-2)\times(D-2)} \end{array}\right]. \tag{3.0.51}$$

But this is not doing much: we are still only boosting in the $1-$direction. What if we wish to boost in v^i direction, where v^i is now some arbitrary spatial vector? To this end, we may view the $(0,1)$ and $(1,0)$ components of eq. (3.0.49) to generalize to the spatial vectors $\Lambda^0{}_i$ and $\Lambda^i{}_0$ parallel to v^i. Whereas the $(1,1)$ component of eq. (3.0.49) to be acting on the 1D space parallel to v^i, namely the operator $v^i v^j/\vec{v}^2$. (As a check: When $v^i = v(1,\vec{0})$, $v^i v^j/\vec{v}^2 = \delta^i_1\delta^j_1$.) The identity operator acting on the orthogonal $(D-2)\times(D-2)$ space, i.e. the analog of $\mathbb{I}_{(D-2)\times(D-2)}$ in eq. (3.0.51), is $\delta^{ij} - v^i v^j/\vec{v}^2$. Altogether, the Lorentz boost in the v^i direction is given by

$$\Lambda^\mu{}_\nu(\vec{v}) \doteq \left[\begin{array}{cc} \gamma & -\gamma v^i \\ -\gamma v^i & \gamma\frac{v^i v^j}{\vec{v}^2} + \left(\delta^{ij} - \frac{v^i v^j}{\vec{v}^2}\right) \end{array}\right], \qquad \vec{v}^2 \equiv \delta_{ab} v^a b^b. \tag{3.0.52}$$

It may be worthwhile to phrase this discussion in terms of the Cartesian coordinates $\{x^\mu\}$ and $\{x'^\mu\}$ parametrizing the two inertial frames. What we have shown is that the Lorentz boost in eq. (3.0.52) describes

$$U'^\mu = \Lambda^\mu{}_\nu(\vec{v}) U^\nu, \tag{3.0.53}$$

$$U^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}, \qquad U'^\mu = \frac{\mathrm{d}x'^\mu}{\mathrm{d}\lambda} = \left(\mathrm{sgn}(U^0)\sqrt{\eta_{\mu\nu}U^\mu U^\nu}, 0\right)^T. \tag{3.0.54}$$

λ is the intrinsic 1D coordinate parametrizing the worldlines, and by definition does not alter under Lorentz boost. The above statement is therefore equivalent to

$$\mathrm{d}x'^\mu = \Lambda^\mu{}_\nu(\vec{v})\mathrm{d}x^\nu, \tag{3.0.55}$$

$$x'^\mu = \Lambda^\mu{}_\nu(\vec{v}) x^\nu. \tag{3.0.56}$$

Problem 3.4. Lorentz boost in $(D+1)-$dimensions If $v^\mu \equiv (1, v^i)$, check via a direction calculation that the $\Lambda^\mu{}_\nu$ in eq. (3.0.52) produces a $\Lambda^\mu{}_\nu v^\nu$ that has no non-trivial spatial components. Also check that eq. (3.0.52) is, in fact, a Lorentz transformation. What is $\Lambda^\mu{}_\sigma(\vec{v})\Lambda^\sigma{}_\nu(-\vec{v})$?

Spacelike case Suppose U were spacelike, $(U^0)^2 < (U^1)^2 \Rightarrow |U^1/U^0| = |\mathrm{d}x^1/\mathrm{d}x^0| \equiv |v| > 1$. Then, it is not possible to find a ξ such that $U'^1 = 0$, because that would amount to solving $\tanh(\xi) = -U^1/U^0$, but tanh lies between -1 and $+1$ whereas $-U^1/U^0 = -v$ is either less than -1 or greater than $+1$. On the other hand, it is certainly possible to have $U'^0 = 0$. Simply do $\tanh(\xi) = -U^0/U^1 = -1/v$. Similar algebra to the timelike case then hands us

$$\cosh(\xi) = \left(1 - v^{-2}\right)^{-1/2}, \tag{3.0.57}$$

$$\sinh(\xi) = -(1/v)\left(1 - v^{-2}\right)^{-1/2}, \tag{3.0.58}$$

$$U' = \left(0, \mathrm{sgn}(U^1)\sqrt{-\eta_{\mu\nu}U^\mu U^\nu}\right)^T. \tag{3.0.59}$$

We may interpret U'^μ and U^μ as infinitesimal vectors joining the same pair of spacetime points but in their respective frames. Specifically, U'^μ are the components in the frame where the pair lies on the same constant-time surface ($U'^0 = 0$). While U^μ are the components in a boosted frame.

Null (aka lightlike) case If U were null, that means $(U^0)^2 = (U^1)^2$, which in turn means $U^\mu = \omega(1, \pm 1)$ for some real number ω. Upon a Lorentz boost, eq. (3.0.45) tells us

$$U' \equiv \Lambda \cdot U = \omega \begin{bmatrix} \cosh(\xi) \pm \sinh(\xi) \\ \sinh(\xi) \pm \cosh(\xi) \end{bmatrix}. \tag{3.0.60}$$

As we shall see below, if U^μ describes the $d-$momentum of a photon, so that $|\omega|$ is its frequency in the un-boosted frame, the $\cosh(\xi) \pm \sinh(\xi)$ describes the photon's red- or blue-shift in the boosted frame.

> **Summary** Our analysis of the group of matrices $\{\Lambda\}$ obeying $\Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu \eta_{\alpha\beta} = \eta_{\mu\nu}$ reveals that these Lorentz transformations consists of: time reversals, parity flips, rotations and Lorentz boosts. A timelike vector can always be Lorentz-boosted so that all its spatial components are zero; while a spacelike vector can always be Lorentz-boosted so that its time component is zero.

Problem 3.5. Null, spacelike vs. timelike Do null vectors form a vector space? Similarly, do spacelike or timelike vectors form a vector space?

Problem 3.6. Determinants and discontinuities What are the determinants of the time reversal $\widehat{T}$ and parity flips $\{{}_i\widehat{P}\}$ matrices? What is the determinant of the Lorentz boost matrix in eq. (3.0.43)? Hint: Your answers should tells us, as long as the determinants of Lorentz transformations are real, time-reversals and parity flips cannot be continuously connected to the identity transformation. Whereas, when the rapidity ξ and rotation angle θ are set to zero, Lorentz boosts and rotations respectively become the identity in a continuous manner.

Problem 3.7. Non-singular Coordinate transformations form a group Let us verify explicitly that the Jacobians associated with general non-singular coordinate transformations form a group. Specifically, let us consider transforming from the coordinate system x^α to y^μ, and assume x^α in terms of y^μ has been provided (i.e., $x^\alpha(y^\mu)$ is known). We may also proceed to consider transforming to a third coordinate system, from y^μ to z^κ.

- *Closure* Denote the Jacobian as, for e.g., $\mathcal{J}^\alpha{}_\mu[x \to y] \equiv \partial x^\alpha / \partial y^\mu$. If we define the group operation as simply that of matrix multiplication, verify that

  $$\mathcal{J}^\alpha{}_\sigma[x \to y] \mathcal{J}^\sigma{}_\nu[y \to z] = \mathcal{J}^\alpha{}_\nu[x \to z]. \tag{3.0.61}$$

  In words: multiplying the transformation matrix bringing us from x to y followed by that from y to z, yields the Jacobian that brings us from x directly to z. This composition law is what we would need, if the group operation is to implement coordinate transformations.

- *Associativity* Explain why the composition law for Jacobians is associative.

- *Identity* What is the identity Jacobian? What is the most general coordinate transformation it corresponds to?

- *Inverse* By non-singular, we mean $\det \mathcal{J}^{\alpha}{}_{\mu} \neq 0$. What does this imply about the existence of the inverse $(\mathcal{J}^{-1})^{\alpha}{}_{\mu}$?

