

PHYS415

GENERAL RELATIVITY

*Lecture Notes*

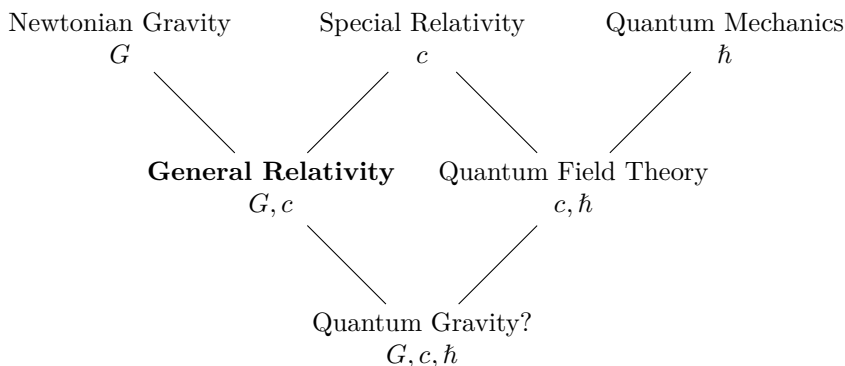
# Contents

|           |  |           |
|-----------|--|-----------|
| <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 . . . . .                            | 4         |
| I.1.1     | Examples of Lorentz Transformations . . . . .                | 5         |
| I.1.2     | The Restricted Lorentz Group . . . . .                       | 7         |
| I.2       | The Scalar Product in Minkowski Spacetime . . . . .          | 8         |
| I.3       | Relativistic Fields . . . . .                                | 9         |
| I.4       | The Causal Structure of Minkowski Spacetime . . . . .        | 10        |
| I.4.1     | 4-velocity . . . . .   | 12        |
| I.4.2     | 4-momentum . . . . .   | 12        |
| I.4.3     | Angular momentum . . . . .                                   | 13        |
| I.5       | Gauß's Theorem in Minkowski Spacetime . . . . .              | 13        |
| I.6       | Conserved Currents . . . . .                                 | 14        |
| <b>II</b> | <b>The Equivalence Principle</b>                             | <b>15</b> |
| II.1      | Newtonian Mechanics & the Weak Equivalence Principle . . . . | 15        |

## 0 Physical Overview of General Relativity

IN GENERAL RELATIVITY, gravity is no longer a “force”, but an effect of the *curvature of spacetime*.

*Space tells matter how to move;  
Matter tells space how to curve.*  
— Misner, Thorne and Wheeler



**Large masses + high speeds  $\Rightarrow$  general relativity**

In Newtonian gravity, escape velocity is given by  $\frac{1}{2}mv^2 = \frac{GMm}{R}$ . Hence, gravity is significant if  $\frac{v^2}{2} \sim \frac{GM}{R}$ . Special relativity is significant when  $v^2 \sim c^2$ . General relativity, at the intersection of the two, is hence significant when  $\frac{c^2}{2} \sim \frac{GM}{R} \iff R \sim \frac{2GM}{c^2}$ ; this is the *Schwarzschild radius*. Where the escape velocity coincides with the speed of light, 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<sup>1</sup>, **spacetime**, or **Minkowski space**  $\mathcal{M}^4$ , which is the set of all **events**  $x^\mu$ ,

$$x^\mu \equiv \mathfrak{x} = (x^0, x^i) \equiv (x^0, \underline{x}) = \underset{\substack{\text{time, space}}}{(ct, \underline{x})} .$$

— *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;  $\{1, 2, 3\}$ .

$\mathfrak{x}$  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  $x^\nu \rightarrow x^{\bar{\mu}}(x^\nu)$  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.

---

<sup>1</sup>Mathematically, a *Lorentzian manifold*...

