

Basic Summary Notes: General Relativity

By Shikang Ni

DISCLAIMER

This set of notes is by far INCOMPLETE representation of the tricks and understanding you need to know about the topic

WISDOM

The purpose of this note is to give you a quick reminder/summary of what was taught when you are revising (KNOWING)
Actual understanding comes after CONNECTING and GENERALISING

This can be done by practicing, mindmapping and reading

Other great and important sources of information are:

1. Einstein Gravity in a Nutshell (Zee)
2. General Relativity - An Introduction for Physicists (Hobson) + [Solutions](#)
3. Additional readings/notes: [Sean Carroll](#)
4. Past exam papers (higher order corrections, beware of over-fitting)
5. Example sheets (build awareness/first-order practice)
6. Pen and paper derivations/proofs

Jacobi's formula: $\frac{\partial}{\partial x_a} \det(M) = \det(M) \cdot \text{Tr} \left(M^{-1} \frac{\partial M}{\partial x^a} \right)$. In indices form:

$$\partial_\mu \det(g) = \det(g) \cdot g^{ab} \partial_\mu g_{ab}$$

Course Content

1. Special relativity

Mathematical pre-requisites

2. Manifolds and coordinates

“Special relativity tells us to think of space and time as a 4-dimensional continuum called Minkowski spacetime, which is an example a manifold. We explore the geometric properties of the types of manifold relevant for GR...”

3. Vectors and tensors algebra

“Our goal is to formulate physical laws in such a way that they reduce to SR in locally-inertial coordinates. the most efficient way to do this is to write down equations that are true in a general coordinate system. Such a coordinate-independent, or geometric, approach naturally give rise to vector-valued fields. These are geometric objects that are independent of the choice of coordinate system. We shall learn to generalise familiar Euclidean ideas to define vector and tensors in general Riemannian manifolds and arbitrary coordinate systems...”

4. Vector and tensor calculus

“The laws of physics are differential equations involving tensor-valued objects. On a general manifold, tensors at different points inhabit different vector space and there is no unique way of comparing tensors at different points. We shall learn to construct covariant derivatives of tensors which connect together spaces at different points...”

- Covariant derivatives
- Intrinsic derivative of vector along curve
- Geodesic curves / parallel transport

Applications

5. Particles dynamics

“With the machinery of tensor algebra and calculus in place, we shall first apply this to SR and express the theory more formally. The spacetime of SR is a pseudo-Euclidean manifold, over which we can globally define Cartesian coordinates...”

6. Electromagnetism

“We shall develop electromagnetism as a relativistic field theory on Minkowski spacetime. We will combine 3D Maxwell’s equations into 4D tensor equations and show that the theory is relativistically covariant...”

Gravity

7. Spacetime curvature

“Gravity enters only via our inability to construct such coordinates, through the intrinsic curvature of a manifold... We now shall learn to formulate laws of physics on a curved spacetime as tensor equations, to ensure consistency with the equivalence principle...”

8. Introduction to Gravitational field equations

“We shall see how the curvature of spacetime is related to the matter that is present. This will lead us to the Einstein equations, relating the energy-momentum tensor of the matter to the Einstein tensor...”

9. Schwarzschild solution

“The Einstein field equations are non-linear PDE which are hard to solve. However, when the spacetime possesses symmetries, we can find exact solutions. Schwarzschild found the first exact solution that represents the spacetime in the vacuum region outside a spherically symmetric mass distribution...”

10. Classic tests of general relativity

“Two classical tests of GR are the perihelion advance of the planet Mercury and the bending of light by the Sun.”

- Shapes of orbits for massive and massless particles
- Precession of planetary orbits (Mercury)
- Bending of light

11. Schwarzschild black holes

“Schwarzschild metric is singular at $r = 2\mu$ and we have explored motion in the region $r > 2\mu$. We now explore the casual structure of the Schwarzschild solution and finds that region $r < 2\mu$ represents a black hole and surface $r = 2\mu$ represents an event horizon. A new coordinate system will be introduced to join the regions $r < 2\mu$ and $r > 2\mu$ in a continuous manner...”

12. Cosmology

“We shall apply GR to model the Universe. On sufficiently large spatial scales, the Universe looks remarkably symmetric in space. This symmetry allows us to find analytic solutions for the spacetime of a smoothed-out universe, forming the starting points for all cosmological studies... where smaller scale structures can be treated as perturbations...”

1 Lorentz transformation/boost

Derivation

Most general form for two Cartesian frames in standard configuration:

$$t' = At + Bx \quad , \quad x' = A(x - vt) \quad , \quad y' = y \quad , \quad z' = z$$

For special relativity, the interval squared is invariant:

$$c^2t^2 - x^2 - y^2 - z^2 = c^2t'^2 - x'^2 - y'^2 - z'^2$$

The constants can be solved for to find Lorentz transformation:

$$ct' = \gamma(ct - \beta x) \quad , \quad x' = \gamma(x - \beta ct) \quad , \quad y' = y \quad , \quad z' = z$$

Matrix form:
$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \quad \text{or} \quad (x^\mu)' = L_\nu^\mu x^\nu$$

where L_ν^μ are the elements of the general Lorentz transformation matrix. The invariance of the interval constrains $\det(L) = \pm 1$ and $L_0^0 = \pm 1$.

