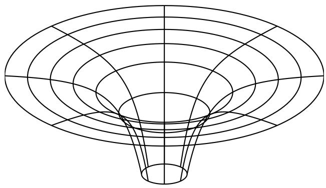


Phys415

General Relativity

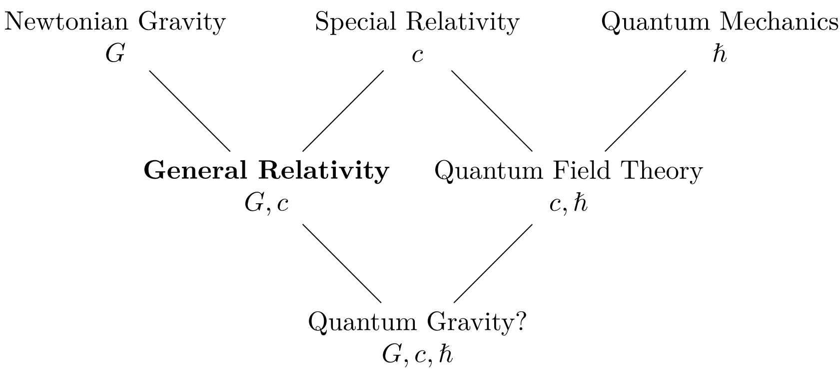
Lecture Notes

Contents

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 if implied.

I Review of Special Relativity

Background

See (?, ch 1) and (?, ch 5, 12, 13).

Assumptions of Special Relativity:

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

$$x^\mu \equiv \underset{\sim}{\boldsymbol{x}} = \left(x^0, x^i\right) \equiv \left(x^0, \vec{x}\right) = \underset{\text{time, space}}{\left(ct, \vec{x}\right)} .$$

Notation

Greek indices, μ, ν run over *spacetime* index values; $\{0, 1, 2, 3\}$.
Latin indices (mid-alphabet), i, j, k run over *spatial* index values; $\{1, 2, 3\}$.
$\underset{\sim}{\boldsymbol{x}}$ is a 4-vector (twiddle); $\vec{x}$ 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 $x^\nu \to x^{\bar{\mu}}(x^\nu)$ 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

$$\left|\frac{\mathrm{d}\vec{x}}{\mathrm{d}t}\right| \le \left|\frac{\mathrm{d}\vec{x}}{\mathrm{d}t}\right|_{\max} = c \quad \text{(speed of light)}$$

and this value is the same in all inertial frames.

[1] Mathematically, a *Lorentzian manifold*...

For photons,

$$c = \underset{\text{frame } K}{\left|\frac{\mathrm{d}\vec{x}}{\mathrm{d}t}\right|} = \underset{\text{frame } K'}{\left|\frac{\mathrm{d}\vec{x}'}{\mathrm{d}t}\right|}. \tag{I.1}$$

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

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

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

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

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

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

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 (**?**, §1.6).

$$\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu = \eta_{\bar{\mu}\bar{\nu}}\mathrm{d}x^{\bar{\mu}}\mathrm{d}x^{\bar{\nu}}. \tag{I.2}$$

Note

- The symbol $\mathrm{d}s^2$ for the interval is purely notational convention, since we may have $\mathrm{d}s^2 < 0$ is some cases.
- The first postulate alone implies either S.R. or its $c \to \infty$ 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 $\mathrm{d}s^2 = \underset{\sim}{\boldsymbol{\eta}}$ is really the *metric tensor*...

I.1 Lorentz Transformations

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

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

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

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

Now,

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

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

$$\boxed{x^{\bar\mu} = L^{\bar\mu}{}_\nu x^\nu + a^{\bar\mu},} \tag{I.3}$$

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

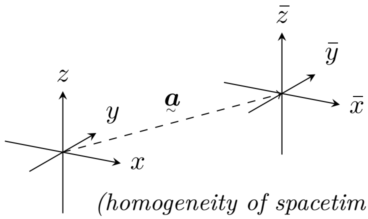
$$\boxed{\eta_{\bar\mu\bar\nu} L^{\bar\mu}{}_\mu L^{\bar\nu}{}_\nu = \eta_{\mu\nu}.} \tag{I.4}$$

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

$$\boxed{\begin{aligned} \underset{\sim}{\bar{\boldsymbol{x}}} &= L\underset{\sim}{\boldsymbol{x}} + \underset{\sim}{\boldsymbol{a}}, && \text{(I.3a)} \\ L^\top \eta L &= \eta. && \text{(I.4a)} \end{aligned}}$$

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



(homogeneity of spacetime)

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 $\mathrm{O}(1,3)$ can be represented as a matrix acting on coordinates x^μ when they are *viewed as vectors.*

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

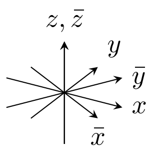
Note

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 $\mathrm{O}(1,3)$ consists of (combinations of) the following:

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

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



(isotropy of spacetime)

- **Boosts,** e.g., by a velocity $\vec{v}$ 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 $\sinh\alpha = \frac{\beta}{\sqrt{1-\beta^2}} \equiv \beta\gamma$.

