§1. SPECIAL RELATIVITY

§1.0 Introduction: Comparison of Newtonian and SR Specifications

- Special Relativity (SR) preceded General Relativity (GR)
  No gravity, no curvature of space-time (space-time is flat)
  Covers all other (non-quantum) physics

- I assume SR is already familiar to you. Thus I shall just present SR as it is, without spending due on its motivation. (Such "motivation" could be confusing, since often such material mixes Newtonian and relativistic ideas. The structure of SR is actually very simple. The only difficult thing about SR is forgetting Newtonian ideas about space and time.)
  The purpose is to prepare for GR.
  - introduce 4-vectors, tensors, etc. (machinery to be used in GR also)

- Geometry and coordinates:
  - geometry: In SR, space-time is flat
  - coordinates: In SR, we shall use inertial (Cartesian) coordinates only
Comparison of Newtonian and SR spacetime

Special Relativity (SR): a theory of structure of spacetime
Newtonian Mechanics (NM): a different theory of

- Spacetime: 4-dimensional set of points: events (or spacetime points)
  label with 4 coordinates: \((t, x, y, z)\)
  many different coordinate systems ("frames")

Worldline: path of particle through spacetime
- Particles can move back and forth in space, but not in time
  (both in NM and SR)

*) In SR, "ord. system" and "frame" are synonymous (as long as we stick to
  inertial ord. systems), in GR they are not ("frame" is just local).
Newtonian spacetime

Structure: divided into uniquely defined
Time grids "Space at time t"

Absolute time: A and B are
simultaneous (at the same t)
No absolute x, y, or z (have A and B
the same y ?)
No absolute space (are A and C at
the same place ?)

3 kinds of spacetime directions:

- worldlines move forward in time
  (direction is future timelike),
  but no other constraint
  (no limit on velocity) \( V < \infty \)

SR spacetime (Minkowski space)

Structure: uniquely defined light cone at
every event = set of special directions
"Speed of light c"

No absolute time or space (in one
coordinate system \( t, x, y, z \), A and B have the same \( t \);
in another \( t', x', y', z' \), A and B have
different \( t' \))
No preferred time sharing \( \Rightarrow \) notion of
spacetime more fundamental

5 kinds of spacetime directions:

- worldlines of massive particles lie inside
  light cone (direction is future timelike) \( V < c \)
- worldlines of massless particles lie along
  light cone (direction is future lightlike): \( V = c \)
- for spacelike direction, no distinction between
  past and future
In the limit $c \to \infty$, $SR \to NM$

as the light cone flattens to become a time slice.

The "speed of light" $c$ has a similar status in SR as an infinite speed has in NM. Thus it is easy to understand and it cannot be exceeded. Or, that travelling faster than speed of light would also make time travel possible in SR. In NM, travelling faster than at infinite speed (covering some distance in zero time), would be covering the distance in negative time.
Geometry refers to quantities that are independent of any system.

Time: Since time is absolute, no non-trivial transformations of the t coordinate.

Space: This leaves us x, y, z. Since we stick to Cartesian 3D space, the only non-trivial transformations are rotations. Example: rotation in xy-plane.

\[ \Delta s = \text{distance between } A \text{ and } B, \quad \text{invariant under change of } \text{nd's} \]

\[ (\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 = (\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2 \]

Because of such invariance, it is more useful to think of 3D space, instead of a set of 1D or 2D spaces.

In Newtonian physics, there is no unidirectional motion of rotating space and time into each other, so the spacetime is just a set of \( t = \text{const} \) 3D spaces. The geometry of Newtonian space is called Euclidean.

*) There is of course the trivial transformation \( t' = t + C \), translation or different choice of origin.
§11. Spacetime of Special Relativity (Minkowski space)

- Need some thought how to form coordinate systems

- Set up a coordinate system
  - Cartesian \( x, y, z \)
  - think of as rigid rods attached to each other at right angles
  - move freely: unaccelerated

\( t \): set of clocks attached (not moving wrt \( x, y, z \))

- Synchronization of clocks by vacuum light rays
  - (local comparisons)

require

\[ t_2 = \frac{1}{2}(t_1 + t_3) \]

(and \( t_4 - t_2 = t_3 - t_1 \))

\( (t, x, y, z) \): inertial coordinates, inertial frame \( K \)

- Construct many such inertial frames \( (t', x', y', z') \) \( K' \)
  - (may move wrt each other)

*) How do you know whether the frame is unaccelerated: Newton's 1st law holds

\[ F = 0 \quad \Rightarrow \quad \frac{d^2x}{dt^2} = \frac{d^2y}{dt^2} = \frac{d^2z}{dt^2} \]
The fundamental geometrical property that defines the structure of Minkowski space is its invariance.

\[
\Delta s^2 = \text{spacetime interval between two events A and B} \\
= - (c\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 \\
= - (c\Delta t')^2 + (\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2
\]

\(c\) = fixed convention factor between units of time and space (distance)

\(\Rightarrow c\) has dimension of velocity.

\[
(\text{ED} \Rightarrow \text{electromagnetic waves propagate in vacuum at this velocity})
\]

\(\Rightarrow c\) usually called the "Speed of Light".

\[
\exists \text{ a constant } c, \text{ such that } \Delta s^2 \text{ is invariant under changes between inertial coordinates.}
\]

\(\therefore\) Useful to think of 4-d spacetime (Minkowski space), instead of a set of 3-d spaces.

\(K \rightarrow K'\) rotate space and time into each other.

\[
\begin{align*}
\Delta t & = \Delta t' \\
(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 & \neq (\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2 \text{ in general}
\end{align*}
\]

\(\ast\) We define the spacetime interval here as "\(\Delta s^2\)", which is to be thought of as just a symbol, not necessarily a square of anything.

It can be positive, negative, or zero.
The invariance of $\Delta s^2$ gives rise to the light cone:

Given a reference event $P$, the set of events whose $\Delta s^2$ from $P$ is $0$, form the light cone of $P$.

- $\Delta s^2 < 0$: timelike separated from $P$ \((A, B)\)
- $\Delta s^2 = 0$: light cone at $P$, null or light-like separated \((C, D)\)
- $\Delta s^2 > 0$: spacelike separated from $P$ \((E, F)\)

$A$ is in the future of $P$, $B$ in the past of $P$.

Space-time diagrams with 1 and 2 space 1ds:

Set origin \((t = x = y = z = 0)\) at $P$.

- Light cone $-t^2 + x^2 = 0$.
- $x = \pm t$ corresponds to motion at the speed of light in the $x$ direction.

- Light cone $-t^2 + x^2 + y^2 = 0$.

---

1.8
Notation

- Coordinates \( x^\mu \): \( x^0 = ct \), \( x^1 = x \), \( x^2 = y \), \( x^3 = z \)
- Choose units where \( c = 1 \) \( \Rightarrow \) \( 1 \text{ s} = 299\,792\,458 \text{ m} \)
  \( 1 \text{ year} = 1 \text{ light year} \)

The metric \( \eta_{\mu\nu} \): a \( 4 \times 4 \) matrix

\[
\begin{bmatrix}
-1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
\end{bmatrix}
\]

\[\Delta s^2 = \sum_{\mu=0}^{3} \sum_{\nu=0}^{3} \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu = \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu \]

Einstein summation convention!