□

4 Local Conservation Laws

Non-relativistic You would be rightly shocked if you had stored a sealed tank of water on your rooftop only to find its contents gradually disappearing over time – the total mass of water ought to be a constant. Assuming a flat space geometry, if you had connected the tank to two pipes, one that pumps water into the tank and the other pumping water out of it, the rate of change of the total mass of the water

$$M \equiv \int_{\text{tank}} \rho(t, \vec{x}) \mathrm{d}^3\vec{x} \tag{4.0.1}$$

in the tank – where t is time, $\vec{x}$ are Cartesian coordinates, and $\rho(t, \vec{x})$ is the water's mass density – is

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\text{tank}} \rho \mathrm{d}^3\vec{x} = -\left(\int_{\text{cross section of 'in' pipe}} + \int_{\text{cross section of 'out' pipe}} \right) \mathrm{d}^2\vec{\Sigma} \cdot (\rho\vec{v}). \tag{4.0.2}$$

Note that $\mathrm{d}^2\vec{\Sigma}$ points *outwards* from the tank, so at the 'in' pipe-tank interface, if the water were indeed following into the pipe, $-\mathrm{d}^2\vec{\Sigma} \cdot (\rho\vec{v}) > 0$ and its contribution to the rate of increase is positive. At the 'out' pipe-tank interface, if the water were indeed following out of the pipe, $-\mathrm{d}^2\vec{\Sigma} \cdot (\rho\vec{v}) < 0$. If we apply Gauss' theorem,

$$\int_{\text{tank}} \dot{\rho} \mathrm{d}^3\vec{x} = -\int_{\text{tank}} \mathrm{d}^3\vec{x} \vec{\nabla} \cdot (\rho\vec{v}). \tag{4.0.3}$$

If we applied the same sort of reasoning to any infinitesimal packet of fluid, with some local mass density ρ, we would find the following local conservation law

$$\dot{\rho} = -\partial_i \left(\rho \cdot v^i \right). \tag{4.0.4}$$

This is a "local" conservation law in the sense that mass cannot simply vanish from one location and re-appear a finite distance away, without first flowing to a neighboring location.

Relativistic We have implicitly assumed a non-relativistic system, where $|\vec{v}| \ll 1$. This is an excellent approximation for most hydrodynamics problems. Strictly speaking, however, relativistic effects – length contraction, in particular – imply that mass density is not a Lorentz scalar. If we define $\rho(t, \vec{x})$ to be the mass density at $(t, \vec{x})$ in a frame instantaneously at rest (aka 'co-moving') with the fluid packet, then the mass density current that is a locally conserved Lorentz vector is given by

$$J^\mu(t, \vec{x}) \equiv \rho(t, \vec{x}) v^\mu(t, \vec{x}). \tag{4.0.5}$$

Along its integral curve v^μ should be viewed as the proper velocity $\mathrm{d}(t, \vec{x})^\mu / \mathrm{d}\tau$ of the fluid packet, where τ is the latter's proper time. Moreover, as long as the velocity v^μ is timelike, which is certainly true for fluids, let us recall it is always possible to find a (local) Lorentz transformation $\Lambda^\mu{}_\nu(t, \vec{x})$ such that

$$(1, \vec{0})^\mu \equiv v'^\mu = \Lambda^\mu{}_\nu(t, \vec{x}) v^\nu(t, \vec{x}). \tag{4.0.6}$$

The local conservation law obeyed by this relativistically covariant current J^μ is now (in Cartesian coordinates)

$$\partial_\mu J^\mu = 0; \tag{4.0.7}$$

which in turn is a Lorentz invariant statement. Total mass M in a given global inertial frame at a fixed time t is

$$M \equiv \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x} J^0. \tag{4.0.8}$$

To show it is a constant, we take the time derivative, and employ eq. (4.0.7):

$$\dot{M} = \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x}\partial_0 J^0 = -\int_{\mathbb{R}^3} \mathrm{d}^3\vec{x}\partial_i J^i. \tag{4.0.9}$$

The divergence theorem tells us that this is equal to the flux of J^i at spatial infinity. But there is no J^i at spatial infinity for physically realistic – i.e., isolated – systems.

Problem 4.1. Local Conservation In Curved Spacetimes The equivalence principle tells us: in a local free-falling frame, physics should be nearly identical to that in Minkowski spacetime. In a generic curved spacetime, the local conservation laws we are examining in this section becomes

$$\nabla_\mu J^\mu = \frac{\partial_\mu \left(\sqrt{|g|} J^\mu\right)}{\sqrt{|g|}} = 0. \tag{4.0.10}$$

(Why does the first equality hold?) Argue that this reduces to eq. (4.0.7) in a FNC coordinate system described by equations (2.2.63)-(2.2.65). □

Problem 4.2. Electric Charge Conservation in a (Spatially Flat) Expanding Universe Let us consider a $d-$dimensional universe described by the line element

$$\mathrm{d}s^2 = a(\eta)^2 \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu. \tag{4.0.11}$$

(The $x^0 = \eta$ in $a(\eta)$ is the time coordinate, not to be confused with the flat metric.) When $d = 4$, this appears to describe our universe at the largest length scales.

Let us now examine an electric current J^μ inhabiting such a universe, where $\nabla_\mu J^\mu = 0$. First verify that the orthonormal frame fields describing a family of co-moving observers is given by

$$\varepsilon^{\widehat{\alpha}}{}_\mu = a(\eta)\delta^\alpha_\mu. \tag{4.0.12}$$

According to this family of observers, they measure a local electric charge density of $J^{\widehat{0}}$ and current flow $J^{\widehat{i}}$. On a constant time η hypersurface, the induced metric can be obtained from eq. (4.0.11) by setting $\mathrm{d}x^0 = 0$ and multiplying throughout by a -1 (so as to get positive distances):

$$H_{ij} = a^2\delta_{ij}. \tag{4.0.13}$$

Define the total charge on a constant η hypersurface as

$$Q \equiv \int_{\mathbb{R}^{d-1}} \mathrm{d}^{d-1}\vec{x}\sqrt{\det H_{ab}} J^{\widehat{0}}. \tag{4.0.14}$$

Show that Q is actually independent of $x^0 = \eta$. □

5 Electromagnetism in Minkowski Spacetime

A symmetry is an operation that leaves an mathematical object or physical system unchanged. Much of our progress in understanding fundamental physics involved the discovery of symmetries that underlie the basic laws of Nature. The Lorentz invariance we have explored in §(3), for instance, is a *spacetime symmetry*: inertial observers moving at constant velocities with respect to each other would find the laws of physics in their respective frames to take the same form. To translate results in one frame to another, all one has to do is to transform their respective spacetime coordinates through eq. .[12] There are also non-spacetime (aka 'internal') symmetries.

In this section we will discuss in some detail Minkowski spacetime electromagnetism to illustrate both its Lorentz and gauge symmetries. It will also provide us the opportunity to introduce the action principle, which is key formulating both classical and quantum field theories.

Maxwell & Lorentz We begin with Maxwell's equations in the following Lorentz covariant form, written in Cartesian coordinates $\{x^\mu\}$ so that $g_{\mu\nu} = \eta_{\mu\nu}$:

$$\partial_\mu F^{\mu\nu} = J^\nu, \qquad \partial_{[\mu} F_{\alpha\beta]} = 0, \qquad F_{\mu\nu} = -F_{\nu\mu}. \tag{5.0.1}$$

The J^μ is the electromagnetic current, where in a fixed inertial frame $J^\mu \equiv \rho v^\mu$. Assuming J^μ is timelike, v^μ is its $d-$velocity; and $\rho \equiv J^\mu v_\mu$ is the electric charge in the (local) rest frame where $v^\mu = \delta^\mu_0$. Defined this way, ρ is a Lorentz scalar and J^μ is a Lorentz vector since v^μ is a Lorentz vector. It is then reasonable to suppose $F_{\mu\nu}$ is a rank-2 Lorentz tensor. Specifically, let two inertial frames $\{x^\mu\}$ and $\{x'^\mu\}$ be related via the Lorentz transformation

$$x^\mu = \Lambda^\mu{}_\alpha x'^\alpha, \qquad \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \eta_{\mu\nu} = \eta_{\alpha\beta}. \tag{5.0.2}$$

Then the Faraday tensor transforms as

$$F_{\alpha'\beta'}(x') = F_{\mu\nu}\left(x(x') = \Lambda \cdot x'\right) \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \tag{5.0.3}$$

Its derivatives are also Lorentz covariant, for keeping in mind eq. (5.0.2),

$$\partial_{\lambda'} F_{\alpha'\beta'}\left(x'\right) = \frac{\partial x^\sigma}{\partial x'^\lambda} \partial_\sigma F_{\mu\nu}\left(x(x') = \Lambda \cdot x'\right) \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \tag{5.0.4}$$

$$= \Lambda^\sigma{}_\lambda \partial_\sigma F_{\mu\nu}\left(x(x')\right) \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta. \tag{5.0.5}$$

This immediately tells us $\partial_\mu F^{\mu\nu} = \eta^{\mu\alpha} \partial_\mu F_{\alpha\beta} \eta^{\beta\nu}$ in eq. (5.0.1) is a Lorentz vector.