Generalised Lorentz Boost

Consider components of \mathbf{r} parallel and perpendicular to \mathbf{v} .

$$ct' = \gamma(ct - \boldsymbol{\beta} \cdot \mathbf{r}) \quad \text{and} \quad \mathbf{r}'_{\parallel} = \gamma(\mathbf{r}_{\parallel} - \boldsymbol{\beta}ct) \quad \text{and} \quad \mathbf{r}'_{\perp} = \mathbf{r}_{\perp}$$

Then,

$$\mathbf{r}' = \mathbf{r}'_{\parallel} + \mathbf{r}'_{\perp}$$

Rapidity, ϕ

Intuition behind rapidity: can be thought of as angles in spacetime, so we can consider Lorentz boost as rotations in 4D spacetime:

$$\phi = \tanh^{-1} \beta = \frac{1}{2} \ln \left(\frac{1 + \beta}{1 - \beta} \right) \quad \text{and} \quad \cosh \phi = \gamma \quad \text{and} \quad \sinh \phi = \gamma\beta$$

We can rewrite the Lorentz transformation in terms of rapidity. This allows us to perform many standard Lorentz boost and link reference frames easily it's just adding up rapidities: $\phi_u = \phi_{u'} + \phi_v$.

If the axis are not aligned, perform rotation to align their axis, now the systems can be related by a standard Lorentz boost.

Consequences

1. Length contraction ($\Delta t = 0$): $l = l_0/\gamma$
2. Time dilation ($\Delta x' = 0$):

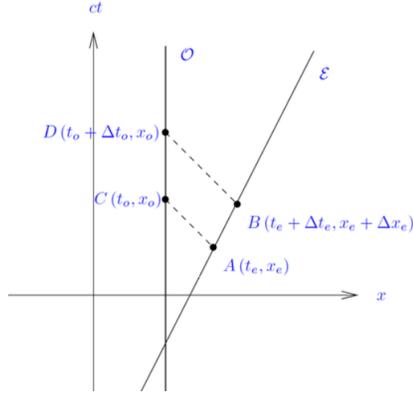
$$\Delta t = \gamma \Delta \tau \quad \text{and also,} \quad \boxed{\frac{dt}{d\tau} = \gamma}$$

3. Similarity transformation of Lorentz transformation for boost in arbitrary direction:

$$L_\theta = RL_xR^{-1}$$

where R are rotation matrix(ices).

4. Doppler shift - ratio of proper time experienced by observer and emitter between two wavecrest
 $\left(\frac{f}{f_0} = \frac{\Delta\tau_{AB}}{\Delta\tau_{CD}}\right)$:



[Derivation] $\frac{f}{f_0} = \sqrt{\frac{1-\beta}{1+\beta}}$

5. Expression for velocity and its addition:

$$u'_x = \frac{dx'}{dt'} = \frac{u_x - v}{1 - u_x v/c^2} \quad \text{and} \quad u'_y = \frac{dy'}{dt'} = \frac{u_y}{\gamma(1 - u_x v/c^2)}$$

6. Acceleration in special relativity (derived by considering differentials of the velocity):

$$du'_x = \frac{du_x}{\gamma^2(1 - u_x v/c^2)^2} \quad \text{and} \quad du'_y = \frac{du_y}{\gamma(1 - u_x v/c^2)} + \frac{u_y v du_x}{c^2 \gamma(1 - u_x v/c^2)^2}$$

Of course not forgetting (using Lorentz transform for time),

$$dt' = \gamma(1 - u_x v/c^2) dt$$

Acceleration is not invariant in SR, but all observers will agree whether a particle is accelerating or not.

7. (Derivation) Rectilinear acceleration: we can Lorentz transform from S to an instantaneous rest frame (IRF) S'.

In IRF: $u' = 0$ and $\frac{du'}{d\tau} := f(\tau)$ and $v = u(t)$

$$\frac{du}{d\tau} = \frac{f(\tau)}{\gamma^2} \Rightarrow c \frac{d\psi}{d\tau} = f(\tau) \quad \text{so} \quad c\psi = \int_0^\tau f(\tau') d\tau'$$

$$\frac{dt}{d\tau} = \cosh \psi(\tau) \quad , \quad \frac{dx}{d\tau} = c \sinh \psi(\tau)$$

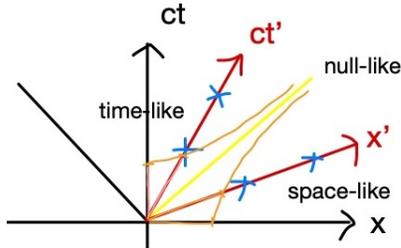
The set of equation can be solved to find $t(\tau)$ and $x(\tau)$.

2 Minkowski spacetime

Special relativity has led us to model the physical world as a four-dimensional spacetime continuum, with Minkowski geometry. It has intrinsic geometry expressed in line element form of:

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2$$

Particle worldlines



$$ds^2 > 0 \quad (\text{timelike})$$

$$ds^2 = 0 \quad (\text{lightlike})$$

$$ds^2 < 0 \quad (\text{spacelike})$$

Transforming from $S \rightarrow S'$:

$$x' = 0 \Rightarrow x = \beta ct \quad (\text{t'axis}) \quad \text{and} \quad ct' = 0 \Rightarrow ct = \beta x \quad (\text{x'axis})$$

Consequences

1. Causality is preserved for $ds^2 \geq 0$

2. Useful identity: $\gamma^2 = \beta^2 \gamma^2 + 1 \Rightarrow \boxed{\gamma^2 v^2 = \gamma^2 c^2 - c^2}$

3 Riemannian Manifolds

Indices upstairs are coordinates and indices downstairs lives in the tangent space. Summation is conducted over one indices being upstairs and the other one being downstairs.

Coordinates relabelling

The choice of coordinates is arbitrary and what's important is the geometrical relationship between them. We can relabel the points using a coordinate transformation:

$$dx^k = \left(\frac{\partial x^k}{\partial x'^b} \right) dx'^b$$

where $J^{-1} = \frac{\partial x^k}{\partial x'^b}$ is an $n \times n$ transformation matrix, with Jacobian of the transformation being $J = \det[J] = \det \left[\frac{\partial x'^k}{\partial x^b} \right]$.

