

PHYS415

GENERAL RELATIVITY

*Lecture Notes*

# Contents

<b>0</b>	<b>Physical Overview of General Relativity</b>	<b>1</b>
<b>I</b>	<b>Review of Special Relativity</b>	<b>2</b>
I.1	Lorentz Transformations . . . . .	3
I.1.1	Examples of Lorentz Transformations . . . . .	5
I.1.2	The Restricted Lorentz Group . . . . .	7
I.2	The Scalar Product in Minkowski Spacetime . . . . .	7
I.3	The Causal Structure of Minkowski Spacetime . . . . .	7
<b>II</b>	<b>The Equivalence Principle</b>	<b>10</b>

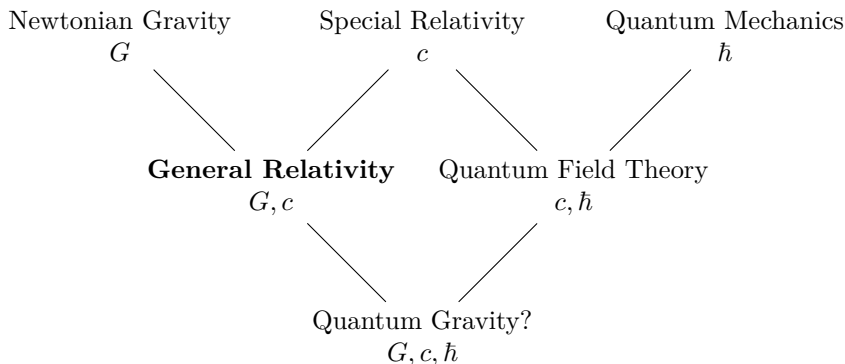
## 0 Physical Overview of General Relativity

IN GENERAL RELATIVITY, gravity is no longer a “force”...

*Space tells matter how to move;*

*Matter tells space how to curve.*

— Misner, Thorne and Wheeler



In Newtonian gravity, escape velocity is given by  $\frac{1}{2}mv^2 = \frac{GMm}{R}$ . Gravity is significant if  $\frac{v^2}{2} \sim \frac{GM}{R}$ . Special relativity is significant when  $v^2 \sim c^2$ . Hence, general relativity is significant when  $\frac{c^2}{2} \sim \frac{GM}{R} \iff R \sim \frac{2GM}{c^2}$ ; this is the *Schwarzschild radius*.

When escape velocity  $\sim$  light velocity, the existence of a black hole is implied.

# I Review of Special Relativity

— *Background* —

See (Schutz, 2009, ch 1) and (Doughty, 2018, ch 5, 12, 13).

Assumptions of Special Relativity:

1. The world is described by a 4-dimensional continuum, **spacetime**, or **Minkowski space**  $\mathcal{M}^4$ , which is the set of all **events**  $x^\mu$ ,

$$x^\mu \equiv \underline{x} = (x^0, x^i) \equiv (x^0, \underline{x}) = (ct, \underline{x}).$$

— *Notation* —

Greek indices,  $\mu, \nu$  run over *spacetime* index values;  $\{0, 1, 2, 3\}$ .

Latin indices (mid-alphabet),  $i, j, k$  run over *spatial* index values;  $\{1, 2, 3\}$ .

$\underline{x}$  is a 4-vector (twiddle);  $\underline{x}$  is a 3-vector (under bar).

2. There exist **inertial frames**; namely, frames in which the measured values of time  $t$  and position  $x^i$  of events result in *linear* equations of motion for *free* particles.

Postulates of Special Relativity:

1. **PRINCIPLE OF RELATIVITY**: The laws of physics are invariant under transformations  $x^\nu \rightarrow x^{\bar{\nu}}(x^\nu)$  from one inertial frame to another (and such transformations form a group).
2. **CONSTANCY OF THE SPEED OF LIGHT**: There exists an invariant upper bound on all velocities

$$\left| \frac{d\underline{x}}{dt} \right| \leq \left| \frac{d\underline{x}}{dt} \right|_{\max} = c \quad (\text{speed of light})$$

and this value is the same in all inertial frames.