Problem 5.1. 4D Maxwell's Equations in term of $(\vec{E}, \vec{B})$ Let us check that eq. (5.0.1) does in fact reproduce Maxwell's equations in terms of electric E^i and magnetic B^i fields in 4D. Given a Lorentzian inertial frame, define

$$F^{i0} \equiv E^i \qquad \text{and} \qquad F^{ij} \equiv \epsilon^{ijk} B^k; \tag{5.0.6}$$

with $\epsilon^{123} \equiv -1$. Show that the $\partial_\mu F^{\mu\nu} = J^\nu$ from eq. (5.0.1) translates to

$$\vec{\nabla} \cdot \vec{E} = J^0 \qquad \text{and} \qquad \vec{\nabla} \times \vec{B} - \partial_t \vec{E} = \vec{J}. \tag{5.0.7}$$

[12] Nature is actually not symmetric under parity flips nor under time reversal; but proper Lorentz symmetry does appear to hold to a high degree of accuracy.

(The over-arrow refers to the spatial components; for instance $\vec{B} = (B^1, B^2, B^3)$.) The $\partial_{[\alpha} F_{\mu\nu} = 0$ from eq. (5.0.1) translates to

$$\partial_t \vec{B} + \vec{\nabla} \times \vec{E} = 0 \qquad \text{and} \qquad \vec{\nabla} \cdot \vec{B} = 0. \tag{5.0.8}$$

Hint: Note that $(\vec{\nabla} \times \vec{A})^i = -\epsilon^{ijk} \partial_j A^k$, for any Cartesian vector $\vec{A}$. Also, when you compute $\partial_{[i} F_{jk]}$, you simply need to set $\{i, j, k\}$ to be any distinct permutation of $\{1, 2, 3\}$. (Why?)

Next, verify the Lorentz invariant relations, with $\epsilon^{0123} \equiv -1$:

$$F_{\mu\nu} F^{\mu\nu} = -2 \left(\vec{E}^2 - \vec{B}^2 \right), \qquad \vec{E}^2 = E^i E^i, \ \vec{B}^2 = B^i B^i, \tag{5.0.9}$$

$$\epsilon^{\mu\nu\alpha\beta} F_{\mu\nu} F_{\alpha\beta} = \partial_\mu \left(\epsilon^{\mu\nu\alpha\beta} A_\nu \partial_\alpha A_\beta \right) = 8 \vec{E} \cdot \vec{B}. \tag{5.0.10}$$

How does $F_{\mu\nu} F^{\mu\nu}$ transform under time reversal, $t \equiv x^0 \to -t$? How does it transform under parity flips, $x^i \to -x^i$ (for a fixed i)? Answer the same questions for $\widetilde{F}^{\mu\nu} F_{\mu\nu}$, where the dual of $F_{\mu\nu}$ is

$$\widetilde{F}^{\mu\nu} \equiv \frac{1}{2} \widetilde{\epsilon}^{\mu\nu\alpha\beta} F_{\alpha\beta}. \tag{5.0.11}$$

$d \neq 4$ Can you comment what the analog of the magnetic field ought to be in spacetime dimensions different from 4 – is it still a 'vector'? – and what is the lowest dimension that the magnetic field still exists? How many components does the electric field have in 1+1 dimensions?

□

Current conservation Taking the divergence of $\partial_\mu F^{\mu\nu} = J^\nu$ yields the conservation of the electric current as a consistency condition. For, by the antisymmetry $F_{\mu\nu} = -F_{\nu\mu}$, $\partial_\nu \partial_\mu F^{\mu\nu} = (1/2) \partial_\nu \partial_\mu F^{\mu\nu} - (1/2) \partial_\mu \partial_\nu F^{\nu\mu} = 0$.

$$\partial_\mu J^\mu = 0. \tag{5.0.12}$$

Problem 5.2. Total charge is constant Even though we defined ρ in the $J^\mu = \rho v^\mu$ as the charge density in the local rest frame of the electric current itself, we may also define the charge density $J^{\widehat{0}} \equiv u_\mu J^\mu$ in the rest frame of an arbitrary family of inertial time-like observers whose worldlines' tangent vector is $u^\mu \partial_\mu = \partial_\tau$. (In other words, in their frame, the spacetime metric is $\mathrm{d}s^2 = (\mathrm{d}\tau)^2 - \mathrm{d}\vec{x} \cdot \mathrm{d}\vec{x}$.) Show that total charge is independent of the Lorentz frame by demonstrating that

$$Q \equiv \int_{\mathbb{R}^D} \mathrm{d}^D \Sigma_\mu J^\mu, \qquad \mathrm{d}^D \Sigma_\mu \equiv \mathrm{d}^D \vec{x} n_\mu, \tag{5.0.13}$$

is a constant. □

Vector Potential & Gauge Symmetry The other Maxwell equation (cf eq. (5.0.1)) leads us to introduce a vector potential A_μ. For $\partial_{[\mu} F_{\alpha\beta]} = 0 \Leftrightarrow \mathrm{d}F = 0$ tells us, by the Poincaré lemma, that

$$F = \mathrm{d}A \qquad \Leftrightarrow \qquad F_{\mu\nu} = \partial_{[\mu} A_{\nu]} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{5.0.14}$$

Notice the dynamics in eq. (5.0.1) is not altered if we add to A_μ any object L_μ that obeys $\mathrm{d}L = 0$, because that does not alter the Faraday tensor: $F = \mathrm{d}(A + L) = \mathrm{d}A + \mathrm{d}L = F$. Now, $\mathrm{d}L = 0$ means, again by the Poincaré lemma, that $L_\mu = \partial_\mu L$, where L on the right hand side is a scalar. *Gauge symmetry*, in the context of electromagnetism, is the statement that the following replacement involving the gauge potential

$$A_\mu(x) \to A_\mu(x) + \partial_\mu L(x) \tag{5.0.15}$$

leaves the dynamics encoded in Maxwell's equations (5.0.1) unchanged.

The use of the gauge potential A_μ makes the $\mathrm{d}F = 0$ portion of the dynamics in eq. (5.0.1) redundant; and what remains is the vector equation

$$\partial_\mu F^{\mu\nu} = \partial_\mu \left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) = J^\nu. \tag{5.0.16}$$

The symmetry under the gauge transformation of eq. (5.0.15) means that solutions to eq. (5.0.16) cannot be unique – in particular, since A_μ and $A_\mu + \partial_\mu L$ are simultaneously solutions, there really is an infinity of solutions parametrized by the arbitrary function L. In this same vein, by going to Fourier space, namely

$$A_\mu(x) \equiv \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \widetilde{A}_\mu(k) e^{-ik_\mu x^\mu} \qquad \text{and} \qquad J_\mu(x) \equiv \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \widetilde{J}_\mu(k) e^{-ik_\mu x^\mu}, \tag{5.0.17}$$

we may see that the differential operator in eq. (5.0.16) cannot be inverted because it has a zero eigenvalue. Firstly, the Fourier version of eq. (5.0.16) reads

$$-K^{\mu\nu} \widetilde{A}_\mu = \widetilde{J}^\nu, \tag{5.0.18}$$

$$K^{\mu\nu} \equiv k_\sigma k^\sigma \eta^{\mu\nu} - k^\nu k^\mu. \tag{5.0.19}$$

If K^{-1} exists, the solution in Fourier space would be (schematically) $\widetilde{A} = -K^{-1}\widetilde{J}$. That $K^{\mu\nu} = K^{\nu\mu}$ has no inverse, however, can be seen by observing that k_μ is in fact its null eigenvector:

$$K^{\mu\nu} k_\mu = (k_\sigma k^\sigma) k^\nu - k^\nu k^\mu k_\mu = 0. \tag{5.0.20}$$

Problem 5.3. Can you explain why eq. (5.0.20) amounts to the statement that $F_{\mu\nu}$ is invariant under the gauge transformation of eq. (5.0.15)?