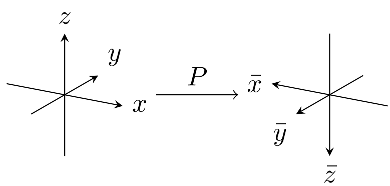
Note

Boosts appear similar to rotations, but differ as a consequence of the indefiniteness of the metric. Formally, an x-boost of rapidity α is equivalent to a rotation by an 'imaginary angle' $\alpha' = i\alpha$ through the τx-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*[2] 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 (θ or α). However, the Lorentz group $\mathrm{O}(1,3)$ also contains *discrete* transformations...

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



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

[2] 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 $\{1, 2, 3, \ldots\} \subset \mathbb{R}$).

Note

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

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

The Lorentz matrix condition (I.4) implies that $(\det L)^2 = 1 \iff \det L = \pm 1$ for any Lorentz transformation $L \in \mathrm{O}(1,3)$. Those with $\det L = +1$ and those with $\det L = -1$ form two disconnected pieces of $\mathrm{O}(1,3)$, 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

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

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.

$$\mathrm{SO}^{+}(1,3) \equiv \left\{ \Lambda \;\middle|\; \Lambda^{\top}\eta\Lambda = \eta, \Lambda^{\bar{0}}{}_{0} \geq 1, \det \Lambda = 1 \right\}$$

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

Notation

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

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

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

$$\mathrm{O}(1,3) = \left\{ \Lambda, \Lambda P, \Lambda T, \Lambda PT \;\middle|\; \Lambda \in \mathrm{SO}^+(1,3) \right\}.$$

I.2 The Scalar Product in Minkowski Spacetime

We define a scalar product

$$\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = \mathrm{d}x_\mu \mathrm{d}x^\mu \equiv \mathrm{d}\underset{\sim}{\boldsymbol{x}} \cdot \mathrm{d}\underset{\sim}{\boldsymbol{x}},$$

where

$$\mathrm{d}x_\mu = \eta_{\mu\nu}\mathrm{d}x^\nu = \left(-\mathrm{d}x^0, \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}x^3\right) \tag{I.5}$$

is the dual component form.

We refer to $\{\mathrm{d}x^\mu\}$ as the *contravariant* components, and $\left\{\mathrm{d}x_\mu\right\}$ as the *covariant* components of the vector $\mathrm{d}\underset{\sim}{\boldsymbol{x}}$. Thus the metric $\eta_{\mu\nu}$ acts as an *index-lowering operator*[3].

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

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

$$\eta^{\mu\lambda}\eta_{\lambda\nu} = \delta^\mu{}_\nu. \tag{I.6}$$

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

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

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

[3] N.B. $\left\{\mathrm{d}x_\mu\right\}$ are *not* differentials of coordinates.

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

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

The inverse transformation taking $x^{\bar\mu}$ back to x^μ 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:

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

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

$$\mathrm{d}x_{\bar\mu} = L^\mu{}_{\bar\mu} \mathrm{d}x_\mu.$$

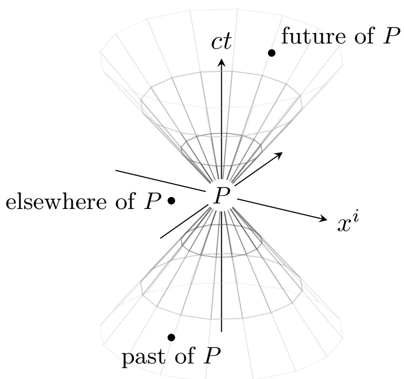
I.3 The Causal Structure of Minkowski Spacetime

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

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 $\underset{\sim}{\boldsymbol{V}}$ given by $V^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}$ is everywhere timelike; $V^\mu V_\mu < 0$. This implies that, in an inertial frame, the particles velocity $\vec{v}$ is subluminal $|\vec{v}| < c$ *(exercise)* . □