Expression of line element

Metric function

ds^2 is invariant. $g_{ab}(x)$ is the metric tensor in a chosen coordinate system. g_{ab} is symmetric. In Riemannian manifold, the interval takes the form:

$$\boxed{ds^2 = g_{ab}(x) dx^a dx^b}$$

E.g. If g_{ab} is diagonal i.e. in an orthogonal coordinate system:
The line element takes the form:

$$ds^2 = g_{11}(dx^1)^2 + g_{22}(dx^2)^2 + \dots + g_{NN}(dx^N)^2 \quad \text{or} \quad ds^2 = dx^2 + dy^2 + dz^2 + dw^2$$

E.g. In Cartesian coordinates, $g_{kk} = 1$: $ds^2 = dx^k dx^k$

E.g. Coordinate transformation affects metric function like:

$$g'_{cd}(x') = g_{ab}(x) \frac{\partial x^a}{\partial x'^c} \frac{\partial x^b}{\partial x'^d} = g_{ab}(x) h_{ac} h_{bd} = h_{ca}^T g_{ab}(x) h_{bd}$$

$$\text{In matrix form: } G' = X^T G X$$

“Volume” element

$$d^N V = \sqrt{|g|} dx^1 dx^2 \dots dx^N$$

Using:

$$G' = X^T G X \quad \text{and} \quad J = \det \left[\frac{\partial x'_a}{\partial x_b} \right] = (\det X)^{-1} \Rightarrow g' = \frac{1}{J^2} g$$

where $\det[G] = g$. Since $g' = \pm 1$ so $g = \pm J^2$. Hence equivalently,

$$d^N V = dx'^1 dx'^2 \dots dx'^N = J dx^1 dx^2 \dots dx^N$$

Tangent space, $T_P(\mathcal{M})$

At each point P of a general ND manifold M, we can construct a ND vector space – the tangent space $T_P(\mathcal{M})$ – whose elements are (local) vectors.

The tangent space is locally Cartesian so $G' = \text{diag}(\pm 1, \pm 1, \dots, \pm 1)$ and so it is characterised by its line element:

$$ds^2 = \pm(dx^1)^2 \pm (dx^2)^2 \pm \dots \pm (dx^N)^2$$

Local Cartesian/inertial coordinates [Zee Chapter 1.6]

$$g'_{\mu\nu} = \begin{cases} \delta_{\mu\nu} + \mathcal{O}[(x - x_P)^2], & \text{if manifold is Riemannian/Euclidean } (\Delta s^2 > 0) \\ \eta_{\mu\nu} + \mathcal{O}[(x - x_P)^2], & \text{if manifold is pseudo-Riemannian } (\Delta s^2 < 0) \end{cases}$$

where $\eta_{\mu\nu} = \text{diag}(\pm 1, \dots, \pm 1)$. For Minkowski spacetime, the line element has the form: $ds^2 = d(ct)^2 - dx^2 - dy^2 - dz^2$, so $g_{ab} = \eta_{ab} = (+1, -1, -1, -1)$.

[Proof] Locally, we can find:

$$\left. \frac{\partial g_{ab}}{\partial x^c} \right|_P = 0 \quad \text{such that} \quad g_{ab}(x) = \delta_{ab} + \mathcal{O}[(x - x_P)^2]$$

The higher order big-O terms are what give rise to curvature as these are the terms that we can't get rid of via the coordinates. By the time we reach third order derivatives of our coordinates, we no longer have enough independent values to account for the number of independent values in the second order derivatives of our metric.

4 Vector and tensor algebra

Change the way you think about vectors: think of a vector at a point P as a **differential operator** there, which maps scalar fields on \mathcal{M} to a number. By extension, a vector field is associated with a differential operator at every point and maps scalar fields to scalar fields. Intuitively, it is the directionality of the differential operator that captures the idea of vectors as having an associated direction.

Local vector field $\mathbf{v}(x)$ can be expanded at each point into contravariant components $v^a(x)$ in terms of basis vector $\mathbf{e}_a(x)$ or covariant components $v_a(x)$ in terms of dual basis vector $\mathbf{e}^a(x)$:

$$\mathbf{v} = \underbrace{v^a}_{\text{component}} \underbrace{\frac{\partial}{\partial x^a}}_{\text{basis}} \stackrel{\text{invariance}}{=} v'^a \frac{\partial}{\partial x'^a} = v^a \mathbf{e}_a = v_a \mathbf{e}^a$$

with $\mathbf{e}^a(x) \cdot \mathbf{e}_b(x) = \delta_b^a$, $\mathbf{v} \cdot \mathbf{e}^a = v^a$ and $\mathbf{v} \cdot \mathbf{e}_a = v_a$. We can see that such operator is closed under addition and scalar multiplication, so the space of all such operators at P is closed and forms a vector space. In this way, we have explicitly constructed the tangent space $T_P(\mathcal{M})$, with $\{\partial/\partial x^1, \dots, \partial/\partial x^N\}$ as the basis vectors and v^a being the components.

Vectors (upstairs) are

$$\text{objects that transform as: } v'^a = \underbrace{\frac{\partial x'^a}{\partial x^b}}_{\text{transformation matrix}} v^b$$

So, any N-tuple that transforms according to this forms the components of a vector. For example: the vector/differential operator associated with tangent vector to a curve $x^a(u)$, which has components, dx^a/du is $\frac{dx^a}{du} \frac{\partial}{\partial x^a} = \frac{d}{du}$.

Dual vectors (downstairs) are linear maps (vectors \rightarrow numbers),

$$\text{it transforms as: } X'_a = \frac{\partial x^b}{\partial x'^a} X_b$$

Dual vectors at P live in a different vector space, called the dual vector space, $T_P^*(\mathcal{M})$.