Lorenz gauge To make $K^{\mu\nu}$ invertible, one *fixes a gauge*. A common choice is the Lorenz gauge; in Fourier spacetime:

$$k^\mu \widetilde{A}_\mu = 0. \tag{5.0.21}$$

In 'position'/real spacetime, this reads instead

$$\partial^\mu A_\mu = 0 \qquad \text{(Lorenz gauge)}. \tag{5.0.22}$$

With the constraint in eq. (5.0.21), Maxwell's equations in eq. (5.0.18) becomes

$$-\left(k_\sigma k^\sigma \widetilde{A}^\nu - k^\mu (k^\nu \widetilde{A}_\nu)\right) = -k_\sigma k^\sigma \widetilde{A}^\nu = \widetilde{J}^\nu. \tag{5.0.23}$$

Now, Maxwell's equations have become invertible:

$$\widetilde{A}_\mu(k) = \frac{\widetilde{J}_\mu(k)}{-k^2}, \qquad k^2 \equiv k_\sigma k^\sigma, \qquad \text{(Lorenz gauge).} \tag{5.0.24}$$

In position/real spacetime, eq. (5.0.23) is equivalent to

$$\partial^2 A^\nu(x) = J^\nu(x) \qquad \partial^2 \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu. \tag{5.0.25}$$

[13]In the Lorenz gauge, we have d Minkowski scalar wave equations, one for each Cartesian component. We may express its position spacetime solution by inverting the Fourier transform in eq. (5.0.24):

$$A_\mu(x) = \int_{\mathbb{R}^{d-1,1}} \mathrm{d}^d x' G^+_d(x-x') J_\mu(x'), \tag{5.0.26}$$

$$G^+_d(x-x') \equiv \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{e^{-ik\cdot(x-x')}}{-k^2}. \tag{5.0.27}$$

Because A_μ is not gauge-invariant, its physical interpretation can be ambiguous. Classically it is the electromagnetic fields $F_{\mu\nu}$ that exert forces on charges/currents, so we need its solution. In fact, we may take the curl of eq. (5.0.25) to see that $\partial^2 F_{\mu\nu} = \partial_{[\mu} J_{\nu]}$; this means, using the same Green's function in eq. (5.0.27):

$$F_{\mu\nu}(x) = \int_{\mathbb{R}^{d-1,1}} \mathrm{d}^d x' G^+_d(x-x') \partial_{[\mu} J_{\nu]}(x'). \tag{5.0.28}$$

We may verify that equations (5.0.26) and (5.0.27) solve eq. (5.0.25) readily:

$$\begin{aligned}
\partial_x^2 G^+_d(x-x') &= \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{\partial_\sigma \partial^\sigma e^{-ik\cdot(x-x')}}{-k^2} \\
&= \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{\partial_\sigma(-ik_\rho \delta^\sigma_\sigma e^{-ik\cdot(x-x')})}{-k^2} \\
&= \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{\partial_\sigma(-ik^\sigma e^{-ik\cdot(x-x')})}{-k^2} \\
&= \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{(-ik_\sigma)(-ik^\sigma) e^{-ik\cdot(x-x')}}{-k^2} \\
&= \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} e^{-ik\cdot(x-x')} = \delta^{(d)}(x-x');
\end{aligned} \tag{5.0.29}$$

with a similar calculation showing $\partial_{x'}^2 G^+_d(x-x') = \delta^{(d)}(x-x')$. Therefore

$$\begin{aligned}
\partial_x^2 A_\mu(x) = \partial_x^2 \left(\int_{\mathbb{R}^{d-1,1}} \mathrm{d}^d x' G^+_d(x-x') J_\mu(x') \right) &= \int_{\mathbb{R}^{d-1,1}} \mathrm{d}^d x' \delta^{(d)}(x-x') J_\mu(x') \\
&= J_\mu(x).
\end{aligned} \tag{5.0.30}$$

[13]Eq. (5.0.25) is valid in any dimension $d \geq 3$. In 2D, the $\mathrm{d}F = 0$ portion of Maxwell's equations is trivial – i.e., *any* F would satisfy it – because there cannot be three distinct indices in $\partial_{[\mu} F_{\alpha\beta]} = 0$.

Lorenz gauge: Existence That we have managed to solve Maxwell's equations using the Lorenz gauge, likely convinces the practical physicist that the Lorenz gauge itself surely exists. However, it is certainly possible to provide a general argument. For suppose $\partial^\mu A_\mu$ were not zero, then all one has to show is that we may perform a gauge transformation (cf. (5.0.15)) that would render the new gauge potential $A'_\mu \equiv A_\mu - \partial_\mu F$ satisfy

$$\partial^\mu A'_\mu = \partial^\mu A_\mu - \partial^2 F = 0. \tag{5.0.31}$$

But all that means is, we have to solve $\partial^2 F = \partial^\mu A_\mu$; and since the Green's function $1/\partial^2$ exists, we have proved the assertion.

Lorenz gauge and current conservation You may have noticed, by taking the divergence of both sides of eq. (5.0.25),

$$\partial^2 \left(\partial^\sigma A_\sigma\right) = \partial^\sigma J_\sigma. \tag{5.0.32}$$

This teaches us the consistency of the Lorenz gauge is intimately tied to the conservation of the electric current $\partial^\sigma J_\sigma = 0$. Another way to see this, is to take the time derivative of the divergence of the vector potential, followed by subtracting and adding the spatial Laplacian of A_0 so that $\partial^2 A_0 = J_0$ may be employed:

$$\begin{aligned}
\partial^\sigma \dot{A}_\sigma = \ddot{A}_0 + \partial^i \dot{A}_i &= \partial^0\partial_0 A_0 + \partial^i\partial_i A_0 + \partial^i\partial_0 A_i - \partial^i\partial_i A_0) \\
&= \partial^2 A_0 - \partial^i\left(\partial_i A_0 - \partial_0 A_i\right) \\
\partial_0\left(\partial^\sigma A_\sigma\right) &= J_0 - \partial^i F_{i0}.
\end{aligned} \tag{5.0.33}$$

Notice the right hand side of the last line is zero if the $\nu = 0$ component of $\partial_\mu F^{\mu\nu} = J^\nu$ is obeyed.

5.1 4 dimensions

4D Maxwell We now focus on the physically most relevant case of $(3+1)D$. In 4D, the wave operator ∂^2 has the following inverse – i.e., retarded Green's function – that obeys causality:

$$G_4^+(x-x') \equiv \frac{\delta(t-t'-|\vec{x}-\vec{x}'|)}{4\pi|\vec{x}-\vec{x}'|}, \qquad x^\mu = (t,\vec{x}),\ x'^\mu = (t',\vec{x}'), \tag{5.1.1}$$

$$\partial_x^2 G_4^+(x-x') = \partial_{x'}^2 G_4^+(x-x') = \delta^{(4)}(x-x'), \tag{5.1.2}$$

$$\partial_x^2 \equiv \eta^{\mu\nu}\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu} \qquad \partial_{x'}^2 \equiv \eta^{\mu\nu}\frac{\partial}{\partial x'^\mu}\frac{\partial}{\partial x'^\nu}. \tag{5.1.3}$$

Problem 5.4. Lorentz covariance Suppose $\Lambda^\alpha{}_\mu$ is a Lorentz transformation; let two inertial frames $\{x^\mu\}$ and $\{x'^\mu\}$ be related via

$$x^\mu = \Lambda^\mu{}_\alpha x'^\alpha. \tag{5.1.4}$$

Suppose we solved the Lorenz gauge Maxwell's equations in the $\{x^\mu\}$ frame, namely

$$\frac{\partial A^\mu(x)}{\partial x^\mu} = 0, \qquad \eta^{\mu\nu}\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu}A_\alpha(x) = J_\alpha(x). \tag{5.1.5}$$

Explain how to solve $A_{\alpha'}(x')$, the solution in the $\{x'^\mu\}$ frame. $\square$

Problem 5.5. Analogy: Driven Simple Harmonic Oscillator Suppose we only Fourier-transformed the spatial coordinates in the Lorenz gauge Maxwell eq. (5.0.25). Show that this leads to

$$\ddot{\widetilde{A}}_\mu(t,\vec{k}) + k^2 \widetilde{A}_\mu(t,\vec{k}) = \widetilde{J}_\mu(t,\vec{k}), \qquad k \equiv |\vec{k}|. \tag{5.1.6}$$

[14]Compare this to the simple harmonic oscillator (in flat space), with Cartesian coordinate vector $\vec{q}(t)$, mass m, spring constant σ, and driven by an external force $\vec{f}$:

$$m\ddot{\vec{q}} + \sigma\vec{q} = \vec{f}, \tag{5.1.7}$$

where each over-dot corresponds to a time derivative. Identify k^2 and $\widetilde{J}$ in eq. (5.1.6) with the appropriate quantities in eq. (5.1.7). Even though the Lorenz gauge Maxwell equations are fully relativistic, notice the analogy with the non-relativistic driven harmonic oscillator! In particular, when the electric current is not present (i.e., $J_\mu = 0$), the 'mixed-space' equations of (5.1.6) are in fact a collection of free simple harmonic oscillators.