(an index, which occurs twice, once downstairs, once upstairs, is summed over)

The indices and sums have the following ranges:

Greek: \( \alpha, \beta, \gamma, \ldots, \mu, \nu, \sigma, \ldots \) = 0, 1, 2, 3  

Spacetime

Latin: \( i, j, k, \ldots \) = 1, 2, 3  

Space

Those indices that are not summed over, must be in the same position (up or down) on both sides of the equation, and appear exactly once in each term. The same index must not appear more than twice in any term. (If an exception to these rules must be made, it needs to be explained.) Note, that you are free to choose the name of a dummy index (and often must do this, to keep w these rules).

*) the meaning of this up/down position will be explained a little later

**) another convention, common in particle physics, has the opposite sign for \( \eta_{\mu\nu} \)
For timelike intervals only, define
\[ \Delta \tau = \sqrt{-\Delta s^2} = \sqrt{-\eta_{\mu\nu} \Delta x^\mu \Delta x^\nu}, \]
the proper time between two events (e.g., \( P \) and \( A \))
= time measured by an observer that moves from \( P \) to \( A \) on a straight line.

Show this: In this observer's frame, \( x^i(A) = x^i(P) \) \( \Rightarrow \) \( \Delta x^i = 0 \)
\[ \Rightarrow \Delta \tau = \sqrt{-\eta_{\mu\nu} \Delta t \Delta t'} = \Delta t' \]
\( \Delta \tau \) is invariant under \( k \rightarrow k' \), but \( \Delta t \) is not \( \Rightarrow \) in general, \( \Delta \tau \neq \Delta t \)

(1.10a)

Proper time along timelike paths: Consider first a path (worldline) made
out of straight timelike pieces

\[ \Delta \tau = \sum_a \Delta \tau_{(a)} = \sum_a \sqrt{-\eta_{\mu\nu} \Delta x_{(a)}^\mu \Delta x_{(a)}^\nu}, \]
A straight line maximizes \( \Delta \tau \) from \( P \) to \( A \)
(Exercise: e.g., the twin "paradox")

*) Straight in spacetime, i.e., also constant velocity (an inertial observer)
The set of events, which are the same proper time away from the origin \( O \). (We can choose any spacelike event to be the origin.)

For any such point, if inertial frame, where the point has the same (space) location as \( O \), and the (coordinate) time difference from \( O \) is \( \Delta t = \pm \Delta z \) then

\[
\Delta s = \sqrt{\Delta \eta^2} = \sqrt{\eta_{\mu \nu} \Delta x^\mu \Delta x^\nu}
\]

The set of events, which are the same proper distance \( \Delta s \) away from the origin \( O \).

For any such event, if inertial frame where the event has the same time (coordinate) as \( O \), and the space distance \( \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2} \) is \( \Delta s \).
The Line Element

\[ ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu \]

(line element)

(infinitesimal spacetime interval)

parameterized curve \( x^\mu(\lambda) \):

\[ dx^\mu = \frac{dx^\mu}{d\lambda} \]

(\( \lambda \) not necessarily identified w the time and \( \lambda \))

\[ \Delta \tau = \int_{\lambda_1}^{\lambda_2} \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda \]

= time measured by the observer, whose worldline this is \( * \)

For \( ds^2 < 0 \) can define \( d\tau = \sqrt{-\eta_{\mu\nu} dx^\mu dx^\nu} \)

\( \Rightarrow \) proper time along timelike curve (world line)

For \( ds^2 > 0 \) can define \( ds = \sqrt{\eta_{\mu\nu} dx^\mu dx^\nu} \)

\( \Rightarrow \) proper length along spacelike curve

\[ \Delta s = \int_{\lambda_1}^{\lambda_2} \sqrt{\eta_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda \]

For a timelike (null) curve, \( ds^2 = 0 \) \( \Rightarrow \Delta \tau = \Delta s = 0 \)

\( *\) This holds for arbitrary (timelike, i.e., at less than speed of light) motion, which may contain all kinds of acceleration. We shall use only inertial rel’s in SR, but we are free to consider all kinds of observers, both inertial and non-inertial.

(The use of non-inertial rel’s in SR is also possible; but then we would need to introduce half the math at GR, who only real taught. So we gave those to GR.)
1.3 Lorentz Transformations

- How to transform from one inertial frame to another? \( K \rightarrow K' \)

Suppose we know the event's \( x^m \) at event A in K.

What are the event's \( x'^m \) in \( K' \)?

(note we put the prime on the index)

The transformation must leave \( \Delta s^2 = g_{\mu\nu} \Delta x^\mu \Delta x^\nu \) invariant

- Translations (change at origin) \( x^m \rightarrow x'^m = x^m + a^m \) \hspace{1cm} (1)
  \[ \Rightarrow \Delta x'^m = \Delta x^m \Rightarrow \Delta s^2 \text{ invariant} \]
  \[ \Rightarrow \eta \text{ constants} \]

- Rotations (K and \( K' \) have different orientation of the \( x,y,z \) axes) \hspace{1cm} \{ \text{linear} \}

- Boosts (K and \( K' \) — velocity —- ) \hspace{1cm} \{ \text{transformations} \}

  \[ x'^m = \Lambda^{\mu\nu} x^\nu \] \hspace{1cm} (2)

  where \( \Lambda \) is a \( 4\times4 \) matrix.

\[ \Rightarrow \Delta x'^m = \Lambda^{\mu\nu} (\Delta x^\nu) \]

- In matrix notation \( \Delta x' = \Lambda (\Delta x) \) \hspace{1cm} (3)

  The invariance

  \[ \Delta s^2 = (\Delta x)^T \eta (\Delta x) = (\Delta x')^T \eta (\Delta x') \]

  \[ \Rightarrow \eta = \Lambda^T \eta \Lambda \] \hspace{1cm} (4)

  (This equivalence requires that \( \eta \) and \( \Lambda^T \eta \Lambda \) are symmetric; they are)

  In index notation (4) reads

  \[ \eta_{\mu\nu} = \Lambda^\mu_\alpha \eta_{\alpha\beta} \Lambda^\beta_\nu \]

  \[ = \Lambda^\mu_\alpha \eta_{\alpha\beta} \Lambda^\beta_\nu \]

  (in index notation the order does not matter, the corresponding information is carried by which indices are paired, to be summed over)

- Matrices that satisfy (4) are called Lorentz transformations.

They form a group under matrix multiplication, the Lorentz group.
Analogous to the rotation group in 3-d space:

Orthogonal matrices \( R \) satisfy \( R^T R = I \) (6)

Orthogonal 3x3 matrices form the orthogonal group \( O(3) \).

It includes rotations and parity transformations → \[
\begin{bmatrix}
\pm 1 & \pm 1 & \pm 1 \\
\end{bmatrix}
\]

Changing 1 or 3 signs is a parity transformation,
changing 2 signs is actually a 180° rotation:

\( \text{det} R = 1 \) excludes parity transforms

→ Special orthogonal group = rotation group = \( SO(3) \)

To see the analogy to Lorentz group, write (6) as

\[
I = R^T I R
\]

and compare to (4): the difference between \( O(3) \) and Lorentz group is

\[
I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \rightarrow \quad \gamma = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\]

The Lorentz group is denoted \( O(3,1) \).