For photons,

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

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

$$(dx)^2 + (dy)^2 + (dz)^2 - c^2(dt)^2 = (dx')^2 + (dy')^2 + (dz')^2 - c^2(dt')^2 = 0.$$

$\text{frame } K$  $\text{frame } K'$

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

$$\begin{aligned} ds^2 &\equiv -c^2 dt^2 + dx^2 + dy^2 + dz^2 \\ &= \eta_{\mu\nu} dx^\mu dx^\nu \end{aligned} \quad \dots \text{pseudo-Riemannian structure}$$

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,  $ds^2 = 0$ . Using the two postulates, one may show that the interval  $ds^2$  is invariant with respect to coordinates based in any inertial frame (Schutz, 2009, §1.6).

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu = \eta_{\bar{\mu}\bar{\nu}} dx^{\bar{\mu}} dx^{\bar{\nu}}. \quad (\text{I.2})$$

*Note*

- The symbol  $ds^2$  for the interval is purely notational convention (it is not the square of ‘ $ds$ ’) since we may have  $ds^2 < 0$  in some cases.
- The first postulate alone implies either S.R. or its  $c \rightarrow \infty$  limit (a.k.a. Galilean relativity).
- Formally,  $dx^\mu dx^\nu$  is shorthand for the *tensor product*  $dx^\mu \otimes dx^\nu$ , and you will see that  $ds^2 \equiv \tilde{\eta}$  is really the *metric tensor*...

## I.1 Lorentz Transformations

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

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

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  $\forall dx^\mu$ ,

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

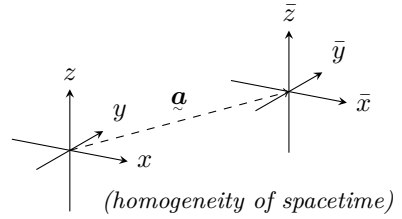
In matrix form, (I.3) and (I.4) are

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

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

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

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

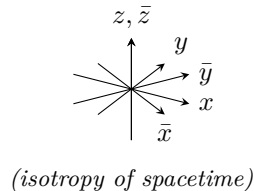
#### Note

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* —

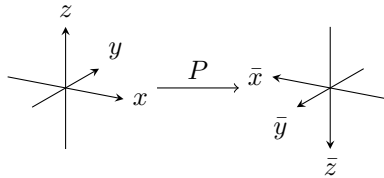
Boosts appear similar to rotations, but differ as a consequence of the indefiniteness of the metric: they are *hyperbolic rotations*. E.g., consider a vector space with signature  $(-+)$ ; a ‘rotation’ which preserves the inner product  $(\begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} c \\ d \end{bmatrix}) \mapsto -ac + bd$  has the same form as the  $tx$ -submatrix of  $L_B(\alpha)$ .

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 useful picture physically!

An important difference between rotations and boosts is that, whereas  $0 \leq \theta < 2\pi$  for rotations, we have  $-\infty < \alpha < \infty$  for boosts, i.e., rotations form a *compact*<sup>2</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  $\det P = -1$ .

<sup>2</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  $\{1, 2, 3, \dots\} \subset \mathbb{R}$ ).



*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 has  $\det R = 1$ . In all dimensions, inversions  $x_i \mapsto -x_i$  of a *single* spatial coordinate are each 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  $L \in \text{O}(1, 3)$ . Those with  $\det L = +1$  and those with  $\det L = -1$  form two disconnected pieces of  $\text{O}(1, 3)$ , but only the first piece contains the identity transformation  $\delta^\mu_\nu$ .

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

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

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 orthochronality. 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;

$$\text{O}(1, 3) = \{A, AP, AT, APT \mid A \in \text{SO}^+(1, 3)\}.$$

## I.2 The Scalar Product in Minkowski Space-time

We define a scalar product

$$\text{d}s^2 = \eta_{\mu\nu} \text{d}x^\mu \text{d}x^\nu = \text{d}x_\mu \text{d}x^\mu \equiv \underline{\text{d}\mathbf{x}} \cdot \underline{\text{d}\mathbf{x}},$$

where

$$\text{d}x_\mu = \eta_{\mu\nu} \text{d}x^\nu = (-\text{d}x^0, \text{d}x^1, \text{d}x^2, \text{d}x^3) \quad (\text{I.5})$$

is the dual component form.