[Exercise] Contraction of: $X_a v^a$ is invariant under coordinate transformation

Tensor fields

Tensor of type (k, l) has k “upstairs” indices (contravariant) and l “downstairs” indices (covariant). It has rank $k + l$. It is a geometrical object with the transformation property:

$$\text{tensors transform as: } T_{c\dots d}^{\prime a\dots b} = \left(\frac{\partial x'^a}{\partial x^p} \dots \frac{\partial x'^b}{\partial x^q} \right) \left(\frac{\partial x^r}{\partial x'^c} \dots \frac{\partial x^s}{\partial x'^d} \right) T_{r\dots s}^{p\dots q}$$

Rank-0 tensors are scalars, type-(1,0) tensors are vector and type-(0,1) are dual vectors. Think of tensors as geometric objects that are invariant under changes of coordinates.

Symmetrisation of tensors

$$S_{ab} = \begin{cases} S_{ba}, & \text{symmetric} \\ -S_{ba}, & \text{antisymmetric} \end{cases}$$

We can always decompose a tensor into a sum of symmetric $S_{[ab]}$ and antisymmetric $S_{(ab)}$ parts: $S_{ab} = S_{(ab)} + S_{[ab]}$ such that:

$$S_{(ab\dots c)} = \frac{1}{n!}(\text{sum over all permutations of a,b,\dots c})$$

$$S_{[ab\dots c]} = \frac{1}{n!}(\text{alternating sum over all permutations})$$

where term is *+ve* for even permutation and *-ve* for odd permutation.

We can also (anti-)symmetrise on a subset of indices. E.g. $S_{(ab)c} = \frac{1}{2}(S_{abc} + S_{bac})$.

Quotient theorem: is it a tensor?

If an object - contract with an any tensor is a tensor. Then the original object is a tensor.

Covariant derivatives

[Exercise] Consider the transformation of $\partial v^{tb} / \partial x'^a$, we realise it does not fully transform as a tensor due to an annoying extra term. So the derivative is not a tensor.

$$\partial'_a v^{tb} = \partial'_a \left(\frac{\partial x^{tb}}{\partial x^c} v^c \right) = \dots \text{product rule} \dots = \left(\frac{\partial x^d}{\partial x'^a} \partial_d \right) \left(\frac{\partial x^{tb}}{\partial x^c} v^c \right) + \frac{\partial x^d}{\partial x'^a} \frac{\partial^2 x^{tb}}{\partial x^d \partial x^c} v^c$$

We want to work with derivatives that transforms as a tensor. Thus, this motivates the construction of covariant derivative designed to cancel the unwanted final term on the right.

$$\text{We just want to do: } \partial_a(v^b \mathbf{e}_b) = (\partial_a v^b) \mathbf{e}_b + (\partial_a \mathbf{e}_b) v^b$$

On a general manifold, tensors at different points inhabit separate tangent spaces so the derivative of the basis vectors is non-zero. We take care of that by introducing the connection, Γ_{ac}^b .

$$\nabla_a v^b = \partial_a v^b + \Gamma_{ac}^b v^c \quad \text{and} \quad \nabla_a X_b = \partial_a X_b - \Gamma_{ab}^c X_c$$

where $\Gamma_{ac}^b (= \mathbf{e}^b \cdot \partial_c \mathbf{e}_a)$ is called the **connection**. Neither of the two terms individually are tensors, but together they transform as a tensor.

[Exercise] To see the form of the connection, consider forming the covariant derivative in the new coordinates:

$$\begin{aligned} \nabla'_a v^{tb} &= (\partial'_a v^{tb}) + \Gamma_{ac}^{tb} v^c = \left[(\partial'_a x^d) (\partial_c x^{tb}) (\partial_d v^c) + (\partial'_a x^d) (\partial_d \partial_c x^{tb}) \right] + \Gamma_{ac}^{tb} (\partial_d x'^c) v^d \\ &= (\partial'_a x^d) (\partial_c x^{tb}) \nabla_d v^c - \underbrace{(\partial'_a x^d) (\partial_c x^{tb}) \Gamma_{de}^c v^e + (\partial'_a x^d) (\partial_d \partial_c x^{tb}) + \Gamma_{ac}^{tb} (\partial_d x'^c) v^d}_{\text{must vanish}} \end{aligned}$$

Properties:

- Covariant derivative of a scalar is the gradient:

$$\nabla_a \phi = \frac{\partial \phi}{\partial x^a}$$

- Linearity:

$$\nabla_c(\alpha T + \beta S) = \alpha \nabla_c T + \beta \nabla_c S$$

- Leibnitz rule (product rule):

$$\nabla_c(TS) = (\nabla T) \otimes S + T \otimes (\nabla S)$$

Applying to a general rank-2 tensor:

$$\nabla_c T^{ab} = \partial_c T^{ab} + \Gamma_{cd}^a T^{db} + \Gamma_{cd}^b T^{ad}$$

$$\nabla_c T_b^a = \partial_c T_b^a + \Gamma_{cd}^a T_b^d - \Gamma_{cb}^d T_d^a = 0$$

$$\nabla_c T_{ab} = \partial_c T_{ab} - \Gamma_{ca}^d T_{db} - \Gamma_{cb}^d T_{ad}$$

Metric connection

On a manifold equipped with a metric, there is two further constraints which singles out a natural connection. They are:

1. Metric compatibility i.e. metric stays the same:

$$\nabla_a g_{bc} = 0$$

2. Commutative action on scalar fields \Rightarrow symmetric in its lower indices i.e. torsion-free:

$$\nabla_a \nabla_b \phi = \nabla_b \nabla_a \phi \quad \Rightarrow \quad \Gamma_{ab}^c = \Gamma_{ba}^c$$

$$\text{Proof: } \nabla_{[a} \nabla_{b]} \phi = 0 \text{ and } \nabla_a \nabla_b \phi = \partial_a \partial_b \phi - \Gamma_{ab}^c \partial_c \phi$$