It includes rotations, boosts, time reversals \( (t^2 = -t) \) and parity transformations.

\[ \text{det} \Lambda = 1 \quad \rightarrow \quad "\text{proper Lorentz group}" \quad SO(3,1) \]

Leaves a combination of time reversal and parity transf

One can easily show that always \( |\Lambda_0^0| = 1 \). Requires

\[ \text{det} \Lambda = 1 \quad \text{and} \quad \Lambda_0^0 \geq 1 \quad (8) \]

leaves the "proper orthochronous" or "restricted" Lorentz group \( SO(3,1)^+ \)

\[
I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\]

is the metric of ordinary 3-d flat space: \( ds^2 = (dx)^T I (dx) \)

A metric with all eigenvalues positive, has Euclidean signature

... one sign different from the others, has Lorentzian signature
Examples of Lorentz transformations

- **Rotations**, e.g.
  \[
  \Lambda^\mu_\nu = \begin{bmatrix}
  1 & \cos \theta & \sin \theta \\
  -\sin \theta & \cos \theta & 0 \\
  0 & 0 & 1
  \end{bmatrix}
  \]
  in \ xy\text{-plane} \hspace{1cm} \text{rotation angle } \theta, \hspace{1cm} \text{periodic (}2\pi\text{) variable (9)}

- **Boosts**, e.g.
  \[
  \Lambda^\mu_\nu = \begin{bmatrix}
  \cosh \gamma & -\sinh \gamma & 0 & 0 \\
  -\sinh \gamma & \cosh \gamma & 0 & 0 \\
  0 & 0 & 1 & 0 \\
  0 & 0 & 0 & 1
  \end{bmatrix}
  \]
  "rotation" in \ ex\text{-plane} \hspace{1cm} \text{boost in } x \text{ direction (10)}

  the boost parameter \( \gamma \in (-\infty, \infty) \), not periodic

- **Exercise**: Show that these \( \Lambda^\mu_\nu \) satisfy Eq. (4).

A general SO(3,1) matrix has 6 parameters (3 rotation, 3 boost)

- Lorentz transformations do not commute \( \Rightarrow \) non-abelian group

- Include translations \( q^\mu \Rightarrow 10\text{-parameter non-abelian group, the Poincaré group}

- Boosts give a transformation to an inertial frame \( w \) with a different velocity:

  \[
  \begin{aligned}
  t' &= t \cosh \gamma - x \sinh \gamma \\
  x' &= -t \sinh \gamma + x \cosh \gamma
  \end{aligned}
  \]

  \( t'\text{-axis: } x' = 0 \Rightarrow x \cosh \gamma = t \sinh \gamma \) \hspace{1cm} \Rightarrow \hspace{1cm} x = t \cdot \tanh \gamma

\[
\therefore \quad \begin{cases}
  t' = t \cosh \gamma - x \sinh \gamma \\
  x' = -t \sinh \gamma + x \cosh \gamma
\end{cases}
\]

\[
\therefore \quad \begin{cases}
  \text{The } x' = 0 \text{ point is moving in } K \text{ with velocity } V
\end{cases}
\]

\[
V = \frac{x}{t} = \frac{v}{\cosh \gamma} \Rightarrow \gamma = \tanh^{-1} \frac{V}{c}
\]

\( \therefore \quad \gamma = \tanh^{-1} \frac{V}{c} \) \hspace{1cm} \text{rapidly}

\[\text{velocity (11)}\]

\[
\cosh^2 \gamma - \sinh^2 \gamma = 1 \Rightarrow 1 - \tanh^2 \gamma = \frac{1}{\cosh^2 \gamma}
\]

\[
\cosh \gamma = \frac{1}{\sqrt{1 - \tanh^2 \gamma}} = \frac{1}{\sqrt{1 - V^2}} = \gamma
\]

\[
\sinh \gamma = \cosh \gamma \cdot \tanh \gamma = \gamma V
\]

\( (12) \)

\[
(14) \quad \text{The conventional way to write the Lorentz transformation}
\]

\[
K' \text{ is moving with velocity } V \text{ in } K.
\]

\[
(13)
\]
Boostra and Spacetime Diagrams

We already saw how the $t'$ axis looks on a spacetime diagram drawn in K.

$x^2$-axis: $t' = 0 \Rightarrow t \cosh \psi = x \sinh \psi \Rightarrow t = x \cdot \tanh \psi$

The $x'$ and $t'$ axes are orthogonal in the Lorentzian sense (as we shall see).

Consider an object moving at speed c at light $x = \pm ct = \pm t$ in K

$\Rightarrow t' = \cosh \psi - x \sinh \psi = t (\cosh \psi \mp \sinh \psi)$

$x' = -t \sinh \psi + x \cosh \psi = -t (\sinh \psi \mp \cosh \psi) = \pm t'$

$\Rightarrow$ it also moves at speed at light in $K'$.

Inverse Lorentz Transformation (belong to p. 1-17)

$K \rightarrow K': x^\mu = \Lambda^\mu_{\nu} x^\nu$

$K' \rightarrow K: x^\mu = \Lambda^\mu_{\nu} x'^\nu$

Since $x^\mu = \Lambda^\mu_{\nu} x'^\nu = \Lambda^\mu_{\nu} \Lambda^\nu_{\sigma} x^\sigma = x^\mu \forall \rho$

must have $\delta^\mu_{\sigma}$ (Kronecker delta $= 1$ when $\mu = \sigma$)

$= 0$ when $\mu \neq \sigma$

$\therefore \Lambda^\mu_{\nu}$ and $\Lambda^\mu_{\nu}$ are inverse matrices.

$\Lambda^\mu_{\nu}$
- Draw the same thing in $k'$:

\[ \text{K is moving with velocity } -v \text{ in } k' \]

Both diagrams are equally good representations of Minkowski space.

Although we say that the Minkowski space is flat (= not curved), its geometry is not Euclidean, and thus we cannot faithfully reproduce its geometry at the $tx$-plane on an Euclidean 2-d sheet of paper.

- How do the positions of spacetime points (events) on a spacetime diagram change if we vary the boost parameter $\gamma$ continuously?

Some figures for rotation in the $xy$-plane.
The Poincaré transformation is
\[ x'^\mu = \Lambda^\mu_\nu x^\nu + a^\mu \]

It is a 1st transformation. The Jacobian matrix \( \frac{\partial x'^\mu}{\partial x^\nu} \) at this 1st transformation is
\[ \frac{\partial x'^\mu}{\partial x^\nu} = \Lambda^\mu_\nu \]

**Inverse Lorentz Transformation**

\[ K \rightarrow K': \quad x'^\mu = \Lambda^\mu_\nu x^\nu \quad \text{ (note: } \Lambda^\mu_\nu \text{ and } \Lambda^\nu_\mu \text{ are}
\]
\[ K' \rightarrow K: \quad x^\mu = \Lambda^\nu_\mu x'^\nu \quad \text{ two different matrices!} \]

Since \( x^\mu = \Lambda^\mu_\nu x'^\nu = \frac{\Lambda^\mu_\nu}{\delta^\nu_\nu} x'^\nu = x^\mu \quad \forall \rho \)

must have
\[ x^\mu = \delta^\mu_\nu \quad \text{ (Kronecker delta = 1 when } \mu = \nu \text{ )}
\]
\[ = 0 \quad \text{ when } \mu \neq \nu \]

\( \therefore \) \( \Lambda^\mu_\nu \) and \( \Lambda^\nu_\mu \) are inverse matrices.
\( \Lambda \quad \Lambda^{-1} \)
1.4 Vectors (contravariant vectors)

- The general mathematical concept of vector space: a collection of objects, "vectors," which can be added together and multiplied by real numbers in a linear way:

\[(a+b)(V+W) = aV + bV + aW + bW\] (1)

- Vector field: a set of vectors, one for each point of spacetime (manifold M), or some subset of spacetime, e.g., worldline.