□

For photons, (*exercise*)

$$c = \left| \frac{d\underline{x}}{dt} \right|_{\text{frame } K} = \left| \frac{d\underline{x}'}{dt'} \right|_{\text{frame } K'}. \quad (\text{I.1})$$

Rewriting (I.1) we see that, for photons, (*exercise*) □

$$\underbrace{(\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2}_{\text{frame } K} - c^2(\mathrm{d}t)^2 = \underbrace{(\mathrm{d}x')^2 + (\mathrm{d}y')^2 + (\mathrm{d}z')^2}_{\text{frame } K'} - c^2(\mathrm{d}t')^2 = 0.$$

This suggests the definition of the **spacetime interval** between any neighbouring events  $x^\mu$  and  $x^\mu + \mathrm{d}x^\mu$  as

$$\begin{aligned} \mathrm{d}s^2 &\equiv -c^2 \mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \\ &= \eta_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu \end{aligned} \quad \dots \text{pseudo-Riemannian structure}$$

where  $x^\mu = (x^0, x^i) = (ct, x, y, z)$  and where

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & & & \\ & +1 & & \\ & & +1 & \\ & & & +1 \end{pmatrix} \equiv \text{diag}(-1, +1, +1, +1) \equiv -1 \oplus \mathbb{1}_3.$$

For photons,  $\mathrm{d}s^2 = 0$ . Using the two postulates, one may show that the interval  $\mathrm{d}s^2$  is invariant with respect to coordinates based in any inertial frame (Schutz, 2009, §1.6).

$$\mathrm{d}s^2 = \eta_{\mu\nu} \mathrm{d}x^\mu \mathrm{d}x^\nu = \eta_{\bar{\mu}\bar{\nu}} \mathrm{d}x^{\bar{\mu}} \mathrm{d}x^{\bar{\nu}}. \quad (\text{I.2})$$

*Note*

- The symbol  $\mathrm{d}s^2$  for the interval is purely notational convention, since we may have  $\mathrm{d}s^2 < 0$  in some cases.
- The first postulate alone implies either S.R. or its  $c \rightarrow \infty$  limit (a.k.a. Galilean relativity).
- Formally,  $\mathrm{d}x^\mu \mathrm{d}x^\nu$  is shorthand for the *tensor product*  $\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu$ , and you will see that  $\mathrm{d}s^2 = \tilde{\eta}$  is really the *metric tensor*...

## I.1 Lorentz Transformations

In an inertial frame  $K$ , the equations of motion of a free particle are linear;

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\lambda^2} = 0 \quad \Longleftrightarrow \quad x^\mu = x_0^\mu + u^\mu \lambda,$$

where  $x_0^\mu, u^\mu$  are constant and  $\lambda$  is a parameter. Similarly, in any other inertial frame  $\bar{K}$ ,

$$\frac{d^2 x^{\bar{\mu}}}{d\lambda^2} = 0 \quad \Longleftrightarrow \quad x^{\bar{\mu}} = x_0^{\bar{\mu}} + u^{\bar{\mu}} \lambda.$$

Now,

$$\begin{aligned} \frac{dx^{\bar{\mu}}}{d\lambda} &= \frac{\partial x^{\bar{\mu}}}{\partial x^\nu} \frac{dx^\nu}{d\lambda} = \frac{\partial x^{\bar{\mu}}}{\partial x^\nu} u^\nu && \dots \text{chain rule} \\ \Rightarrow 0 &= \frac{d^2 x^{\bar{\mu}}}{d\lambda^2} = \frac{\partial}{\partial x^\alpha} \left( \frac{\partial x^{\bar{\mu}}}{\partial x^\nu} u^\nu \right) u^\alpha = \frac{\partial^2 x^{\bar{\mu}}}{\partial x^\alpha \partial x^\nu} u^\nu u^\alpha && \because u^\nu \text{ constant} \\ \Rightarrow 0 &= \frac{\partial^2 x^{\bar{\mu}}}{\partial x^\alpha \partial x^\nu}. && \dots \text{since true } \forall u^\alpha \end{aligned}$$