[Derivation] We can now find the **metric connection/Christoffel symbol**, by first writing down $0 = \nabla_c g_{ab}$:

$$\Gamma_{bc}^a = \frac{1}{2} g^{ad} (\partial_b g_{dc} + \partial_c g_{db} - \partial_d g_{bc})$$

Other useful properties of the metric connection:

1. Can interchange the order of raising/lowering indices and covariant differentiation: $\nabla_c T^{ab} = \nabla_c (g^{bd} T_d^a) = \dots$ product rule $\dots = g^{bd} (\nabla_c T_d^a)$
2. $\nabla^a = g^{ac} \nabla_c$
3. By Jacobi's formula: $g^{ab} \partial_c g_{ab} = g^{-1} \partial_c g = 2\Gamma_{ac}^a$ where $g = \det(g_{ab})$
4. $\nabla_a = \partial_a$ in local Cartesian coordinates because the derivative of the metric vanishes so the connection vanishes (equivalence principle).

Defining divergence, curl and the Laplacian

The familiar operations of taking the divergence and curl of a vector field, and the Laplacian, generalise to tensor calculus on manifolds.

$$\begin{aligned}\nabla \cdot \mathbf{v} &= \nabla_a v^a = \text{scalar} \\ (\nabla \times \mathbf{X})_{ab} &= \nabla_a X_b - \nabla_b X_a = \partial_a X_b - \partial_b X_a \quad \text{for a symmetric connection} \\ \nabla^2 \varphi &= \nabla^a \nabla_a \varphi = g^{ab} \nabla_b \nabla_a \varphi \quad \text{and} \quad \nabla^2 T^{ab} = \nabla^c \nabla_c T^{ab} = g^{cd} \nabla_c \nabla_d T^{ab}\end{aligned}$$

Intrinsic derivative - is just covariant derivative with parametrisation

Given curve parameterised by parameter u , the intrinsic derivative of \mathbf{v} along curve $x^a(u)$ is obtained by contracting the tangent vector to the curve with the covariant derivative of \mathbf{v} (**chain rule**):

$$\frac{Dv^a}{Du} = \frac{dx^b}{du} \nabla_b v^a = \underbrace{\frac{dx^b}{du} \frac{\partial v^a}{\partial x^b}}_{\frac{dv^a}{du}} + \frac{dx^b}{du} \Gamma_{bc}^a v^c = \frac{dv^a}{du} + \frac{dx^b}{du} \Gamma_{bc}^a v^c$$

In local Cartesian coord. $\frac{Dv^a}{Du} \rightarrow \frac{dv^a}{du}$. For type (1,1) tensor as an example of its application,

$$\frac{DT_b^a}{Du} = \frac{dx^c}{du} \nabla_c T_b^a = \frac{dT_b^a}{du} + \frac{dx^c}{du} (\Gamma_{cd}^a T_b^d - \Gamma_{cb}^d T_d^a)$$

Parallel transport

Parallel transport is defined where the intrinsic derivative $\frac{D}{Du}$ vanishes.

$$\frac{Dv^a}{Du} = 0 = \frac{dv^a}{du} + \frac{dx^b}{du} \Gamma_{bc}^a v^c$$

- Change in components are independent of parametrisation:

$$\frac{Dv^a}{Du} = 0 = \frac{dv^a}{du} + \frac{dx^b}{du} \Gamma_{bc}^a v^c \quad \Rightarrow \quad \delta v^a = -\delta x^b \Gamma_{bc}^a v^c$$

- Length, $g_{ab} v^a v^b$ is preserved:

$$\frac{D}{Du} (\underbrace{g_{ab} v^a v^b}_{\text{inner product}}) = 2g_{ab} v^a \frac{Dv^b}{Du} = 0$$

More generally, inner product of two parallel transported vector is preserved.

- On a surface embedded in Euclidean space, parallel transport from a point A to an infinitesimally-separated point B corresponds to parallel transport in the embedding space followed by projection into the surface at B (General Theory of Relativity by Dirac)

Geodesic curves

Geodesic on a manifold are the **generalisation of “straight lines”**. They are defined as:

- (1) Curves of extremal distance between two points OR
- (2) Curves that parallel transport their tangent vector $t^a = dx^a/du = \dot{x}^a$.

They are important because free particles and massless particles move along geodesics in spacetime.

Tangent vector:

$$|\mathbf{t}|^2 = g_{ab} \frac{dx^a}{du} \frac{dx^b}{du} = \begin{cases} g(\mathbf{t}, \mathbf{t}) > 0, & \text{timelike} \\ g(\mathbf{t}, \mathbf{t}) < 0, & \text{spacelike} \\ g(\mathbf{t}, \mathbf{t}) = 0, & \text{null} \end{cases}$$

$|\mathbf{t}| = \left| \frac{ds}{du} \right|$ where s is the proper path length s along the curve.

(1) Stationary property of non-null geodesics: affinely-parameterised geodesic equation:

Use the Lagrangian approach to find extremal distance.

$$L = g_{ab} \dot{x}^a \dot{x}^b \quad \Rightarrow \quad \frac{\partial L}{\partial x^c} = \frac{d}{du} \left(\frac{\partial L}{\partial \dot{x}^c} \right)$$

[Exercise] Then, just do the math,

$$\partial_c g_{ab} \dot{x}^a \dot{x}^b = \frac{d}{du} (g_{ca} \dot{x}^a + g_{ac} \dot{x}^a) = \frac{d}{du} (2g_{ac} \dot{x}^a) = 2g_{ac} \ddot{x}^a + 2\dot{x}^a \partial_d g_{ac} \dot{x}^d$$

Note that: $\frac{dg_{ac}}{du} = \partial_d g_{ac} \dot{x}^d$ when doing chain rule in the 3rd equal sign. Rearrange and multiple both sides by g^{dc} to find the geodesic equation.