We refer to  $\{\text{d}x^\mu\}$  as the *contravariant* components, and  $\{\text{d}x_\mu\}$  as the *covariant* components of the vector  $\underline{\text{d}\mathbf{x}}$ . Thus the metric  $\eta_{\mu\nu}$  acts as an *index-lowering operator*<sup>3</sup>.

The identity operation (a Lorentz transformation of course) is the Kronecker delta  $\delta^\mu{}_\nu$  in indexed form.

Thus  $\delta^\mu{}_\nu = (\eta^{-1}\eta)^\mu{}_\nu = (\eta^{-1})^{\mu\lambda} \eta_{\lambda\nu}$  by applying the rule that the indices of  $\eta$  are downstairs. As a matrix,  $\eta^{-1} = \eta = \text{diag}(-1, \mathbb{1})$ , but we write  $\eta^{-1} = (\eta^{-1})^{\mu\lambda} = \eta^{\mu\lambda}$  (with indices raised) so that levels of indices are consistent, i.e.

$$\eta^{\mu\lambda} \eta_{\lambda\nu} = \delta^\mu{}_\nu. \quad (\text{I.6})$$

By (I.5) and (I.6) it follows that

$$\text{d}x^\mu = \eta^{\mu\lambda} \text{d}x_\lambda,$$

i.e.  $\eta^{\mu\lambda}$  is an index-raising operator.

---

<sup>3</sup>N.B.  $\{\text{d}x_\mu\}$  are *not* differentials of coordinates.

Under a Poincaré transformation  $x^\mu \rightarrow x^{\bar{\mu}}(x^\mu)$ ,

$$dx^{\bar{\mu}} = L^{\bar{\mu}}{}_{\mu} dx^{\mu}, \quad \text{where} \quad L^{\bar{\mu}}{}_{\mu} = \frac{\partial x^{\bar{\mu}}}{\partial x^{\mu}} \in O(1, 3). \quad (\text{I.7})$$

The inverse transformation taking  $x^{\bar{\mu}}$  back to  $x^\mu$  just swaps indices  $\bar{\mu} \leftrightarrow \mu$  in above. Thus the natural notation for the Lorentz matrix inverse to  $L^{\bar{\mu}}{}_{\mu}$  is  $(L^{-1})^{\mu}{}_{\bar{\mu}} \equiv L^{\mu}{}_{\bar{\mu}} = \frac{\partial x^{\mu}}{\partial x^{\bar{\mu}}}$ . By the chain rule:

$$L^{\bar{\mu}}{}_{\mu} L^{\mu}{}_{\bar{\nu}} = \delta^{\bar{\mu}}{}_{\bar{\nu}}, \quad L^{\mu}{}_{\bar{\mu}} L^{\bar{\mu}}{}_{\nu} = \delta^{\mu}{}_{\nu}.$$

From (I.5), (I.7), and (I.4a) it follows that (*exercise*) under (I.7) □

$$dx_{\bar{\mu}} = L^{\mu}{}_{\bar{\mu}} dx_{\mu}.$$

## I.3 Relativistic Fields

We define relativistic fields by their transformation properties under *restricted* Lorentz (or restricted Poincaré) transformations. Suppose such a transformation  $x^\mu \rightarrow x^{\bar{\mu}} = \Lambda^{\bar{\mu}}{}_{\mu} x^{\mu} + a^{\bar{\mu}}$  is defined, where  $\Lambda^{\bar{\mu}}{}_{\mu} = \frac{\partial x^{\bar{\mu}}}{\partial x^{\mu}} \in SO(1, 3)$ .