So the required transformation between inertial frames is *linear*;

$$\boxed{x^{\bar{\mu}} = L^{\bar{\mu}}{}_\nu x^\nu + a^{\bar{\mu}}}, \quad (\text{I.3})$$

where  $a^{\bar{\mu}}$  and  $L^{\bar{\mu}}{}_\nu \equiv \frac{\partial x^{\bar{\mu}}}{\partial x^\nu}$  are constants. Differentiate (I.3) and substitute into (I.2) to give  $\eta_{\mu\nu} dx^\mu dx^\nu = \eta_{\bar{\mu}\bar{\nu}} L^{\bar{\mu}}{}_\mu L^{\bar{\nu}}{}_\nu dx^\mu dx^\nu$ . Since this is true  $\forall dx^\mu$ ,

$$\boxed{\eta_{\bar{\mu}\bar{\nu}} L^{\bar{\mu}}{}_\mu L^{\bar{\nu}}{}_\nu = \eta_{\mu\nu}}. \quad (\text{I.4})$$

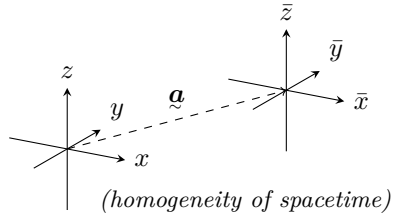
In matrix form, (I.3) and (I.4) are

$$\boxed{\tilde{\mathbf{x}} = \mathbf{L}\tilde{\mathbf{x}} + \mathbf{a}}, \quad (\text{I.3a})$$

$$\boxed{L^\top \eta L = \eta}. \quad (\text{I.4a})$$

Transformations  $\mathbf{x} \mapsto \tilde{\mathbf{x}}$  defined by (I.3), (I.4) are the *inhomogeneous* Lorentz transformations, or **Poincaré transformations**, and form the Poincaré group IO(1, 3) (pronounced Inhomogeneous Orthogonal group).

“Inhomogeneous” refers to the inclusion of spacetime translations  $x^\nu \mapsto x^{\bar{\mu}} = \delta^{\bar{\mu}}{}_\nu x^\nu + a^{\bar{\mu}}$ , which form a subgroup  $T^4$  of the Poincaré group. If we set  $a^{\bar{\mu}} = 0$  in (I.3), we are left with *homogeneous* transformations, called simply the **Lorentz transformations**.



In G.R., our task is to generalise these ideas to general coordinate frames for which  $L^{\bar{\mu}}{}_\nu$  are not necessarily constant.

### I.1.1 Examples of Lorentz Transformations

A transformation belonging to the (homogeneous) Lorentz group  $O(1, 3)$  can be represented as a matrix acting on coordinates  $x^\mu$  when they are *viewed as vectors*.

$$x^\mu \cong \underline{x} = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}$$

*Note*

Spacetime events (i.e., *points* in spacetime) are not themselves vectors—neither addition nor scalar multiplication of events makes physical sense (i.e., spacetime itself is not a *vector space*). However, in S.R. we may represent events by their associated displacement vector relative to a chosen orthogonal inertial frame.

The Lorentz group  $O(1, 3)$  consists of (combinations of) the following:

- **Rotations**, e.g., about the  $z$ -axis (in the  $xy$ -plane) by an angle  $\theta$ ;

$$L_R(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

- **Boosts**, e.g., by a velocity  $\underline{v}$  in the  $x$ -direction;

$$L_B(\alpha) = \begin{pmatrix} \cosh \alpha & -\sinh \alpha & 0 & 0 \\ -\sinh \alpha & \cosh \alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos i\alpha & i \sin i\alpha & 0 & 0 \\ i \sin i\alpha & \cos i\alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where  $\alpha = \tanh^{-1} \frac{v}{c}$  is the *rapidity parameter*. In terms of the velocity  $\beta \equiv \frac{v}{c}$ , one has  $\cosh \alpha = \frac{1}{\sqrt{1-\beta^2}} \equiv \gamma$  and  $\sinh \alpha = \frac{\beta}{\sqrt{1-\beta^2}} \equiv \beta\gamma$ .

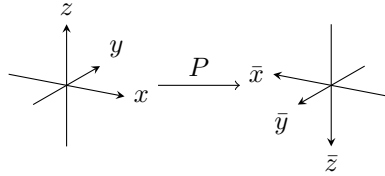
*Note*

Boosts appear similar to rotations, but differ as a consequence of the indefiniteness of the metric. Formally, an  $x$ -boost of rapidity  $\alpha$  is equivalent to a rotation by an ‘imaginary angle’  $\alpha' = i\alpha$  through the  $\tau x$ -plane, where  $\tau = it$  is ‘imaginary time’... though this is not a good picture physically!