Geodesic equation:

$$\ddot{x}^d + \Gamma_{ab}^d \dot{x}^a \dot{x}^b = 0 = \frac{Dt^a}{Du}$$

(2) We identify that the geodesic equation is also the parallel transport of the tangent vector.

There is a preferred class of parameters called affine parameters that make this expression true. Affine parameters are linearly related to length along the line, $u = as + b$.

This procedure is useful to find **metric connection** quickly.

[Exercise] Alternative form obtained by first lowering the index on the equation of parallel transport for the tangent vector i.e. start from $g_{ab} \frac{Dt^a}{du} = 0$:

$$\frac{dt_a}{du} = \frac{1}{2} \partial_a g_{bc} t^b t^a$$

Conserved quantities along geodesic:

1. $L = \text{constant}$ since L does not depend explicitly on u.
2. $\partial_a g_{bc} = 0 \Rightarrow t_a = \text{constant}$. If the metric does not depend on a coordinate x^a , then the a^{th} component of the tangent (dual) vector is conserved along an affinely-parameterised geodesic. Also follows from Euler-Lagrange if L does not depend explicitly on x^c , getting conservation of conjugate momenta, $\pi_a = \frac{\partial L}{\partial \dot{x}^a}$.

Summary

- Metric function: from $ds = \mathbf{e}_a(x)dx^a$, we have $ds^2 = \mathbf{e}_a \cdot \mathbf{e}_b dx^a dx^b$,

$$g_{ab}(x) = \mathbf{e}_a(x) \cdot \mathbf{e}_b(x) \quad \text{and similarly} \quad g^{ab}(x) = \mathbf{e}^a(x) \cdot \mathbf{e}^b(x)$$

$$g'_{ab} = \frac{\partial x^c}{\partial x'^a} \frac{\partial x^d}{\partial x'^b} g_{cd} \quad \Rightarrow \quad g' = J^{-1} J^{-1} g \quad \text{and} \quad g'^{-1} = g^{-1} J J$$

- Coordinate transformation:

$$\text{upstairs: } v'^a = \frac{\partial x'^a}{\partial x^b} v^b \quad \text{and} \quad \text{downstairs: } v'_a = \frac{\partial x^b}{\partial x'^a} v_b$$

- Choosing orthonormal basis vectors: $\hat{\mathbf{e}}_a \cdot \hat{\mathbf{e}}_b \eta_{ab}$
- Raising and lowering indices: from inner product: $\mathbf{v} \cdot \mathbf{w} = g_{ab} v^a w^b$, we have

$$g_{ab} v^a w^b = v_a w^a = g^{ab} v_a w_b = v^a w_a$$

where:

$$g^{ab} g_{bc} = \delta_c^a$$

Metric and dual metric are inverses of each other. i.e. $h^{-1}h = \mathbb{I}$.

- Special tensor:

$$g'^a_b = \delta^a_b = g^a_b$$

The only rank-2 tensor whose components are the same in all coordinate systems.

- Covariant derivative:

$$\nabla_a v^b = \frac{\partial v^b}{\partial x^a} + \Gamma^b_{ac} v^c$$

- Affine connection:

$$\Gamma^a_{bc} = \mathbf{e}^a \cdot \partial_c \mathbf{e}_b \quad \text{or} \quad \Gamma'^a_{bc} = \mathbf{e}'^a \cdot \partial_c \mathbf{e}'_b$$

Computing metric coefficients quickly using the Lagrangian method:

1. Write down $L = g_{ab} \dot{x}^a \dot{x}^b$.
2. Use Euler-Lagrange equation.
3. Compare with geodesic equation: $\ddot{x}^d + \Gamma^d_{ab} \dot{x}^a \dot{x}^b = 0$.
4. Write down connection coefficients quickly.

5 Theory of Relativistic Mechanics

$$\text{Event: } x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z)$$

Minkowski spacetime is a 4D pseudo-Euclidean manifold. We can adopt a global inertial coordinates:

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu \quad \text{where} \quad \eta_{\mu\nu} = \text{diag}(+1, -1, -1, -1) = \eta^{\mu\nu}$$

where $\eta_{\mu\nu}$ is the **Minkowski metric**. In Cartesian coordinates, $\Gamma_{\nu\sigma}^\mu = 0$ so $\Delta = \partial$ and $\frac{D}{Du} = \frac{d}{du}$. Notation for dot product: $g(u, v) = g_{ab} u^a v^b$.

Lorentz Transformation

Definition of Lorentz transformation:

$$\eta_{\mu\nu} = \frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial x'^\sigma}{\partial x^\nu} \eta_{\rho\sigma} \quad \iff \quad \eta_{\mu\nu} = \Lambda^\rho{}_\mu \Lambda^\sigma{}_\nu \eta_{\rho\sigma} \quad (1)$$

where

$$\Lambda^\mu{}_\nu = \frac{\partial x'^\mu}{\partial x^\nu} \quad \text{and} \quad \Lambda_\mu{}^\nu = \frac{\partial x^\nu}{\partial x'^\mu}$$

It can be shown that Lorentz transformation must be linear:

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu \quad (\text{Poincare transformation})$$

for suitable constant Λ_ν^μ and constant a^μ corresponds to changing the space-time origin, dropping this term gives us **homogeneous Lorentz transformation**. The standard Lorentz boost (along x) is:

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad ; \quad \Lambda_\mu{}^\nu = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Using eqn (1), the inverse of the transformation matrix is $\Lambda_\mu{}^\nu = (\Lambda^{-1})^\mu{}_\nu = \eta_{\mu\rho} \eta^{\nu\sigma} \Lambda^\rho{}_\sigma$.