Now, how does one solve eq. (5.1.7)? Explain why the inverse of $(\mathrm{d}/\mathrm{d}t)^2 + k^2$ is

$$G_{\text{SHO}}(t-t',k) = -\int_{\mathbb{R}} \frac{\mathrm{d}\omega}{2\pi} \frac{e^{-i\omega(t-t')}}{\omega^2 - k^2}. \tag{5.1.8}$$

That is, verify that this equation satisfies

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}t^2} + k^2\right) G_{\text{SHO}}(t-t',k) = \left(\frac{\mathrm{d}^2}{\mathrm{d}t'^2} + k^2\right) G_{\text{SHO}}(t-t',k) = \delta(t-t'). \tag{5.1.9}$$

If one tries to integrate ω over the real line in eq. (5.1.8), one runs into trouble – explain the issue. Now evaluate the Green's function G^+_{SHO} in eq. (5.1.8) using the contour running just slightly above the real line, i.e., $\omega \in (-\infty + i0^+, +\infty + i0^+)$. You should find

$$G^+_{\text{SHO}}(t-t',k) = \Theta(t-t')\frac{\sin\left(k(t-t')\right)}{k}. \tag{5.1.10}$$

Hence, the mixed-space Maxwell's equations have the solution

$$\widetilde{A}_\mu(t,\vec{k}) = \int_{-\infty}^{t} \mathrm{d}t' G^+_{\text{SHO}}(t-t',k)\widetilde{J}(t',\vec{k}). \tag{5.1.11}$$

By performing an inverse-Fourier transform, namely

$$A_\mu(x) = \int_{\mathbb{R}^{3,1}} \mathrm{d}^4x' G^+_4(x-x')J_\mu(x'), \tag{5.1.12}$$

arrive at the expression in eq. (5.1.1) □

[14]This equation actually holds in all dimensions $d \geq 3$.

Vacuum solution & Spin/Helicity-1 Let us examine the simplest situation in 4D flat spacetime, where there are no electric charges nor currents present: $J_\nu = 0$. In Fourier space, setting $\widetilde{J} = 0$ in eq. (5.1.6) leads us to

$$\ddot{\widetilde{A}}_\mu(t, \vec{k}) + k^2 \widetilde{A}_\mu(t, \vec{k}) = 0, \qquad k \equiv |\vec{k}|. \tag{5.1.13}$$

The solutions are $\widetilde{A}_\mu(t, \vec{k}) = \exp(\pm ikt)$. Hence, the general solution is the superposition

$$A_\mu = \int_{\mathbb{R}^3} \frac{d^3\vec{k}}{(2\pi)^3} \left(a_\mu(\vec{k}) \exp(-ikt + i\vec{k}\cdot\vec{x}) + b_\mu(\vec{k}) \exp(ikt + i\vec{k}\cdot\vec{x}) \right) \tag{5.1.14}$$

$$= \int_{\mathbb{R}^3} \frac{d^3\vec{k}}{(2\pi)^3} \left(a_\mu(\vec{k}) \exp(-ikt + i\vec{k}\cdot\vec{x}) + b_\mu(-\vec{k}) \exp(ikt - i\vec{k}\cdot\vec{x}) \right). \tag{5.1.15}$$

But since A_μ is real it must be that $a_\mu(\vec{k})^* = b_\mu(-\vec{k})$:

$$A_\mu = \int_{\mathbb{R}^3} \frac{d^3\vec{k}}{(2\pi)^3} \left(a_\mu(\vec{k}) e^{-ik\cdot x} + a_\mu(\vec{k})^* e^{ik\cdot x} \right). \tag{5.1.16}$$

Since a_μ has been arbitrary thus far, we may write a single plane wave solution to eq. (5.1.6) as

$$\widetilde{A}_\mu(t, \vec{k}) e^{i\vec{k}\cdot\vec{x}} = \mathrm{Re}\left\{ \epsilon_\mu(\vec{k}) e^{-ik\cdot t} e^{i\vec{k}\cdot\vec{x}} \right\} = \mathrm{Re}\left\{ \epsilon_\mu(\vec{k}) e^{-ik\cdot x} \right\},$$
$$k_\mu \equiv (k, k_i), \qquad k \equiv |\vec{k}| = \sqrt{\delta^{ab} k_a k_b}. \tag{5.1.17}$$

The Lorenz gauge says $k^\mu \widetilde{A}_\mu = 0$. Since the $\exp(-ik_\mu x^\mu)$ are basis functions, it must be that the polarization vector ϵ_μ itself is orthogonal to the momentum vector k^μ:

$$k^\mu \epsilon_\mu(\vec{k}) = 0. \tag{5.1.18}$$

Let us suppose k_i points in the positive 3−axis, so that

$$k_\mu = k(1, 0, 0, -1) \qquad \text{and} \qquad k^\mu = k(1, 0, 0, 1). \tag{5.1.19}$$

This means the plane wave itself becomes

$$\exp(-ik_\mu x^\mu) = \exp(-ik(t - x^3)); \tag{5.1.20}$$

i.e., it indeed describes propagation in the positive 3−direction. The polarization tensor may then be decomposed into the following basis vectors,

$$\epsilon_\mu = a_{0,+}\epsilon_\mu^{(0,+)} + a_{0,-}\epsilon_\mu^{(0,-)} + a_+\epsilon_\mu^+ + a_-\epsilon_\mu^-; \tag{5.1.21}$$

where the a's are (scalar) complex amplitudes while

$$\epsilon_\mu^{(0,\pm)} \equiv \frac{1}{\sqrt{2}} (1, 0, 0, \pm 1)^T, \tag{5.1.22}$$

$$\epsilon_\mu^\pm \equiv \frac{1}{\sqrt{2}} (0, \mp 1, i, 0)^T. \tag{5.1.23}$$

Now, under the following rotation on the $(1,2)$-plane orthogonal to $\vec{k}$, namely

$$\widehat{R}(\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) & 0 \\ 0 & \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{5.1.24}$$

the polarizations $\epsilon_\mu^{(0,\pm)}$ in eq. (5.1.22) remain unchanged – they are the spin-0 modes – while the polarizations $\epsilon_\mu^\pm$ in eq. (5.1.22) transform as

$$\widehat{R}(\theta)^\mu{}_\nu \epsilon^{\pm\ \nu} = e^{\pm i\theta}\epsilon^{\pm\ \mu}. \tag{5.1.25}$$

These $\epsilon_\mu^\pm$ are the spin-1 modes.

Problem 5.6. Verify eq. (5.1.25).

We now turn to imposing the Lorenz gauge condition $k^\mu \widetilde{A}_\mu = 0$.

$$k\epsilon_0 + k\epsilon_3 = 0 \qquad \Rightarrow \qquad \epsilon_3 = -\epsilon_0. \tag{5.1.26}$$

Since the 0th component has to be negative the 3rd, the $\epsilon_\mu^{(0,+)}$ cannot occur in the decomposition of eq. (5.1.21). But since $\epsilon_\mu^{(0,-)}$ is proportional to k_μ (cf. eq. (5.1.22)) and $k^2 \equiv k_\nu k^\nu = 0$, we see this remaining spin-0 piece of the polarization tensor simultaneously satisfies the Lorenz gauge and is a gradient term – and hence 'pure gauge' (cf. the $\partial_\mu L$ terms of eq. (5.0.15)) – in position spacetime:

$$a_{0,-}\epsilon_\mu^{(0,-)} = a_{0,-}\frac{k_\nu}{k}. \tag{5.1.27}$$

Since this term will not contribute to the electromagnetic fields $F_{\mu\nu}$, we may perform a Lorenz-gauge-preserving gauge transformation to cancel this term:

$$\widetilde{A}'_\mu(t,\vec{k}) \equiv \widetilde{A}_\mu(t,\vec{k}) - a_{0,-}\frac{k_\mu}{k}e^{-ikt}. \tag{5.1.28}$$