- *Scalar field.* A single function of coordinates  $x^\mu$  which is Lorentz (or Poincaré) invariant,

$$\varphi(x) \rightarrow \bar{\varphi}(\bar{x}) = \varphi(x).$$

— *Example* —

If  $x \rightarrow \bar{x} = x + a$ , then  $\bar{\varphi}(x + a) = \varphi(x) \iff \bar{\varphi}(x) = \varphi(x - a)$ .

If  $x \rightarrow \bar{x} = \Lambda x$ , then  $\bar{\varphi}(\Lambda x) = \varphi(x) \iff \bar{\varphi}(x) = \varphi(\Lambda^{-1}x)$ .

- *4-vector field.* A quantity  $\mathbf{V}$  with 4 components labelled  $V^\mu$  or  $V_\mu$  (related by  $V_\mu = \eta_{\mu\nu} V^\nu$  and  $V^\mu = \eta^{\mu\nu} V_\nu$ ), where
  - *contravariant* components transform as  $V^{\bar{\mu}} = \Lambda^{\bar{\mu}}{}_{\mu} V^\mu$ ; and
  - *covariant* components transform as  $V_{\bar{\mu}} = \Lambda_{\bar{\mu}}{}^{\mu} V_\mu$ .

*Example*

If  $x \rightarrow \bar{x} = \Lambda x$ , then  $V^{\bar{\mu}}(\bar{x}) = \Lambda^{\bar{\mu}}_{\mu} V^{\mu}(x)$  or  $V^{\bar{\mu}}(x) = \Lambda^{\bar{\mu}}_{\mu} V^{\mu}(\Lambda^{-1}x)$ . An example of a covariant 4-vector (“co-vector”) is the partial derivative operator,  $\partial_{\bar{\mu}} = \Lambda^{\mu}_{\bar{\mu}} \partial_{\mu}$ .

- *4-tensor field.* A rank  $n$  tensor is an object with  $4^n$  components transforming as

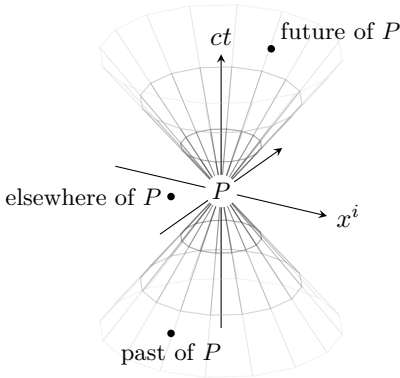
$$T^{\bar{\mu}_1 \bar{\mu}_2 \dots \bar{\nu}_1 \bar{\nu}_2 \dots} = \Lambda^{\bar{\mu}_1}_{\mu_1} \Lambda^{\bar{\mu}_2}_{\mu_2} \dots \Lambda^{\nu_1}_{\bar{\nu}_1} \Lambda^{\nu_2}_{\bar{\nu}_2} \dots T^{\mu_1 \mu_2 \dots \nu_1 \nu_2 \dots},$$

where all indices are raised (or lowered) by contraction with the Minkowski metric  $\eta^{\mu\nu}$  (or  $\eta_{\mu\nu}$ ), respectively.

## I.4 The Causal Structure of Minkowski Space-time

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



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  $\underline{V}$  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 particles velocity  $\underline{v}$  is subluminal  $|\underline{v}| < c$  (*exercise*).

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

Consider a freely moving particle whose (timelike) worldline unit

tangent vector is  $\underline{V}$ . There always 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. □

$$c^2 d\tau^2 = -\eta_{\mu\nu} dx^\mu dx^\nu \quad (\text{I.8})$$

*Note*

Since  $d\mathbf{x}$  is timelike here,  $\eta_{\mu\nu} dx^\mu dx^\nu < 0$ , so (I.8) contains a minus sign. Proper time is not defined for spacelike intervals, since this would involve  $\sqrt{-1}$ . Instead, proper distance is defined by  $dl^2 = \eta_{\mu\nu} dx^\mu dx^\nu$ .