- In physics, "vector" has a more specific meaning: a vector has a direction, corresponding to a direction in space/spacetime (and a magnitude).

\[\Rightarrow\] dimension (number of components) = dimension of space/spacetime

- In Newtonian physics, vectors are 3-dimensional: \(\mathbb{R}^{3}\)

- In SR and GR, vectors are 4-dimensional: \(\mathbb{R}^{4}\)

  (quite defined, e.g. \(\times\) cross product)

  - Often they are called four-vectors (4-vector, contravariant vectors)

  They will belong to vector fields: a given 4-vector is associated with a given point (event) at spacetime.

\[p\]

- To get the correspondence between vector and spacetime directions, 4-vectors are defined in terms of tangent vectors to spacetime curves.

**In Minkowski space (SR) it is also possible to introduce position 4-vectors, extending from one event \(P\) to another event \(Q\). Its components are \(x^M(Q) - x^M(P)\). In curved spacetime (GR) this no longer works, and a given vector is necessarily associated with just a single point.**
A parametrized curve is a map $\mathbb{R} \rightarrow M$

(map a real number $\lambda$ to a point $P$ in spacetime)

For a given real system, where $P$ has coords $x^\mu(P)$, the curve has a coord. representation $x^\mu(\lambda)$, a map $\mathbb{R} \rightarrow \mathbb{R}^n$.

The tangent vector to this curve has components

$$\frac{dx^\mu}{d\lambda}$$

Tangent space at point $P$: $T_P$: the set of all possible tangent vectors at $P$ at parametrized curves passing through $P$.

This is a $\mathbb{R}^n$-vector located at $P$.

The vector space at $P$ (of a particle) is a $\mathbb{R}^n$-vector.

Example: $\gamma$-velocity $\gamma$ = tangent vector at worldline parametrized by proper time $\tau$.

In coords: worldline $x^\mu(\tau)$

components at $\gamma$: $\gamma^\mu = \frac{dx^\mu}{d\tau}$ (index upstairs)

To get dimensional $\gamma$-vector, we allow multiplication by dimensional numbers, e.g.

$\gamma$-momentum $p = m\gamma$ $p^\mu = m\gamma^\mu = m\frac{dx^\mu}{d\tau}$

$m$ = mass at no particle (also called the rest mass, a Lorentz invariant quantity)

The concept of worldline and $\gamma$-vector are independent at all systems.

The components of a $\gamma$-vector are wrt some real system, and change with a real transformation $K \rightarrow K'$. 

$$\gamma^\mu = \frac{dx^\mu}{d\tau} \quad \gamma^\mu = \frac{dx^\mu}{d\tau} = \frac{\partial x^\mu}{\partial x^\nu} \frac{dx^\nu}{d\tau} = \Lambda^\mu_\nu \gamma^\nu$$

One way to define $\gamma$-vector: Obj. w $\gamma$ components defined in terms of coordinates, such that the components transform as (4).

(Then in a real dependent definition. Will give a more "elevated" real-independent definition, based on the above idea of a tangent vector, in Chapter 7.)
The tangent space $T_p$ has the same geometry as Minkowski space (vectors $v \in T_p$ corresponding to position vectors with fixed origin $O \in \mathbb{M}$).

In GR the spacetime is curved, if no more how the Minkowski geometry, but the tangent space $T_p$ still has the Minkowski geometry!

The world line $x^m(\tau)$ of a massive particle is timelike

\[ u^u \equiv \frac{dx^m}{d\tau} \] is timelike

since

\[ ds^2 = -\eta_{\mu\nu} dx^\mu dx^\nu \] (p. 1-11)

\[ u \cdot u = \eta_{\mu\nu} u^\mu u^\nu = \frac{\eta_{\mu\nu} dx^\mu dx^\nu}{ds^2} = -1 \]
Basis at a vector space: a set of linearly independent vectors \( \{ e_\mu \} \)

\[
V = V^\mu e_\mu
\]

(5) \( \mu \) labels basis vector, not component

We shall only use (at least in SR) coordinate bases, defined so that

\[
(\text{wth component of } \mu \text{th basis vector}) \quad (e_\mu)^\nu = \frac{dx^\nu}{dx^\mu} \bigg|_{x^\nu = \text{const}} = \frac{\partial x^\nu}{\partial x^\mu} = \delta^\nu_\mu
\]

(6) basis vectors are tangent vectors at coordinates (e.g. \( t \)-line has \( x, y, z = \text{const} \))

\[
(\text{wth component at } \mathcal{V} = (\mathcal{V})^\nu = (V^\mu e_\mu)^\nu = V^\mu (e_\mu)^\nu = V^\mu \delta^\nu_\mu = \delta^\nu_\mu)
\]

\[
(\text{wth component at } \mathcal{V} = (\mathcal{V})^\nu = (V^\mu e_\mu)^\nu = V^\mu (e_\mu)^\nu = V^\mu \delta^\nu_\mu = \delta^\nu_\mu)
\]

\[\Rightarrow \quad e_\mu = \Lambda^\nu_\mu e_\nu \Rightarrow \boxed{e_\mu = \Lambda^\nu_\mu e_\nu} \]

(6) \( \nu \) \( \downarrow \)

- How are the end bases at \( K \) and \( K' \) related?

\[
V = V^\mu e_\mu = V^{\nu'} e_{\nu'} = \Lambda^\nu_{\nu'} V^\mu e_{\nu'}
\]

\[\Rightarrow \quad e_\mu = \Lambda^\nu_\mu e_{\nu'} \Rightarrow \boxed{e_\mu = \Lambda^\nu_\mu e_{\nu'}}\]

The logic at the position of index:

\[
(\text{in a dot transformation } K \rightarrow K') \text{ an object with index } \mu
\]

must be multiplied with \( \Lambda^\nu_\mu \) : index is put upstairs and \( \Lambda^\mu_{\nu} \) : index is put downstairs

\[
(\text{in a dot transformation } K \rightarrow K') \text{ an object with index } \mu
\]

\[\Rightarrow \quad (V^\nu e_{\nu'}) = \Lambda^\nu_{\nu'} V^\mu e_{\nu'} = \boxed{V^\mu e_\mu} = \delta^\mu_\mu \quad \text{(unit index)}
\]

\[\Rightarrow \quad (V = V^\nu e_{\nu'}) = \Lambda^\nu_{\nu'} V^\mu e_{\nu'} = \boxed{V^\mu e_\mu} = \delta^\mu_\mu \quad \text{(unit index)}
\]

\[
(\text{i.e., is unchanged})
\]

\[
V = V^\nu e_{\nu'} = \Lambda^\nu_{\nu'} V^\mu e_{\nu'} = \boxed{\Lambda^\nu_{\nu'} \Lambda^\mu_{\nu} V^\mu e_\mu = V^\mu e_\mu}
\]
1.5. Dual Vectors (co-variant vectors, one-forms)

- For a given vector space, we can associate another vector space (of equal) dimension, its dual vector space.