And now that we have canceled the 0th and 3rd component of the polarization vector,

$$\widetilde{A}'_\nu(t,\vec{k})e^{i\vec{k}\cdot\vec{x}} = \left(a_+\epsilon_\mu^+ + a_-\epsilon_\mu^-\right)e^{-ik\cdot x}. \tag{5.1.29}$$

To sum, given an inertial frame, the electromagnetic vector potential A_μ in vacuum is given by the following superposition of spin-1 waves:

$$A_\mu(x) = \text{Re}\int_{\mathbb{R}^3}\frac{\mathrm{d}^3\vec{k}}{(2\pi)^3}\left(a_+\epsilon_\mu^+(\vec{k}) + a_-\epsilon_\mu^-(\vec{k})\right)e^{-ik\cdot x}, \tag{5.1.30}$$

where $\epsilon_\mu^\pm$ are purely spatial polarization tensors orthogonal to the k_i; and, under a rotation by an angle θ around the plane perpendicular to k_i transforms as $\epsilon_\mu^\pm \to \exp(\pm i\theta)\epsilon_\mu^\pm$.[15]

[15] For a given inertial frame and within the Lorenz gauge, we have been able to get rid of the 'pure gauge' spin-0 mode by a gauge transformation, leaving only the spin-1 (simple-harmonic) waves. Note, however, these waves in eq. (5.1.30) would no longer be a admixture of pure spin-1 modes – simply by viewing them in a different reference frame, i.e., upon a Lorentz boost.

Problem 5.7. Circularly polarized light from 4D spin-1 Consider a single spin-1 plane wave (cf. (5.1.23)) propagating along the 3−axis, with $k_\mu = k(1, 0, 0, 1)$:

$$A^{\pm}_{\mu}(t, x, y, z) \equiv \text{Re}\left\{a_{\pm}\epsilon^{\pm}_{\mu} e^{-ik(t-z)}\right\}, \qquad a_{\pm} \in \mathbb{R}. \tag{5.1.31}$$

Compute the electric field ${}_{\pm}E^i = F^{i0}$ and show that these plane waves give rise to circularly polarized light, i.e., for either a fixed time t or spatial location z – the electric field direction rotates in a circular fashion:

$$_{\pm}E^i = \frac{k \cdot a_{\pm}}{\sqrt{2}} \left(\pm \sin(k(t-z))\widehat{x}^i + \cos(k(t-z))\widehat{y}^i\right), \tag{5.1.32}$$

where $\widehat{x}$ and $\widehat{y}$ are the unit vectors in the 1− and 2−directions:

$$\widehat{x}^i \doteq (1, 0, 0) \qquad \text{and} \qquad \widehat{y}^i \doteq (0, 1, 0). \tag{5.1.33}$$

□

Redshift For each Lorenz-gauge plane wave in an inertial frame $\{x^\mu = (t, \vec{x})\}$,

$$\epsilon^{\pm}_{\mu}(k) \exp(-ik \cdot x) = \epsilon^{\pm}_{\mu}(k) \exp(-ik_j x^j) \exp(-i\omega t), \qquad \omega \equiv |\vec{k}|, \tag{5.1.34}$$

we may read off its frequency ω as the coefficient of the time coordinate t. Quantum mechanics tells us ω is also the energy of the associated photon. In a different Lorentz inertial frame $\{x'\}$ related to the previous through the Lorentz transformation $\Lambda^\alpha{}_\mu$: $x^\alpha = \Lambda^\alpha{}_\mu x'^\mu$. Because the phase in the plane wave solution of eq. (5.1.34) is a scalar, in the $\{x'\}$ Lorentz frame

$$-ik_\alpha x^\alpha = -ik_\alpha \Lambda^\alpha{}_\mu x'^\mu = -i(k_\alpha \Lambda^\alpha{}_0)t' - i(k_\alpha \Lambda^\alpha{}_i)x'^i. \tag{5.1.35}$$

The frequency ω' and hence the photon's energy in this $\{x'\}$ frame is therefore

$$\omega' = k_\alpha \Lambda^\alpha{}_0 = \omega \left(\Lambda^0{}_0 + \widehat{k}_i \Lambda^i{}_0\right) \tag{5.1.36}$$

$$\widehat{k}_i \equiv k_i/|\vec{k}|. \tag{5.1.37}$$

There is a slightly different way to express this redshift result, that would help us generalize the analysis to curved spacetime. To extract the frequency directly from the phase $S \equiv k \cdot x$, we may take its time derivative using the unit norm vector $u \equiv \partial_t = \partial_0$ that we may associate with the worldlines of observers at rest in the $\{x\}$ frame:

$$u^\mu \partial_\mu S = \partial_0(k_\alpha x^\alpha) = \omega. \tag{5.1.38}$$

The observers at rest in the $\{x'\}$ frame have $u' \equiv \partial_{t'} = \partial_{0'}$ as their timelike unit norm tangent vector. (Note: $x^\alpha = \Lambda^\alpha{}_\mu x'^\mu \Leftrightarrow \partial_{\mu'} = \Lambda^\alpha{}_\mu \partial_\alpha$.) The energy of the photon is then

$$\begin{aligned} u'^\alpha \partial_{\alpha'} S = \partial_{t'} S = \Lambda^\alpha{}_0 \partial_\alpha (k \cdot x) &= \Lambda^\alpha{}_0 k_\alpha \\ &= \omega \left(\Lambda^0{}_0 + \widehat{k}_i \Lambda^i{}_0\right). \end{aligned} \tag{5.1.39}$$

Problem 5.8. Consider a single photon with wave vector $k_\mu = \omega(1, \widehat{n}_i)$ (where $\widehat{n}_i \widehat{n}_j \delta^{ij} = 1$) in some inertial frame $\{x^\mu\}$. Let a family of inertial observers be moving with constant velocity $v^\mu \equiv (1, v^i)$ with respect to the frame $\{x^\mu\}$. What is the photon's frequency ω' in their frame? Compute the redshift formula for ω'/ω. Comment on the redshift result when v^i is (anti)parallel to $\widehat{n}_i$ and when v^i is perpendicular to $\widehat{n}_i$.

Problem 5.9. Dispersion relations Consider the *massive* Klein-Gordon equation in Minkowski spacetime:

$$\left(\partial^2 + m^2\right) \varphi(t, \vec{x}) = 0, \tag{5.1.40}$$

where φ is a real scalar field. Find the general solution for φ in terms of plane waves $\exp(-ik \cdot x)$ and obtain the dispersion relation:

$$k^2 = m^2 \qquad \Leftrightarrow \qquad E^2 = \vec{p}^2 + m^2, \tag{5.1.41}$$

$$E \equiv k^0, \qquad \vec{p} \equiv \vec{k}. \tag{5.1.42}$$

If each plane wave is associated with a particle of $d-$momentum k_μ, this states that it has mass m. The photon, which obeys $k^2 = 0$, has zero mass.

5.2 Gauge Invariant Variables for Vector Potential

Although the vector potential A_μ itself is not a gauge invariant object, we will now exploit the spatial translation symmetry of Minkowski spacetime to seek a gauge-invariant set of partial differential equations involving a "scalar-vector" decomposition of A_μ. There are at least two reasons for doing so.

- This will be a warm-up to an analogous analysis for gravitation linearized about a Minkowski "background" spacetime.
- We will witness how, for a given inertial frame, the only portion of the vector potential A_μ that obeys a wave equation is its gauge-invariant "transverse" spatial portion. (Even though every component of A_μ in the Lorenz gauge (cf. eq. (5.0.25)) obeys the wave equation, remember such a statement is not gauge-invariant.) We shall also identify a gauge-invariant scalar potential sourced by charge density.