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.8) along the path to get the total proper time elapsed between events  $x^\mu(\lambda_i)$  and  $x^\mu(\lambda_f)$ .

$$\Delta\tau = \frac{1}{c} \int_{\lambda_i}^{\lambda_f} d\lambda \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} \quad (\text{I.9})$$

*Note*

The proper time elapsed depends on the path taken between the two events. (See the Twin Paradox.)

Particle worldlines can be characterised by an action principle: variation of (I.9) 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.8),  $c^2 d\tau^2 = -dt^2 + \sum_i (dx^i)^2 = dt^2(-1 + \underline{v} \cdot \underline{v})$  where  $v^i \equiv \frac{dx^i}{dt}$  so

$$d\tau = \sqrt{1 - \frac{v^2}{c^2}} dt \equiv \gamma dt \quad (\text{I.10})$$

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

### I.4.1 4-velocity

Since proper time  $d\tau$  is an *invariant*, we are motivated to define the 4-vector

$$u^\mu \equiv \frac{dx^\mu}{d\tau} = \frac{dt}{d\tau} \left( c, \frac{d\mathbf{x}}{dt} \right) = \gamma(c, \underline{v}),$$

where  $\gamma$  is defined as in (I.10). In the particle's instantaneous rest frame,  $\|\underline{v}\|^2 = u^i u_i = 0$ , so that  $u^\mu u_\mu = u^0 u_0 = -c^2$ . However, since this is an invariant (it is a scalar, which does not transform under Poincaré transformations), we have

$$\boxed{\|\underline{u}\|^2 \equiv \underline{u} \cdot \underline{u} \equiv u^\mu u_\mu = -c^2}$$

in *all frames*.

### I.4.2 4-momentum

Relativistic 4-momentum is defined by

$$\underline{p} \equiv m\underline{u} = \left( \frac{E}{c}, \underline{p} \right), \quad E = \gamma mc^2, \quad \underline{p} = \gamma m\underline{v}. \quad (\text{I.11})$$

This form of 4-momentum may be deduced from an action using the Hamiltonian approach. We want to extremise (I.2), but first multiply by  $mc^2$  to get the units right. The action is

$$S = -mc^2 \int_{\underline{x}_i}^{\underline{x}_f} d\tau = \int_{t_i}^{t_f} dt L$$

where the Lagrangian is

$$L = L(\underline{x}, \dot{\underline{x}}; t) = -mc^2 \sqrt{1 - \frac{v^2}{c^2}} = -mc^2 \sqrt{1 - \frac{\underline{x} \cdot \underline{x}}{c^2}}.$$

Then, the Hamiltonian (total energy) and conjugate momentum are

$$\begin{aligned} E = \underline{p} \cdot \underline{x} - L &= \gamma mv^2 + \frac{mc^2}{\gamma} = \gamma mc^2, \\ \underline{p} = \frac{\partial L}{\partial \dot{\underline{x}}} &= mc^2 \frac{1}{2} \frac{2\dot{\underline{x}}}{\sqrt{1 - v^2/c^2}} = \gamma m\underline{v}, \end{aligned}$$

matching the definition (I.11).

In all rest frames,

$$\|\underline{p}\|^2 = \underline{p} \cdot \underline{p} = p^\mu p_\mu = -m^2 c^2$$

as can be shown by evaluating in the particle's instantaneous rest frame and using scalar invariance.

For photons with  $m = 0$ , the 4-momentum  $\underline{p}$  is a null vector. Photons have no instantaneous rest frame—but we can always find a frame in which  $\underline{p} = \frac{E}{c}(1, 0, 0, 1)$ .

### I.4.3 Angular momentum

The covariant orbital angular momentum of a particle about an event  $\underline{x}_0 = \{x_0^\mu\}$  is defined as

$$j^{\mu\nu} = (x^\mu - x_0^\mu)p^\nu - (x^\nu - x_0^\nu)p^\mu,$$

where  $x^\mu$  and  $p^\mu$  are the particle's position and momentum. This is an antisymmetric tensor,  $j^{\mu\nu} = -j^{\nu\mu} = j^{[\mu\nu]}$ , with 6 independent components:  $j^{0k}, j^{k\ell}, k < \ell$ .