Let's only consider **proper Lorentz transformations** which is a subgroup of the full Lorentz transformations that only include transformations between inertial frames with the same spatial handedness and exclude time reversal.

$$\det(\Lambda^\mu{}_\nu) = 1 \quad \text{and} \quad \Lambda^0{}_0 \geq 1$$

Transformations

Orthonormal basis vectors: $\mathbf{g}(\mathbf{e}_\mu, \mathbf{e}_\nu) = \eta_{\mu\nu}$ (scalar product of basis vectors picks out the appropriate component of the metric in those coordinates).

Basis vectors transform with inverse transformation matrix: $\mathbf{e}'_\mu = \Lambda_\mu{}^\nu \mathbf{e}_\nu$.

Components of a vector transform with transformation matrix: $v'^\mu = \Lambda^\mu{}_\nu v^\nu$.

Components of a dual vector transform with the inverse transformation: $\mathbf{X}'_\mu = \Lambda_\mu{}^\nu \mathbf{X}_\nu$.

Name	Symbol	Definition	Components	Invariance of norm squared
4-displacement	x^μ	x^μ	(ct, \mathbf{r})	$ds^2 = c^2\tau^2$
4-velocity	u^μ	$u^\mu = \frac{Dx^\mu}{D\tau}$	$\gamma(c, \mathbf{u})$	c^2
4-momentum	p^μ	$p^\mu = mu^\mu$	$(E/c, \mathbf{p})$ OR $\gamma(mc, m\mathbf{u})$	m^2c^2
4-acceleration	a^μ	$a^\mu = \frac{Du^\mu}{D\tau}$	$\gamma(\dot{\gamma}c, \dot{\gamma}\mathbf{u} + \gamma\mathbf{a})$	$-a_{IRF}^2$
4-force	f^μ	$f^\mu = \frac{Dp^\mu}{D\tau} = ma^\mu$	$\gamma(\mathbf{f} \cdot \mathbf{u}/c, \mathbf{f})$	
4-current	j^μ	$j^\mu = \rho u^\mu$	$(c\rho, \mathbf{j})$	$c^2\rho^2$
4-potential	A^μ	A^μ	$(\phi/c, \mathbf{A})$	

[**Reminder**] Important relation to know: $\frac{dt}{d\tau} = \gamma$, $|A|^2 = \eta_{\mu\nu}A^\mu A^\nu$ and $A'^\mu = \Lambda^\mu{}_\nu A^\nu$.

4-velocity

It is convenient to parameterise the worldline of a massive particle with its proper time, τ . It is the affine parameter for the worldline.

4-acceleration

For free particles, \mathbf{u} and γ_u is constant, so

$$a^\mu = \frac{Du^\mu}{D\tau} = 0$$

Since u^μ is the tangent vector to the worldline in an affine parameterisation, we see that free massive particles move along timelike geodesics in Minkowski space.

$\mathbf{g}(\mathbf{a}, \mathbf{u}) = 0 = \frac{1}{2} \frac{d}{d\tau} (\eta_{\mu\nu} u^\mu u^\nu)$. **Acceleration 4-vector is always orthogonal to 4-velocity.**

In the instantaneous rest frame of the particle, $\mathbf{u} = 0$, so $a^\mu = (0, \mathbf{a}_{IRF})$, the norm squared invariant quantity is $|\mathbf{a}|^2 = -|\mathbf{a}_{IRF}|^2$. This shows that the 4-acceleration is a spacelike vector.

4-momentum

Forming the invariant $|\mathbf{p}|^2$ gives us the energy-momentum invariant:

$$E^2 - |\mathbf{p}|^2 c^2 = m^2 c^4$$

In scattering problems, note that **the total 4-momentum is constant**¹ and the fact that **squared magnitude of the total 4-momentum is Lorentz invariant**. $\frac{Dp^\mu}{D\tau} = 0$.

For photons $[p^\mu = (E/c, \vec{p})$ where $E = |\vec{p}|c$], it moves on null geodesics in Minkowski space i.e. $\mathbf{g}(\mathbf{p}, \mathbf{p}) = 0$. $d\tau = 0$, so we cannot use proper time τ as a parameter. We can always adopt a (dimensional) parameterisation such that $p^\mu = \frac{dx^\mu}{d\lambda}$.

Examples of relativistic mechanics problem: Doppler effect, Compton scattering.

¹This combines both conservation of 3-momentum and energy into a Lorentz invariant (i.e., 4-vector) law

6 Electromagnetism

$$(M1) \nabla \cdot \mathbf{E} = \rho/\epsilon_0 \quad ; \quad (M2) \nabla \cdot \mathbf{B} = 0 \quad ; \quad (M3) \nabla \times \mathbf{E} + \dot{\mathbf{B}} = 0 \quad ; \quad (M4) \nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \dot{\mathbf{E}}$$

$$\mathbf{f} = q(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \quad ; \quad \nabla \cdot \mathbf{J} = -\dot{\rho}$$

Lorentz force law in 4D & Finding the components of the Maxwell stress tensor

To generalise the Lorentz force law to 4D-dimensional spacetime, we want to construct an object that is linear in velocity, electric field and magnetic field. The most general thing I can write down is:

$$f_\mu = qF_{\mu\nu}u^\nu \quad \text{and} \quad f^\mu = qF^\mu_\nu u^\nu$$

where $F_{\mu\nu}$ is the **Maxwell stress tensor**. Now, let's investigate the properties of this new "object":

1. Since $f_\mu u^\mu = 0 = qF_{\mu\nu}u^\nu u^\mu$. The Maxwell stress tensor is antisymmetric, $F_{\mu\nu} = -F_{\nu\mu}$. The **diagonal terms are zero**. We are now left with finding $F_{01}, F_{02}, F_{03}, F_{12}, F_{13}, F_{23}$.
2. To find the remaining components, we compare with the 3D Lorentz force law:

$$f_\mu = \gamma \left(\frac{\mathbf{f} \cdot \mathbf{u}}{c}, \mathbf{f} \right) = q\gamma \left(\frac{\mathbf{E} \cdot \mathbf{u}}{c}, -(\mathbf{E} + \mathbf{u} \times \mathbf{B}) \right) = qF_{\mu\nu}u^\nu$$