Scalar-Vector Decomposition The scalar-vector decomposition is the statement that the spatial components of the vector potential may be expressed as a gradient of a scalar α plus a transverse vector α_i:

$$A_i = \partial_i \alpha + \alpha_i, \tag{5.2.1}$$

where by "transverse" we mean

$$\partial_i \alpha_i = 0. \tag{5.2.2}$$

To demonstrate the generality of eq. (5.2.1) we shall first write A_i in Fourier space

$$A_i(t, \vec{x}) = \int_{\mathbb{R}^D} \frac{\mathrm{d}^D \vec{k}}{(2\pi)^D} \widetilde{A}_i(t, \vec{k}) e^{i\vec{k}\cdot\vec{x}}; \tag{5.2.3}$$

where $\vec{k} \cdot \vec{x} \equiv \delta_{ij} k^i x^j = -k_j x^j$. Every spatial derivative ∂_j acting on $A_i(t, \vec{x})$ becomes in Fourier space a $-ik_j$, since

$$
\begin{aligned}
\partial_j A_i &= \int_{\mathbb{R}^D} \frac{\mathrm{d}^D \vec{k}}{(2\pi)^D} \partial_j \left(i \delta_{ab} k^a x^b \right) \widetilde{A}_i(t, \vec{k}) e^{i \vec{k} \cdot \vec{x}} \\
&= \int_{\mathbb{R}^D} \frac{\mathrm{d}^D \vec{k}}{(2\pi)^D} \left(i \delta_{ab} k^a \delta^b_j \right) \widetilde{A}_i(t, \vec{k}) e^{i \vec{k} \cdot \vec{x}} \\
&= \int_{\mathbb{R}^D} \frac{\mathrm{d}^D \vec{k}}{(2\pi)^D} i k^j \widetilde{A}_i(t, \vec{k}) e^{i \vec{k} \cdot \vec{x}} \\
&= \int_{\mathbb{R}^D} \frac{\mathrm{d}^D \vec{k}}{(2\pi)^D} (-i k_j) \widetilde{A}_i(t, \vec{k}) e^{i \vec{k} \cdot \vec{x}}.
\end{aligned} \tag{5.2.4}
$$

As such, the transverse property of $\alpha_i(t, \vec{x})$ would in Fourier space become

$$
-i k_i \widetilde{\alpha}_i(t, \vec{k}) = 0. \tag{5.2.5}
$$

At this point we simply write down

$$
\widetilde{A}_i(t, \vec{k}) = \left(\delta_{ij} - \frac{k_i k_j}{\vec{k}^2} \right) \widetilde{A}_j(t, \vec{k}) + \frac{k_i k_j}{\vec{k}^2} \widetilde{A}_j(t, \vec{k}). \tag{5.2.6}
$$

This is mere tautology, of course. However, we may now check that the first term on the left hand side of eq. (5.2.6) is transverse:

$$
-i k_i \left(\delta_{ij} - \frac{k_i k_j}{\vec{k}^2} \right) \widetilde{A}_j(t, \vec{k}) = -i \left(k_j - \frac{\vec{k}^2 k_j}{\vec{k}^2} \right) \widetilde{A}_j(t, \vec{k}) = 0. \tag{5.2.7}
$$

The second term on the right hand side of eq. (5.2.6) is a gradient because it is

$$
-i k_i \left(\frac{i k_j}{\vec{k}^2} \widetilde{A}_j \right). \tag{5.2.8}
$$

To sum, we have identified the α and α_i terms of eq. (5.2.1) as

$$
\alpha(t, \vec{x}) = \int_{\mathbb{R}^D} \frac{\mathrm{d}^D \vec{k}}{(2\pi)^D} \frac{i k_j}{\vec{k}^2} \widetilde{A}_j(t, \vec{k}) e^{i \vec{k} \cdot \vec{x}}; \tag{5.2.9}
$$

and the transverse portion as

$$
\begin{aligned}
\alpha_i(t, \vec{x}) &= \int_{\mathbb{R}^D} \frac{\mathrm{d}^D \vec{k}}{(2\pi)^D} P_{ij}(\vec{k}) \widetilde{A}_j(t, \vec{k}) e^{i \vec{k} \cdot \vec{x}}, \\
P_{ij}(\vec{k}) &\equiv \delta_{ij} - \frac{k_i k_j}{\vec{k}^2}.
\end{aligned} \tag{5.2.10}
$$

Notice it is really the projector P_{ij} that is "transverse"; i.e.

$$
k_i P_{ij}(\vec{k}) = 0. \tag{5.2.11}
$$

Let us also note that this scalar-vector decomposition is unique, in that – if we have the Fourier-space equation

$$-ik_i\widetilde{\alpha} + \widetilde{\alpha}_i = -ik_i\widetilde{\beta} + \widetilde{\beta}_i, \tag{5.2.12}$$

where $k_i\widetilde{\alpha}_i = k_i\widetilde{\beta}_i = 0$, then

$$\widetilde{\alpha} = \widetilde{\beta} \qquad \text{and} \qquad \widetilde{\alpha}_i = \widetilde{\beta}_i. \tag{5.2.13}$$

For, we may first "dot" both sides of eq. (5.2.12) with $\vec{k}$ and see that – for $\vec{k} \neq \vec{0}$,

$$\vec{k}^2\widetilde{\alpha} = \vec{k}^2\widetilde{\beta} \qquad \Leftrightarrow \qquad \widetilde{\alpha} = \widetilde{\beta}. \tag{5.2.14}$$

Plugging this result back into eq. (5.2.12), we also conclude $\widetilde{\alpha}_i = \widetilde{\beta}_i$.

Now, this scalar-vector decomposition is really just a mathematical fact, and may even be performed in a curved space – as long as the latter is infinite – since it depends on the existence of the Fourier transform and not on the metric structure. (A finite space would call for a discrete Fourier-like series of sorts.) However, to determine its usefulness, we would need to insert it into the partial differential equations obeyed by A_i, where the metric structure does matter. As we now turn to examine, because of the spatial translation symmetry of Minkowski spacetime, Maxwell's equations themselves admit a scalar-vector decomposition. This, in turn, would lead to PDEs for the gauge-invariant portions of A_μ.

Gauge transformations We first examine how the gauge transformation of eq. (5.0.15) is implemented on a scalar-vector decomposed A_μ.

$$A_0 \to A_0 + \dot{L} \tag{5.2.15}$$

$$A_i = \partial_i\alpha + \alpha_i \to \partial_i\alpha + \alpha_i + \partial_i L \tag{5.2.16}$$

$$= \partial_i(\alpha + L) + \alpha_i. \tag{5.2.17}$$

From the uniqueness discussion above, we may thus identify the gauge-transformed "scalar" portion of A_i

$$\alpha \to \alpha' \equiv \alpha + L \tag{5.2.18}$$

and the "transverse-vector" portion of A_i to be gauge-invariant:

$$\alpha_i \to \alpha_i. \tag{5.2.19}$$

Let us now identify

$$\Phi \equiv A_0 - \dot{\alpha} \tag{5.2.20}$$

because it is gauge-invariant; for, according to eq. (5.2.18)

$$\Phi \to A_0 + \dot{L} - \partial_0(\alpha + L) = A_0 - \dot{\alpha}. \tag{5.2.21}$$

In terms of Φ and α_i, the components of the electromagnetic tensor read

$$F_{0i} \equiv \dot{A}_i - \partial_i A_0 = \dot{\alpha}_i + \partial_i \dot{\alpha} - \partial_i A_0 \tag{5.2.22}$$

$$= \dot{\alpha}_i - \partial_i \Phi \tag{5.2.23}$$

$$F_{ij} = \partial_{[i} A_{j]} = \partial_{[i} \alpha_{j]}. \tag{5.2.24}$$

Electric current We also need to perform a scalar-vector decomposition of the electric current

$$J_\mu \equiv \left(\rho_{\mathrm{E}}, \partial_i \mathcal{J} + \mathcal{J}_i\right). \tag{5.2.25}$$

Its conservation $\partial^\mu J_\mu = 0$ now reads

$$\dot{\rho}_{\mathrm{E}} - \partial_i \left(\partial_i \mathcal{J} + \mathcal{J}_i\right) = 0 \tag{5.2.26}$$

$$\dot{\rho}_{\mathrm{E}} = \vec{\nabla}^2 \mathcal{J}. \tag{5.2.27}$$

Maxwell's Equations At this point, we are ready to write down Maxwell's equations $\partial^\mu F_{\mu\nu} = J_\nu$. From eq. (5.2.23), the $\nu = 0$ component is

$$-\partial_i F_{i0} = \partial_i (\dot{\alpha}_i - \partial_i \Phi) = -\vec{\nabla}^2 \Phi = \rho_{\mathrm{E}}. \tag{5.2.28}$$

The $\nu = i$ component of $\partial^\mu F_{\mu\nu} = J_\nu$, according to eq. (5.2.23) and (5.2.24),

$$\partial_0 F_{0i} - \partial_j F_{ji} = \partial_i \mathcal{J} + \mathcal{J}_i \tag{5.2.29}$$

$$\ddot{\alpha}_i - \partial_i \dot{\Phi} - \partial_j \left(\partial_j \alpha_i - \partial_i \alpha_j\right) = \partial_i \mathcal{J} + \mathcal{J}_i \tag{5.2.30}$$

$$\partial^2 \alpha_i - \partial_i \dot{\Phi} = \mathcal{J}_i + \partial_i \mathcal{J}. \tag{5.2.31}$$