#### Example

About the origin  $x_0^\mu = 0$ , we have  $j^{ij} = x^i p^j - x^j p^i$  which is equivalent to  $j^{ij} = \epsilon^{ijk} \ell_k$  where  $\underline{\ell} = \underline{x} \times \underline{p}$ . The  $j^{0k}$  components represent the motion of the centre of mass:  $j^{0k} = x^0 p^k - x^k p^0 = \gamma m c (v^k t - x^k)$ .

## I.5 Gauß's Theorem in Minkowski Spacetime

In three dimensional Euclidean space, Gauß's theorem is

$$\int_R d^3x \nabla \cdot \underline{V} = \oint_{\partial R} d\underline{S} \cdot \underline{V},$$

where  $\underline{V}$  is a 3-vector and  $R \subset \mathbb{R}^3$  is a region of Euclidean space with boundary  $\partial R$ . The surface element  $d\underline{S}$  is normal to the boundary  $\partial R$ . In index form,

$$\int_R d^3x \partial_i V^{i\dots} = \oint_{\partial R} dS_i V^{i\dots},$$

where dots denote possible free indices, which are not involved. The proof of Gauß's theorem does not require 3 dimensions or a  $+++$  metric signature. In Minkowski spacetime  $\mathcal{M}^4$ , Gauß's theorem reads

$$\int_R d^4x \partial_\mu V^{\mu\dots} = \oint_{\partial R} d\Sigma_\mu V^{\mu\dots},$$

where  $\partial R$  is now the boundary of a simply-connection region  $R \subset \mathcal{M}^4$  with normal surface element  $d\underline{\Sigma}$ .

In 4-dimensional spacetime, we often choose the boundary  $\partial R$  so that it has three parts:  $\Sigma^+$ , spacelike surface toward future;  $\Sigma^-$ , spacelike surface toward past;  $\partial R^\infty$ , timelike surface at spatial infinity, joining  $\Sigma^+$  and  $\Sigma^-$ .

*Note*

The outward normal to  $\Sigma^-$  is *past-pointing*.

This decomposition is convenient if the fields involved vanish at spatial infinity, in which case the contribution from  $\partial R^\infty$  vanishes.

## I.6 Conserved Currents

A 4-vector field  $\underline{J} = J^\mu \partial_\mu$  which is divergence free

$$\operatorname{div} \underline{J} \equiv \frac{\partial J^\mu}{\partial x^\mu} \equiv \partial_\mu J^\mu \equiv J^\mu{}_{,\mu} = 0 \quad (\text{I.12})$$

is called a **4-current**. Equation (I.12) is a conservation law. By Gauß's theorem, supposing that contributions at spatial infinity vanish,

$$0 = \int_R d^4x \partial_\mu J^\mu = \int_{\Sigma^-} d\Sigma_\mu J^\mu + \int_{\Sigma^+} d\Sigma_\mu J^\mu.$$

If we define  $\Sigma_f := \Sigma^+$  and  $\Sigma_p$  to be  $\Sigma^-$  but future-pointing instead of past-pointing, i.e.,  $\int_{\Sigma_p} d\Sigma_\mu J^\mu = -\int_{\Sigma^-} d\Sigma_\mu J^\mu$ , then the equation above becomes

$$\int_{\Sigma_p} d\Sigma_\mu J^\mu = \int_{\Sigma_f} d\Sigma_\mu J^\mu = \int_{t \text{ const.}} d^3x J^0$$



## II The Equivalence Principle

### II.1 Newtonian Mechanics & the Weak Equivalence Principle

In the Newtonian theory, particles affected by gravity alone exhibit the equations of motion

$$\ddot{\underline{x}} = -\nabla\Phi(x) \qquad \text{or} \qquad \ddot{x}^k = -\partial^k\Phi$$

## 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.