An important difference between rotations and boosts is that, where  $0 \leq \theta < 2\pi$  for rotations, we have  $-\infty < \alpha < \infty$  for boosts, i.e., rotations form a *compact*<sup>1</sup> subgroup of the Lorentz (or Poincaré) group, whereas boosts are *non-compact*—and in fact do not form a subgroup (because, in general, the composition of two boosts forms a combination of a rotation and a boost).

Both rotations and boosts depend on continuous parameters ( $\theta$  or  $\alpha$ ). However, the Lorentz group  $O(1, 3)$  also contains *discrete* transformations...

- **Parity inversion;**  $P = \text{diag}(1, -\mathbf{1}_3) \equiv \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}$ .



Notice that  $P$  is *not* equivalent to a rotation; it transforms a right-handed frame into a left and vice versa, since  $\det P = -1$ .

— *Note* —

In even spatial dimensions, the transformation  $R = \text{diag}(1, -\mathbf{1}_{2n})$  is *not* a parity transformation; it is a rotation by  $\pi$  and  $\det R = 1$ . In these cases, inversions  $x_i \mapsto -x_i$  of a *single* spatial coordinate are parity transformations.

- **Time reversal;**  $T = \text{diag}(-1, \mathbf{1}_3)$ .

The Lorentz matrix condition (I.4) implies that  $(\det L)^2 = 1 \iff \det L = \pm 1$  for any Lorentz transformation  $L \in O(1, 3)$ . Those with  $\det L = +1$  and those with  $\det L = -1$  form two disconnected pieces of  $O(1, 3)$ , but only the first piece contains the identity transformation  $\delta^{\bar{\mu}}_{\nu}$ .

Inspecting the  $\bar{\mu}\nu = \bar{0}0$  component of (I.4) gives

$$\left(L^{\bar{0}}_0\right)^2 - \sum_{\bar{k}=1}^3 \left(L^{\bar{k}}_0\right)^2 = 1 \implies \left(L^{\bar{0}}_0\right)^2 \geq 1,$$

---

<sup>1</sup>A *compact* set is one for which any infinite sequence of elements contains a convergent subsequence. E.g.,  $[0, 2\pi]$  is compact, but  $\mathbb{R}$  is not (consider the sequence  $\{1, 2, 3, \dots\} \subset \mathbb{R}$ ).



which shows that there exists two disconnected classes of Lorentz transformation with  $L^{\bar{0}}_0 \geq 1$  and  $L^{\bar{0}}_0 \leq -1$ . Those with  $L^{\bar{0}}_0 \geq 1$  are called **orthochronous**.

### I.1.2 The Restricted Lorentz Group

We define the subgroup of restricted Lorentz transformations by adding two conditions to the Lorentz matrix condition (I.4); that they be 1) orthochronous and 2) have determinant unity.

$$\text{SO}^+(1, 3) \equiv \left\{ \Lambda \mid \Lambda^\top \eta \Lambda = \eta, \Lambda^{\bar{0}}_0 \geq 1, \det \Lambda = 1 \right\}$$

We have removed the discrete transformations involving  $P$  and  $T$ , so that  $\text{SO}^+(1, 3)$  is *continuous* and *connected*, unlike  $\text{O}(1, 3)$ .

— Notation —

The  $S$  in  $\text{SO}^+(1, 3)$  refers to the condition  $\det \Lambda = 1$ , and the  $^+$  refers to the orthochronous condition. Sometimes  $\text{SO}^+(1, 3)$  is simply written as  $\text{SO}(1, 3)$ .

Groups whose elements may be continuously parametrised are **Lie groups**. The translation group  $T^4$  (parametrised continuously by  $\Delta x^\mu$ ) and the rotation group  $\text{SO}(3)$  (parametrised continuously by three angles) are examples of *connected* Lie groups.

Reintroducing translations to the restricted Lorentz group gives the restricted Poincaré group  $\text{ISO}^+(1, 3)$ —also a connected Lie group. Unrestricted groups may be reconstructed by reintroducing the discrete transformations;

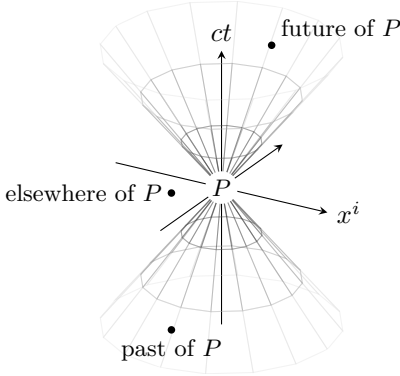
$$\text{O}(1, 3) = \{ \Lambda, \Lambda P, \Lambda T, \Lambda P T \mid \Lambda \in \text{SO}^+(1, 3) \}.$$

## I.2 The Scalar Product in Minkowski Spacetime

## I.3 The Causal Structure of Minkowski Spacetime

Because of the indefiniteness of the Minkowski metric (and hence of the Minkowski scalar product), 4-vectors can have positive, zero or negative norm.