As already advertised, we see that the spatial components of Maxwell's equations does admit a scalar-vector decomposition. By the uniqueness argument above, we may read off the "transverse-vector" portion to be

$$\partial^2 \alpha_i = \mathcal{J}_i. \tag{5.2.32}$$

and the "scalar" portion to be

$$-\dot{\Phi} = \mathcal{J}. \tag{5.2.33}$$

We have gotten 3 (groups of) equations – (5.2.28) through (5.2.33) – for 2 sets of variables (Φ, α_i). Let us argue that eq. (5.2.33) is actually redundant. Taking into account eq. (5.2.27), we may take a time derivative of both sides of eq. (5.2.28),

$$-\vec{\nabla}^2 \dot{\Phi} = \dot{\rho}_{\mathrm{E}} = \vec{\nabla}^2 \mathcal{J}. \tag{5.2.34}$$

For the physically realistic case of isolated electric currents, where we may assume implies both $\dot{\Phi} \to 0$ and $\mathcal{J} \to 0$ as the observer-J_i distance goes to infinity, the solution to this above Poisson equation is then unique. This hands us eq. (5.2.33).

To sum: for physically realistic situations in Minkowski spacetime, if we perform a scalar-vector decomposition of the photon vector potential A_μ through eq. (5.2.1) and that of the current J_μ through eq. (5.2.25), we find a gauge-invariant Poisson equation

$$\vec{\nabla}^2\Phi = \rho_{\mathrm{E}}, \qquad \Phi \equiv A_0 - \dot{\alpha}; \tag{5.2.35}$$

as well as a gauge-invariant wave equation

$$\partial^2 \alpha_i = \mathcal{J}_i. \tag{5.2.36}$$

These illuminate the theoretical structure of electromagnetism. As you may recall, our explicit discussions in 4D leading up to the spin-1 modes of eq. (5.1.30) led us to conclude that the non-trivial homogeneous wave solutions of Maxwell's equations are in fact of the "transverse-vector" type. The gauge-invariant formalism for this section thus allows us to identify the source of these spin-1 waves – they are the "transverse-vector" portion of the spatial electric current.

A Acknowledgments

B Conventions

Function argument There is a notational ambiguity whenever we write "f is a function of the variable x" as $f(x)$. If you did not know f were meant to be a function, what is $f(x+\sin(\theta))$? Is it some number f times $x + \sin\theta$? For this reason, in my personal notes and research papers I reserve square brackets exclusively to denote the argument of functions – I would always write $f[x + \sin[\theta]]$, for instance. (This is a notation I borrowed from the software Mathematica.) However, in these lecture notes I will stick to the usual convention of using parenthesis; but I wish to raise awareness of this imprecision in our mathematical notation.

Einstein summation and index notation Repeated indices are always summed over, unless otherwise stated:

$$\xi^i p_i \equiv \sum_i \xi^i p_i. \tag{B.0.1}$$

Often I will remain agnostic about the range of summation, unless absolutely necessary.

In such contexts when the Einstein summation is in force – unless otherwise stated – both the superscript and subscript are enumeration labels. ξ^i is the ith component of $(\xi^1, \xi^2, \xi^3, \dots)$, not some variable ξ raised to the ith power. The position of the index, whether it is super- or sub-script, usually represents how it transforms under the change of basis or coordinate system used. For instance, instead of calling the 3D Cartesian coordinates (x, y, z), we may now denote them collectively as x^i, where $i = 1, 2, 3$. When you rotate your coordinate system $x^i \to R^i{}_j y^j$, the derivative transforms as $\partial_i \equiv \partial/\partial x^i \to (R^{-1})^j{}_i \partial_j$.

Dimensions Unless stated explicitly, the number of space dimensions is D; it is an arbitrary positive integer greater or equal to one. Unless stated explicitly, the number of spacetime dimensions is $d = D + 1$; it is an arbitrary positive integer greater or equal to 2.

Spatial vs. spacetime indices I will employ the common notation that spatial indices are denoted with Latin/English alphabets whereas spacetime ones with Greek letters. Spacetime indices begin with 0; the 0th index is in fact time. Spatial indices start at 1. I will also use the "mostly minus" convention for the metric; for e.g., the flat spacetime geometry in Cartesian coordinates reads

$$\eta_{\mu\nu} = \text{diag}\left[1, -1, \dots, -1\right], \tag{B.0.2}$$

where "diag$[a_1, \dots, a_N]$" refers to the diagonal matrix, whose diagonal elements (from the top left to the bottom right) are respectively a_1, a_2, $\dots$, a_N. Spatial derivatives are $\partial_i \equiv \partial/\partial x^i$; and spacetime ones are $\partial_\mu \equiv \partial/\partial x^\mu$. The scalar wave operator in flat spacetime, in Cartesian coordinates, read

$$\partial^2 = \Box = \eta^{\mu\nu}\partial_\mu\partial_\nu. \tag{B.0.3}$$

The Laplacian in flat space, in Cartesian coordinates, read instead

$$\vec{\nabla}^2 = \delta^{ij}\partial_i\partial_i, \tag{B.0.4}$$

where δ_{ij} is the Kronecker delta, the unit $D \times D$ matrix $\mathbb{I}$:

$$\begin{aligned} \delta_{ij} &= 1, \qquad i = j \\ &= 0, \qquad i \neq j. \end{aligned} \tag{B.0.5}$$

C Last update: November 16, 2017

References

[1] P. A. R. Ade *et al.* [Planck Collaboration], "Planck 2015 results. XIII. Cosmological parameters," Astron. Astrophys. **594**, A13 (2016) doi:10.1051/0004-6361/201525830 [arXiv:1502.01589 [astro-ph.CO]].

[2] E. Poisson, A. Pound and I. Vega, "The Motion of point particles in curved spacetime," Living Rev. Rel. **14**, 7 (2011) doi:10.12942/lrr-2011-7 [arXiv:1102.0529 [gr-qc]].

[3] S. Weinberg, "Gravitation and Cosmology : Principles and Applications of the General Theory of Relativity,"

[4] B. Schutz, "Geometrical Methods of Mathematical Physics," Cambridge University Press; (January 28, 1980)

[5] C. M. Will, "The Confrontation between General Relativity and Experiment," Living Rev. Rel. **17**, 4 (2014) doi:10.12942/lrr-2014-4 [arXiv:1403.7377 [gr-qc]].

[6] G. W. Gibbons, H. Lu, D. N. Page and C. N. Pope, "The General Kerr-de Sitter metrics in all dimensions," J. Geom. Phys. **53**, 49 (2005) doi:10.1016/j.geomphys.2004.05.001 [hep-th/0404008].

[7] I. I. Shapiro, “Fourth Test of General Relativity,” Phys. Rev. Lett. **13**, 789 (1964). doi:10.1103/PhysRevLett.13.789

[8] I. I. Shapiro, G. H. Pettengill, M. E. Ash, M. L. Stone, W. B. Smith, R. P. Ingalls and R. A. Brockelman, “Fourth Test of General Relativity: Preliminary Results,” Phys. Rev. Lett. **20**, 1265 (1968). doi:10.1103/PhysRevLett.20.1265

[9] I. I. Shapiro, M. E. Ash, R. P. Ingalls, W. B. Smith, D. B. Campbell, R. B. Dyce, R. F. Jurgens and G. H. Pettengill, “Fourth test of general relativity - new radar result,” Phys. Rev. Lett. **26**, 1132 (1971). doi:10.1103/PhysRevLett.26.1132

[10] R. P. Woodard, “Ostrogradsky’s theorem on Hamiltonian instability,” Scholarpedia **10**, no. 8, 32243 (2015) doi:10.4249/scholarpedia.32243 [arXiv:1506.02210 [hep-th]].

[11] J. W. York, Jr., “Role of conformal three geometry in the dynamics of gravitation,” Phys. Rev. Lett. **28**, 1082 (1972). doi:10.1103/PhysRevLett.28.1082

[12] G. W. Gibbons and S. W. Hawking, “Action Integrals and Partition Functions in Quantum Gravity,” Phys. Rev. D **15**, 2752 (1977). doi:10.1103/PhysRevD.15.2752