The dual space of the tangent space $T_p$ is also called the cotangent space $T_p^*$: the set of linear maps (functions) from the vector space to real numbers,

$$\tilde{\omega} : T_p \to \mathbb{R}$$

linear $\Rightarrow \tilde{\omega}(av + bw) = a\tilde{\omega}(v) + b\tilde{\omega}(w)$ \hspace{1cm} (1)

- A dual vector $\tilde{\omega}$ is completely defined by the 4 numbers

$$\tilde{\omega}(e_\mu) = \omega_\mu \hspace{1cm} (2)$$

- Its components in a given orthonormal basis $K$ ($\{e_\mu\}$ being the orthonormal basis of $T_p$),

$$\tilde{\omega}(v) = \tilde{\omega}(v^\mu e_\mu) = v^\mu \tilde{\omega}(e_\mu) = \omega_\mu v^\mu \hspace{1cm} (3)$$

- The basis $\{\tilde{e}^\mu\}$ of the dual vector space is defined by

$$\tilde{e}^\mu(v) = \tilde{e}^\mu(e_\nu) \equiv \delta_\nu^\mu \hspace{1cm} (4)$$

$$\Rightarrow \tilde{\omega} = \omega_\mu \tilde{e}^\mu \hspace{1cm} (5)$$

- Since $\omega_\mu \tilde{e}^\mu(v) = \omega_\mu \tilde{e}^\mu(v^\nu e_\nu) = \omega_\mu v^\nu \tilde{e}^\mu(e_\nu) = \omega_\mu v^\nu \tilde{e}^\mu = \omega_\mu v^\nu \tilde{\omega}(v) = \omega_\mu v^\nu \tilde{\omega}(v)$

- How do the components $\omega_\mu$ transform in $K \rightarrow K'$?

Use the transformation law of vectors: $v^\mu = \Lambda^\mu_\nu v'^\nu$

Since $\tilde{\omega}(v)$ must be independent of vector's, i.e.,

$$\omega_\mu v^\nu = \omega_\mu' v'^\nu = \omega_\mu' \Lambda^\mu_\nu v'^\nu$$

we get $\omega_\mu = \omega_\mu' \Lambda^\mu_\nu \Rightarrow \omega_\mu' = \omega_\mu \Lambda^\mu_\nu$ \hspace{1cm} (6)
One way to define dual vector (or covariant vector): Objects with 4 components defined in terms of coordinates, such that the components transform as (6).

Thus we have:

**Contravariant vectors** \( \mathbf{V} : V^n = \Lambda^n_\mu V^\mu \)

**Covariant vectors** \( \omega_n : \omega_n^\mu = \Lambda^\mu_\nu \omega_\nu \)

As an example of a contravariant vector we had the tangent vector:

\[
V^\mu = \frac{dx^\mu}{d\lambda} : \frac{dx^\mu}{d\lambda} = \frac{\partial x^\mu}{\partial x^\nu} \frac{dx^\nu}{d\lambda} = \Lambda^\mu_\nu \frac{dx^\nu}{d\lambda}
\]

A prototypical example of a covariant vector is the gradient of some scalar function \( \Phi : M \rightarrow \mathbb{R} \) at spacetime point \( \omega_n \)

\[
\omega^\nu = \frac{\partial \Phi}{\partial x^\nu} : \frac{\partial \Phi}{\partial x^\nu} = \frac{\partial \Phi}{\partial x^\mu} \frac{\partial x^\mu}{\partial x^\nu} = \Lambda^\nu_\mu \frac{\partial \Phi}{\partial x^\nu}
\]

Now how is a gradient \( \frac{\partial \Phi}{\partial x^\mu} \) a linear map \( T_\xi \rightarrow \mathbb{R} \)?

Consider gradient acting on a vector, e.g., a 4-velocity

\[
\frac{\partial \Phi}{\partial x^\mu} \, u^\mu = \frac{\partial \Phi}{\partial x^\mu} \frac{dx^\mu}{dt} = \frac{d\Phi}{dt}
\]

= the observed rate of change of quantity \( \Phi \) with \( t \)

(a coordinate-dependent quantity, i.e., a Lorentz scalar)
1.6 Tensors in Minkowski Space

\* Linear functions \( T_p : \mathbb{R} \rightarrow \mathbb{R} \) \( \omega(V) \in \mathbb{R} \)

**Dual vectors**

form also a vector space \( \omega = \omega_\mu \epsilon^\mu \)

transform as \( \omega_\nu = \Lambda^\nu_\mu \omega_\mu \)

Duals and dual vectors are just vectors, since they will again transform as \( \delta^\nu_\mu = \Lambda^\nu_\mu \delta^\nu_\mu \)

We identify \( \delta^\nu_\mu \) with \( \delta \) by \( \delta(V) = \delta(\omega) = \delta^\nu_\mu = \delta_\nu^\mu \)

- We can now proceed to define higher-order tensors in the same way:

  - **What are tensors?**

  My definition: Tensors are (multi)linear tensor-valued maps \( \Lambda^{(\text{functions})} \) at other tensors. This definition sounds too recursive, but we can proceed from lower order tensors already defined (vectors, and scalars, which are 0th order tensors). Operation of this map on its arguments is called multiplication by the tensor. Linear maps are so simple that we don't usually think of tensors as maps, but just as objects that can be multiplied with each other.

  Scalars \( a \in \mathbb{R} \) are real numbers (typically with some unit) associated with points of spacetime, which are independent of (no sense in all) coordinates.

  Operation of a scalar on another scalar

  \( a(b) = ab = c \in \mathbb{R} \)

  Operation of a scalar on a vector

  \( a(V) = aV = W \) (where \( W \parallel V \) )

*) Much of this stuff applies as is for the general spacetimes at GR also.

**) When I say "vector" here, I mean a 4-vector. (Obviously this math can be applied to manifolds of other dimensionality, too.)
1st order tensors: Vectors and Dual vectors

After having defined \( \text{vectors} \) (contravariant vectors) as \( \text{elements of } T_p \), we defined \( \text{dual vectors} \) (covariant vectors) as \( \text{linear functions} \ \tilde{\omega} : T_p \rightarrow \mathbb{R} \)

\[ \tilde{\omega}(v) = \omega^\mu v_\mu \in \mathbb{R} \]

They form a vector space, the cotangent space \( T_p^* \).

Vectors can also be thought of as linear functions \( \nabla : T_p^* \rightarrow \mathbb{R} \) defined by

\[ \nabla(\tilde{\omega}) = \tilde{\omega}(\nabla) = \nabla^\mu \omega_\mu = \omega_\mu v^\mu \quad \text{"duals of duals"} \]

2nd order tensors: Covariant, Contravariant, and Mixed

Covariant tensors can be defined as \( \text{linear functions} \ A : T_p \times T_p \rightarrow \mathbb{R} \)

\[ A(u, v) = b \in \mathbb{R} \]

but they can also be viewed as \( T_p \times T_p^* \)

by telling them operate only on one vector:

\[ A(u, \_ ) \in T_p^* \quad \text{since it remains free to operate on another vector:} \]

if \( \tilde{\omega} = A(u, \_ ) \), then \( \tilde{\omega}(v) = A(u, v) \)

Contravariant tensors can be defined as \( \text{linear functions} \ B : T_p \rightarrow T_p \)

(probably the most common use of a mixed 2nd order tensor), so that

\[ B(v) = w \quad \text{where usually } w \parallel v \]

but they can also be viewed as \( T_p^* \times T_p \rightarrow \mathbb{R} \), i.e., one should have written

\[ B(\_, v) = w \]

and now \( B(\tilde{\omega}, v) = \tilde{\omega}(B(\_, v)) = \tilde{\omega}(w) \quad (\text{This may look complicated,}) \]

but everything is really simple, because of the linearity, as becomes manifest when we switch to component notation.