Compare 0 and spatial terms:

- Temporal: $F_{0i} = \frac{\vec{E}^i}{c}$ and $F_{00} = 0$.
- Spatial: $f_i = qF_{i0}u^0 + q\sum_j F_{ij}u^j = -q\gamma E^i - q\gamma(\vec{u} \times \vec{B})^i$.

$$\Rightarrow F_{ij}u^j = -[\gamma\vec{u} \times \vec{B}]^i = \begin{pmatrix} -(u^2\vec{B}^3 - \vec{B}^2u^3) \\ +(u^1\vec{B}^3 - \vec{B}^1u^3) \\ -(u^1\vec{B}^2 - \vec{B}^1u^2) \end{pmatrix}^i. \text{ So, } F_{12} = -\vec{B}^3, F_{13} = \vec{B}^2 \text{ and } F_{23} = -\vec{B}^1.$$

3. Also, another form of the Maxwell tensor in terms of 4-potential: $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$

We can now write the form of the Maxwell tensor:

$$F_{\mu\nu} = \begin{pmatrix} 0 & \vec{E}^1/c & \vec{E}^2/c & \vec{E}^3/c \\ \cdot & 0 & -\vec{B}^3 & \vec{B}^2 \\ \cdot & \cdot & 0 & -\vec{B}^1 \\ \cdot & \cdot & \cdot & 0 \end{pmatrix} \quad ; \quad F^{\mu\nu} = \begin{pmatrix} 0 & -\vec{E}^1/c & -\vec{E}^2/c & -\vec{E}^3/c \\ \cdot & 0 & -\vec{B}^3 & \vec{B}^2 \\ \cdot & \cdot & 0 & -\vec{B}^1 \\ \cdot & \cdot & \cdot & 0 \end{pmatrix}$$

The bottom half is just whatever is needed to make $F_{\mu\nu}$ antisymmetric. We can also raise both indices with $\eta^{\mu\nu}$, $F^{\sigma\gamma} = \eta^{\sigma\mu}\eta^{\gamma\nu}F_{\mu\nu}$. The Maxwell tensor transforms as²:

$$F'^{\mu\nu} = \Lambda^\mu_\rho \Lambda^\nu_\sigma F^{\rho\sigma} \quad \text{or} \quad F' = \Lambda F \Lambda$$

[ES3 Qn 5] With this, we can relate the components of the electric and magnetic fields in frame S and S': $E'_\parallel = E_\parallel$, $B'_\parallel = B_\parallel$, $\mathbf{E}'_\perp = \gamma(\mathbf{E}_\perp + \vec{v} \times \mathbf{B}_\perp)$, $\mathbf{B}'_\perp = \gamma(\mathbf{B}_\perp + \vec{v} \times \mathbf{E}_\perp/c^2)$.

The Lorentz-invariant is: $F^2 = \frac{1}{2}F^{\mu\nu}F_{\mu\nu} = c^2|\mathbf{B}|^2 - |\mathbf{E}|^2$.

²https://hepweb.ucsd.edu/ph110b/110b_notes/node69.html

Relativistic field equations and getting the 4 Maxwell's equations

Relativistic field equations:

We expect the spacetime-derivative of field-strength to be linearly related to current 4-vector:

$$\nabla_\mu F^{\mu\nu} = \mu_0 j^\nu \quad (1)$$

By antisymmetry $\nabla_{[\sigma} F_{\nu\rho]} = 0$, equivalently:

$$\nabla_\sigma F_{\mu\nu} + \nabla_\nu F_{\sigma\mu} + \nabla_\mu F_{\nu\sigma} = 0 \quad (2)$$

Consider in Cartesian coordinates,

$$(1) \quad \nabla_\mu F^{\mu\nu} = k j^\nu = \partial_\mu F^{\mu\nu}$$

1. Taking divergence and using the antisymmetry of $F^{\mu\nu}$, we get the **continuity equation**: $\partial_\nu j^\nu = 0$.
2. Consider the 0-component, $\frac{F^{00}}{\partial(ct)} + \frac{\partial F^{i0}}{\partial x^i} = k j^0 \Rightarrow \frac{1}{c} \nabla \cdot \mathbf{E} = k c \rho$. Taking $k = \mu_0$, we recover (M1).
3. Consider spatial component, $\frac{F^{0i}}{\partial(ct)} + \frac{\partial F^{ji}}{\partial x^j} = \mu_0 j^i \Rightarrow -\mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t} + \nabla \times \vec{B} = \mu_0 \vec{J}$, we recover (M4).
4. Consider $\nabla_\nu \nabla_\mu F^{\mu\nu} = -\nabla_\nu \nabla_\mu F^{\mu\nu} = \mu_0 \nabla_\nu j^\nu = 0$. We recover charge conservation equation.

Also,

$$(2) \quad \partial_\sigma F_{\mu\nu} + \partial_\nu F_{\sigma\mu} + \partial_\mu F_{\nu\sigma} = 0$$

There are $\binom{4}{3} = 4$ independent components. $(\mu, \nu, \rho) = (0, 1, 2), (0, 1, 3), (0, 2, 3), (1, 2, 3)$.

For $(0, 1, 2)$, we have $\nabla \times \mathbf{E} = -\dot{\mathbf{B}}$, we recover (M3).

For $(1, 2, 3)$, we have $\nabla \cdot \mathbf{B} = 0$, we recover (M2).

4-vector potential, A_μ

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

In this definition, there is a residual gauge freedom to add the gradient of a scalar, $\partial_\mu \psi$ to A_μ without altering $F_{\mu\nu}$. We can hence, choose our gauge. A common one is the Lorenz gauge: $\partial_\mu A^\mu = 0$.