A 4-vector  $\tilde{V}$  is called  $\begin{cases} \text{timelike} \\ \text{null} \\ \text{spacelike} \end{cases}$  if  $\tilde{V} \cdot \tilde{V} = V_\mu V^\mu = \eta_{\mu\nu} V^\mu V^\nu \begin{cases} < 0 \\ = 0 \\ > 0 \end{cases}$ .



□

One spatial dimension suppressed; lightcone in (3 + 1)-d spacetime is a continuum of spheres.

exists a Lorentz transformation which sends a timelike unit vector to  $V^\mu \mapsto V^{\bar{\mu}} = (1, 0, 0, 0)$  (*exercise*). The **proper time interval**  $d\tau$  between two neighbouring events along the particle's worldline is defined as the interval of time measured in the particle's instantaneous rest frame.

□

By the second postulate, all *causally connected* events relative to an event  $P$  lie in its future or past lightcone, because  $P$  cannot causally influence events outside the lightcone without transmitting superluminal data.

A particle's **worldline** is a curve  $x^\mu = x^\mu(\lambda)$  whose tangent  $\tilde{V}$  given by  $V^\mu = \frac{dx^\mu}{d\lambda}$  is everywhere timelike;  $V^\mu V_\mu < 0$ . This implies that, in an inertial frame, the particle's velocity  $\underline{v}$  is subluminal  $|\underline{v}| < c$  (*exercise*).

Consider a freely moving particle whose (timelike) worldline unit tangent vector is  $\tilde{V}$ . There always

$$c^2 d\tau^2 = -\eta_{\mu\nu} dx^\mu dx^\nu \quad (\text{I.5})$$

— Note —

Since  $dx$  is timelike,  $\eta_{\mu\nu} dx^\mu dx^\nu < 0$ , hence the minus sign. Proper time is not defined for spacelike intervals, since the above definition would yield an imaginary time. Instead, proper distance is defined by  $d\ell^2 = \eta_{\mu\nu} dx^\mu dx^\nu$ .

Proper time is the time *experienced* by the particle; i.e., the time that would be measured by a clock moving on the same worldline.

Suppose we have a worldline  $x^\mu(\lambda)$ . Integrate  $d\tau$  via (I.5) along the path to get the total proper time elapsed between events  $x^\mu(\lambda_i)$  and  $x^\mu(\lambda_f)$ .

$$\Delta\tau = \frac{1}{c} \int_{\lambda_i}^{\lambda_f} d\lambda \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} \quad (\text{I.6})$$

*Note*

The proper time elapsed depends on the path taken between the two events.

Particle worldlines can be characterised by an action principle: variation of (I.6) with respect to the trajectories,  $\delta x^\mu(\lambda)$ , yields the equations of motion  $\frac{\partial^2 x^\mu}{\partial \lambda^2} = 0$  (holding  $\delta x^\mu = 0$  fixed at both ends) (*exercise*). In fact, the extremised trajectory minimises the proper time—*freely falling particles take the path of maximum proper time*. □

By (I.5),  $c^2 d\tau^2 = -dt^2 + \sum_i (dx^i)^2 = dt^2(-1 + \underline{v} \cdot \underline{v})$  where  $v^i \equiv \frac{dx^i}{dt}$  so

$$d\tau = \sqrt{1 - \frac{v^2}{c^2}} dt \equiv \gamma dt$$

is the proper time of a frame moving at velocity  $\underline{v}$  relative to a frame with coordinates  $(ct, x^i)$ .

## II The Equivalence Principle

## References

- N. Doughty. *Lagrangian interaction: an introduction to relativistic symmetry in electrodynamics and gravitation*. CRC Press, 2018.
- B. Schutz. *A first course in general relativity*. Cambridge university press, 2009.