Contravariant tensors can be defined as \( C : T_p^* \times T_p^* \rightarrow \mathbb{R} \)

\[ C(\tilde{\omega}, \tilde{\eta}) = d \in \mathbb{R} \]
That tensor operation is just multiplication, because apparent as we switch to component notation. Because of the linearity, tensors are fully defined by their components.

\[ A_{\mu} = A(e_\mu, e_\nu) \Rightarrow A(y, v) = A_{\mu} U^\mu V^\nu \]

\[ B^\mu = B(e^\mu, e_\nu) \Rightarrow B(\omega, \nu) = B^\mu W^\mu W^\nu \]

\[ C^\nu = C(e^\mu, e^\nu) \Rightarrow C(\omega, \nu) = C^\nu W^\mu \]

The tensor product of two dual vectors \( \omega \otimes \nu \) is a covariant tensor defined so that

\[ (\omega \otimes \nu)(y, v) = \omega(y) \nu(v) = \omega_{\mu} U^\mu V^\nu \]

Its components are just \( \omega_{\mu} V^\nu \).

Likewise we have tensor products \( y \otimes v \) (a contravariant tensor) and \( y \otimes \nu \) (a mixed tensor).

Not every tensor is a tensor product of two vectors; but tensor products of basis vectors give us a basis for the tensor space:

\[ A = A_{\mu} e^\mu \otimes e_\nu, \]

\[ A(e_\nu, e_\sigma) = A_{\mu} e^\mu(\xi_\nu) e^\nu(\xi_\sigma) = A_{\nu \sigma} \]

Likewise \( B^\mu = B^\mu_\nu e_\mu \otimes e_\nu \)

and \( B(y) = B^\mu_\nu e_\mu(\nu \xi_\nu) = B^\mu_\nu V^\mu_\nu \)

Thus we can write the equation \( B(y) = \nu \) in component notation as

\[ B^\mu_\nu V^\mu_\nu = \nu \]
- We shall later encounter also higher-order tensors. E.g. a fourth-order
covariant tensor $R$: $R(\alpha, \beta, \gamma, \delta) = R_{\mu
u\rho\sigma} \alpha^\mu \beta^\nu \gamma^\rho \delta^\sigma$

- Carroll defines tensors as linear maps (functions)

$$T: T_p^* \times \cdots \times T_p^* \times T_p \times \cdots \times T_p \rightarrow \mathbb{R}$$

is a tensor of type $(r, s)$

$r + s$ is the order of the tensor

if $s = 0$, we say the tensor is contravariant

$r = 0$

covariant

otherwise

mixed

- E.g., tensor of type $(3,2)$: $T = T^{\mu\nu\rho} \varepsilon_{\mu\nu\rho} \varepsilon \varepsilon \varepsilon$

$$T(\alpha, \beta, \gamma, \delta, \varepsilon) = T^{\mu\nu\rho} \varepsilon_{\mu\nu\rho} \delta^\alpha \beta^\beta \varepsilon^\gamma$$

- But note we can leave some slots unmapped, arriving at my more general
definition, e.g.

$$T(\alpha, \beta, \gamma) = T^{\mu\nu\rho} \varepsilon_{\mu\nu\rho} \delta^\alpha \beta^\beta \varepsilon^\gamma = \epsilon_{\mu} \epsilon_{\nu} \epsilon_{\alpha} = 0$$

a mixed 2nd order tensor

In component notation this reads

$$T^{\mu\nu\rho} \varepsilon_{\mu\nu\rho} \delta^\alpha = \epsilon_{\mu}$$
The essential property of tensors is how their components transform under coordinate transformations. From our definitions, which require that the scalars, e.g., \( A(y, v) \), obtained from tensors operating on vectors, are independent of coordinates used, and the vector and transformation property \( V^\mu = \Lambda^\mu_{\nu}V^\nu \) follows that tensor components transform as expected from the change of their indices, e.g.,

\[
T^\mu_{\nu\lambda} = \Lambda^\mu_{\alpha} \Lambda^\nu_{\beta} \Lambda^\lambda_{\gamma} T^\alpha_{\beta\gamma}
\]

Tensors are sometimes defined as objects (usually given as some expressions relating to coordinates, e.g., \( U^\mu = dx^\mu/dt \)) whose components satisfy this transformation law.

Note that not all objects with components defined in terms of coordinates are tensors. For example, the coordinate velocity \( U^\mu = dx^\mu/dt \) is not a 4-vector. Later (in GR) we'll encounter the Christoffel symbol \( \Gamma^\mu_{\nu\beta} \), not a tensor.

*) This is the first part that requires modification when we switch to GR. Then the Lorentz transformations \( \Lambda^\mu_{\nu} \) are replaced by general coordinate transformations.
Special tensors, whose components are the same in all inertial frames

- Metric tensor, type (0,2)
  \[ \eta(\mathbf{v}, \mathbf{w}) = \eta_{\mu\nu}v^\mu w^\nu = v \cdot w = -v^0w^0 + v^1w^1 + v^2w^2 + v^3w^3 \]

  the inner product (scalar product, dot product) of \( v \) and \( w \).

  We introduced \( \eta_{\mu\nu} \) as the matrix \( \text{diag}(-1,1,1) \)

  Eq. (5) \( \eta_{\mu\nu} = \Lambda^\mu_\gamma \Lambda^\nu_\sigma \eta_{\gamma\sigma} \Leftrightarrow \eta_{\mu\nu} = \Lambda^\mu_\alpha \Lambda^\nu_\beta \eta_{\alpha\beta} \)

  shows that \( \eta_{\mu\nu} \) is a (0,2) tensor, whose components are the same in all inertial frames!

  The inner product is a scalar, invariant under Lorentz transformations.

If \( v \cdot w = 0 \), \( v \) and \( w \) are orthogonal, \( v \perp w \)

The basis vectors of inertial frames are orthogonal to each other:

\[
\eta_{\mu\nu} (e^\mu) (e^\nu) = \eta_{\mu\nu} e^\mu e^\nu = \eta_{\mu\nu} = 0 \quad \text{for} \quad \mu \neq \nu
\]

The norm of a vector: \( v \cdot v = \eta_{\mu\nu} v^\mu v^\nu \) (also called the "square" of a vector)

- \( v \cdot v < 0 \) : \( v \) is timelike
- \( v \cdot v = 0 \) : \( v \) is lightlike (null), orthogonal to itself!!
- \( v \cdot v > 0 \) : \( v \) is spacelike \( |v| = \sqrt{v \cdot v} \) the length of \( v \)

- Kronecker tensor, type (1,1) : the identity map \( T_p \rightarrow T_p \)
  \( \delta^\mu_\nu = v \) components \( \delta^\mu_\nu = \delta^\mu_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \)
  \( \delta^\mu_\nu v^\nu = v^\mu \)

- Inverse metric, type (2,0)
  \[ \eta(\tilde{v}, \tilde{w}) = \eta_{\mu\nu} \tilde{v}^\mu \tilde{w}^\nu = \tilde{v} \cdot \tilde{w} \] the inner product of dual vectors

\[ \gamma_{\mu\nu} \text{ and } \gamma^{\mu\nu} \text{ have the same components} \]

\[ [\gamma_{\mu\nu}] \text{ is the inverse matrix of } [\gamma_{\mu\nu}] : \gamma^{\mu\nu} \gamma_{\nu\lambda} = \delta^{\mu}_{\lambda} \]

In GR, the metric tensor is denoted \( g_{\mu\nu} \), and is no longer diag(-1,1,1).