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

Consider a freely moving particle whose (timelike) worldline unit tangent vector is $\underset{\sim}{\boldsymbol{V}}$. 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** $\mathrm{d}\tau$ between two neighbouring events along the particle's worldline is defined as the interval of time measured in the particle's instantaneous rest frame.

$$c^2 \mathrm{d}\tau^2 = -\eta_{\mu\nu} \mathrm{d}x^{\mu} \mathrm{d}x^{\nu} \tag{I.8}$$

Note

Since $\mathrm{d}x$ is timelike, $\eta_{\mu\nu}\mathrm{d}x^{\mu}\mathrm{d}x^{\nu} < 0$, so (I.8) contains a minus sign. Proper time is not defined for spacelike intervals, since this would yield an imaginary time. Instead, proper distance is defined by $\mathrm{d}\ell^2 = \eta_{\mu\nu}\mathrm{d}x^{\mu}\mathrm{d}x^{\nu}$.

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 $\mathrm{d}\tau$ via (I.8) along the path to get the total proper time elapsed between events $x^{\mu}(\lambda_i)$ and $x^{\mu}(\lambda_f)$.

$$\Delta\tau = \frac{1}{c} \int_{\lambda_i}^{\lambda_f} \mathrm{d}\lambda \sqrt{-\eta_{\mu\nu} \frac{\mathrm{d}x^{\mu}}{\mathrm{d}\lambda} \frac{\mathrm{d}x^{\nu}}{\mathrm{d}\lambda}} \tag{I.9}$$

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), $c^2\mathrm{d}\tau^2 = -\mathrm{d}t^2 + \sum_i (\mathrm{d}x^i)^2 = \mathrm{d}t^2(-1 + \vec{v} \cdot \vec{v})$ where $v^i \equiv \frac{\mathrm{d}x^i}{\mathrm{d}t}$ so

$$\mathrm{d}\tau = \sqrt{1 - \frac{v^2}{c^2}} \mathrm{d}t \equiv \gamma \mathrm{d}t \tag{I.10}$$

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

I.3.1 4-velocity

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

$$u^\mu \equiv \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}t}{\mathrm{d}\tau}\left(c, \frac{\mathrm{d}\vec{x}}{\mathrm{d}t}\right) = \gamma(c, \vec{v}),$$

where γ is defined as in (I.10). In the particle's instantaneous rest frame, $\|\vec{v}\|^2 = u^i u_i = 0$, 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

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

in *all frames.*

I.3.2 4-momentum

Relativistic 4-momentum is defined by

$$\underset{\sim}{\boldsymbol{p}} \equiv m\underset{\sim}{\boldsymbol{u}} = \left(\frac{E}{c}, \vec{p}\right), \qquad E = \gamma mc^2, \qquad \vec{p} = \gamma m\vec{v}. \tag{I.11}$$

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

$$S = -mc^2 \int_{\underset{\sim}{\boldsymbol{x}}_i}^{\underset{\sim}{\boldsymbol{x}}_f} \mathrm{d}\tau = \int_{t_i}^{t_f} \mathrm{d}t L$$

where the Lagrangian is

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

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

$$E = \vec{p} \cdot \dot{\vec{x}} - L = \gamma mv^2 + \frac{mc^2}{\gamma} = \gamma mc^2,$$

$$\vec{p} = \frac{\partial L}{\partial \dot{\vec{x}}} = mc^2 \frac{1}{2} \frac{2\dot{\vec{x}}}{\sqrt{1 - v^2/c^2}} = \gamma m\vec{v},$$

matching the definition (I.11).

In all rest frames,

$$\boxed{\|\underset{\sim}{\boldsymbol{p}}\|^2 = \underset{\sim}{\boldsymbol{p}} \cdot \underset{\sim}{\boldsymbol{p}} = p^\mu p_\mu = -m^2 c^2}$$

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 $\underset{\sim}{\boldsymbol{p}}$ is a null vector. Photons have no instantaneous rest frame—but we can always find a frame in which $\underset{\sim}{\boldsymbol{p}} = \frac{E}{c}(1, 0, 0, 1)$.

I.3.3 Angular momentum

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

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

where x^μ and p^μ 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 $\vec{l} = \vec{x} \times \vec{p}$. j^{0k} represents the motion of the centre of mass: $j^{0k} = x^0 p^k - x^k p^0 = \gamma mc(v^k t - x^k)$.

I.4 Gauß's Theorem in Minkowski Spacetime

In three dimensional Euclidean space, Gauß's theorem is

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

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

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

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

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

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

II The Equivalence Principle