We still have \( g_{\mu\nu} g^{\mu\nu} = \delta_3 \), but \([g_{\mu\nu}]\) and \([g^{\mu\nu}]\) are different matrices, and different in different coord. systems. In GR we shall still use \( \gamma_{\mu\nu} \), \( \gamma^{\mu\nu} \) as symbols for "arrays" for \([\gamma_{\mu\nu}] = [\gamma^{\mu\nu}] = \text{diag}(-1,1,1), \) no longer tensors."
Levi-Civita symbol, type (0,4)

\[ \tilde{\varepsilon}^{\mu\nu\rho\sigma} = \begin{cases} +1 & \text{even permutation of 0123} \\ -1 & \text{odd} \\ 0 & \text{otherwise (two or more of } \nu\rho\sigma \text{ are the same)} \end{cases} \]

(We assume here a right-handed cdt. system; for \( \tilde{\varepsilon}^{\mu\nu\rho\sigma} \) to be a tensor under parity transformations, we need to define it with opposite sign for left-handed cdt. systems.)

odd \} \text{ permutation} \Leftrightarrow \text{odd} \} \text{ number of exchanges of two digits},

even \} \text{ permutation} \Leftrightarrow \text{even} \} \text{ number of exchanges of two digits},

e.g., 0123, 0132, 0312, 3012, 2013

odd \ even \ odd \ even

\[ \tilde{\varepsilon}^{\mu\nu\rho\sigma} \text{ is a tensor for Minkowski space (using inertial frames only).} \]

In GR we shall define the Levi-Civita tensor \( \varepsilon^{\mu\nu\rho\sigma} \) whose components differ in different cdt. systems; but also use the above \( \tilde{\varepsilon}^{\mu\nu\rho\sigma} \) as a symbol, no longer a tensor.
1.7 Manipulating Tensors

We mostly work in component notation. The various tensor operations are just products and sums for components.

- **Tensor product** \((r,s)\) and \((k,l)\) tensors \(\rightarrow\) \((r+k, s+l)\) tensors

\[
T = A \otimes B : \quad T_{\mu \nu \ldots}^{\alpha \beta} = A^{\mu \nu \ldots} B_{\alpha \beta}^{\ldots}
\]

does not commute, \(A \otimes B \neq B \otimes A\) in general

- **Contraction** \((r,s)\) tensor \(\rightarrow\) \((r-1, s-1)\) tensor

\[
S^{\mu \nu} = T^{\mu \nu \ldots} \quad \text{sum over repeated index, one upper, one lower}
\]

- **Tensor operating on a vector** = tensor product + contraction

\[
T(v) = w \quad T_{\alpha}^{\beta} v^{\beta} = w^{\alpha}
\]

(\text{which means: } \sum_{\beta=0}^{3} T_{\alpha}^{\beta} v^{\beta} = w^{\alpha}, \quad \alpha = 0, 1, 2, 3)

- The ordering of the arguments of a tensor (in component notation, the order of the indices) does matter: usually \(T(2,1) \neq T(1,2)\).

- Some tensors are **symmetric**:

\[
S(a,b) = S(b,a) \quad \forall a,b \iff S_{ab} = S_{ba}
\]

or **antisymmetric**:

\[
A(a,b) = -A(b,a) \quad \forall a,b \iff A_{ab} = -A_{ba}
\]

- Further notation: Symmetrization/antisymmetrization of tensors:

\[
T_{[\mu}^{\nu]} = \frac{1}{2} (T_{\mu}^{\nu} + T_{\nu}^{\mu}) = \frac{1}{n!} \text{(sum over permutations)}
\]

\[
T_{[\mu \nu \ldots]} = \frac{1}{2} (T_{\mu \nu \ldots} - T_{\nu \mu \ldots}) = \frac{1}{n!} \text{(alternating sum)}
\]

For example,

\[
T_{[\mu \nu \ldots]}^{\alpha \beta} = \frac{1}{6} (T_{\mu \nu \ldots}^{\alpha \beta} - T_{\mu \nu \ldots}^{\beta \alpha} + T_{\mu \nu \ldots}^{\alpha \beta} - T_{\nu \mu \ldots}^{\alpha \beta} + T_{\nu \mu \ldots}^{\beta \alpha} - T_{\nu \mu \ldots}^{\alpha \beta})
\]
Raising and lowering indices: We can use the Minkowski metric $\eta$ to produce a set of associated tensors (of different type but the same order) for a given tensor:

$$
T^\mu_\nu \overset{\delta}{=} \eta^{\mu \kappa} T_{\kappa \nu} \\
T^\mu_\nu \overset{\delta}{=} \eta^{\mu \kappa} T_{\kappa \nu} \\
T^\mu_\nu \overset{\delta}{=} \eta^{\mu \kappa} \eta^{\nu \lambda} T_{\kappa \lambda} \\
T^\mu_\nu \overset{\delta}{=} \eta^{\mu \kappa} \eta^{\nu \lambda} \eta^{\gamma \delta} \eta_{\gamma \delta} T_{\kappa \lambda} \\
$$

(note that each index keeps its horizontal position).

These are usually thought of "as the same tensor", just different kinds (contravariant, covariant, and various mixed) of components of the tensor.

In particular, the covariant $\eta_\mu$ and the contravariant $\lambda^\mu$ components of the vector $\lambda$:

$$
\lambda_\mu \equiv \eta_\mu \lambda^\nu \\
(\lambda_0, \lambda_1, \lambda_2, \lambda_3) = (-t, x^1, x^2, x^3)
$$

The inner product can now be written:

$$
\begin{align*}
\lambda \cdot \lambda & = \eta_{\mu \nu} \lambda^\mu \lambda^\nu = \eta_0^0 \lambda^0 \lambda^0 + \eta^1_1 \lambda^1 \lambda^1 + \eta^2_2 \lambda^2 \lambda^2 + \eta^3_3 \lambda^3 \lambda^3 \\
& = -\lambda^0 \lambda^0 + \lambda^1 \lambda^1 + \lambda^2 \lambda^2 + \lambda^3 \lambda^3
\end{align*}
$$

(Show that the $\lambda_\mu$ indeed transform as the components of a contravariant (dual) vector:

$$
\begin{align*}
\lambda_\mu \equiv \eta_{\mu \nu} \lambda^\nu & = \eta_{\mu \nu} \lambda^\nu \\
& = \Lambda^\rho_{\mu} \eta_{\rho \sigma} \lambda^\sigma = \lambda_\rho \lambda^\rho
\end{align*}
$$

From the defining property of Lorentz transformation,

$$
\begin{align*}
\eta_{\rho \sigma} \lambda^\rho \lambda^\sigma & = \eta_{\rho \sigma} \Lambda^\rho_\mu \Lambda^\sigma_\nu \lambda^\mu \lambda^\nu \\
& \Rightarrow \eta_{\rho \sigma} \Lambda^\rho_\mu \Lambda^\sigma_\nu \lambda^\mu \lambda^\nu \\
& \Rightarrow \eta_{\rho \sigma} \Lambda^\rho_\mu \Lambda^\sigma_\nu \lambda^\mu \lambda^\nu = \eta_{\rho \sigma} \Lambda^\rho_\mu \Lambda^\sigma_\nu \lambda^\mu \lambda^\nu
\end{align*}
$$

For symmetric tensors $S_{\mu \nu} = \eta^{\alpha \beta} S_{\alpha \beta}$ is also symmetric

and $S_{\mu \beta} = \eta^{\alpha \beta} S_{\alpha \beta} = \eta^{\alpha \beta} S_{\beta \mu} = S^\alpha_\beta = S^\alpha_\beta$ (but $S_{\alpha \beta} \neq S^\alpha_\beta$ !)

*) Really we have defined a covariant vector $\tilde{\lambda}$, associated with $\lambda$, called the dual of $\lambda$; but we shall just talk about a vector $\lambda$, and its contravariant components $\lambda^\mu$ and covariant components $\lambda_\mu$. $\tilde{\lambda}(\lambda) \equiv \lambda \cdot \lambda \in \mathbb{R}$
The trace of a \((1,1)\) tensor \(X = X^\lambda_{\ \lambda}\)

\((0,2)\) tensor \(Y = \eta^{\mu\nu} Y_{\mu\nu} = Y^\lambda_{\ \lambda} \quad (\neq \sum_{\mu} Y_{\mu\mu})\)

Note that the trace of the metric is

\[\eta^{\mu\nu} Y_{\mu\nu} = \gamma_{\mu} = 4 \quad \text{(not } -1+1+1+1 = 2)\]

Tensor fields and partial derivatives:

So far everything was at a single spacetime point (event).

In practice we have tensor fields: In component/coordinate notation, tensor components are functions of coordinates \(A^\mu_{\ \nu}(x^\sigma)\).

Partial derivative \(\frac{\partial}{\partial x^\sigma}: (r,s)\) tensor \(\rightarrow (r,s+1)\) tensor \(\star\)

\[
\frac{\partial}{\partial x^\sigma} A^\mu_{\ \nu} = \delta^\mu_{\ \sigma} = A^\mu_{\ \sigma}
\]

Introduce notation \([\frac{\partial}{\partial x^\mu} \equiv \partial_\mu \equiv \partial_\mu]\)

Partial derivatives commute: \(\partial_\mu \partial_\nu (\cdots) = \partial_\nu \partial_\mu (\cdots)\)

\[A^\mu_{\ \nu,\sigma} = A^\mu_{\ \nu,\sigma}\]

\(\star\) In GR this is not true for partial derivatives (except for scalars: \(f, a\) in a dual vector), which motivates us to define covariant derivatives.

Covariant derivatives do NOT commute.
Some Standard 4-Vectors and Relativistic Mechanics

- Introduce notation \(\Lambda^\mu = (\Lambda^0, \Lambda^1, \Lambda^2, \Lambda^3) = (\lambda^0, \lambda^i) = (\chi^0, \chi^i)\) (*

- 4-velocity \(u^\mu = \frac{dx^\mu}{dt}\)
  \(u \cdot u = \gamma_{\mu\nu} u^\mu u^\nu = \frac{d\mu}{d\tau^2} = -1 = -c^2\)

  Compare this to the coordinate velocity \(V^\mu = \frac{dx^\mu}{dt} = (v^1, v^2)\)
  which is not a 4-vector!

  \[\frac{dt}{d\tau} = \sqrt{1 - v^2} = \gamma\]

  \[\gamma = \sqrt{(v^1)^2 + (v^2)^2 + (v^3)^2}\]

  \[V = (v^1, v^2, v^3)\]

\[u^\mu = (1, v^1, v^2, v^3) \quad u_\mu = (-1, 0, 0, 0)\]

- 4-acceleration \(a^\mu = \frac{du^\mu}{d\tau^2}\)

  Since \(u \cdot u = -1\) = const (length of \(u\), \(|u| = 1\), snot const!),
  \(a\) must always be orthogonal to \(u\)!

  Check this:

  \[a \cdot u = g_{\mu\nu} u^\mu a^\nu = \gamma_{\mu\nu} \frac{du^\mu}{d\tau^2} \left(\gamma_{\nu\nu} u^\nu u^\nu\right) = 0\]

  \[\frac{d}{d\tau} \left(\gamma_{\mu\nu} u^\mu u^\nu\right) = \gamma_{\mu\nu} \frac{du^\mu}{d\tau} u^\nu + \gamma_{\mu\nu} \frac{du^\mu}{d\tau} u^\nu = 2 \gamma_{\mu\nu} \frac{du^\mu}{d\tau} u^\nu\]

  \[= \gamma_{\mu\nu}\]

- Mass \(m\) and charge \(q\) of a particle are scalars. (\(\equiv\) they are invariant under a Lorentz transformation).
  Multiplying a 4-vector with a scalar, or taking \(\frac{d}{d\tau}\) of it (but not \(\frac{d}{dt}\)) gives another 4-vector

*) So the "general component" \(\Lambda^\mu\) represents the full set of components.

We do not write \(\Lambda = \Lambda^\mu\), instead we write \(\Lambda = \Lambda^\mu e_\mu\),
keeping the coordinate-independent concept \(\Lambda\) separate from its component representation \(\Lambda^\mu\). However, I care less about this distinction when discussing the space part, and I do write \(\Lambda = \Lambda^i\), since the separation of \(\Lambda\) into \(\Lambda\) and \(\chi\) already depends on the choice of inertial frame (is not invariant under boosts).
\[ \mathbf{p}^\mu = m u^\mu = (m \gamma, m \gamma v) = (E, \mathbf{p}) \]

Two energy is not a scalar, but it is the 0th component of a 4-vector: \( E = p^0 \)

\[ \mathbf{p} \cdot \mathbf{p} = p^0 p^0 = m^2 u^0 u^0 = m^2 - m^2 c^2 = -E^2 + p^2, \quad \text{where} \quad p = \| \mathbf{p} \| \]

\[ \Rightarrow \quad E = \sqrt{m^2 + p^2} = \sqrt{m^2 c^2 + p^2} \]

Newtonian mechanics is not compatible with the geometry of Minkowski space, and needs to be replaced by relativistic mechanics.

The Newtonian relations \( E = \frac{1}{2} m v^2 \) and \( \mathbf{p} = m \mathbf{v} \)

are replaced by \( E = m \gamma \) and \( \mathbf{p} = m \gamma \mathbf{v} \)

Together, energy \( E \) and momentum \( \mathbf{p} \) form a 4-vector.

Since \( \gamma = \frac{1}{\sqrt{1 - v^2}} = 1 + \frac{1}{2} v^2 + \frac{3}{8} v^4 + \ldots \) we have the expansion

\[ E = m + \frac{1}{2} m v^2 + \frac{3}{8} m v^4 + \ldots \quad (= m c^2 + \frac{1}{2} m v^2 + \frac{3}{8} m v^4 + \ldots) \]

\( \uparrow \text{mass} = \text{"rest energy"} \)

The total 4-momentum of an isolated system is conserved, \( \Sigma p^\mu = \text{const.} \)

For example, in a collision,

\[ \Sigma p^\mu \text{ (before)} = \Sigma p^\mu \text{ (after)} \]

\( \mathbf{f} \)

Newton's 2nd law is replaced by

\[ f^\mu = \frac{\partial p^\mu}{\partial t} = ma^\mu \]

The space part of \( \frac{1}{\gamma} f \) may be called the 3-force \( \mathbf{F} \):

\[ f^\mu = \gamma (\mathbf{F} \cdot \mathbf{J}, \mathbf{F}) = \gamma \frac{dp^\mu}{dt} = \gamma \frac{dE}{dt}, \frac{d\mathbf{p}}{dt} \]

(Equivalence: show that this is consistent)

*) Originally, there was some confusion about terminology: the relation \( \mathbf{p} = m \gamma \mathbf{v} \) suggested that \( m \gamma \mathbf{v} \) could be called the "mass"; giving a velocity-dependent, i.e., a frame-dependent mass. In this older terminology, our bundled-invariant mass in was called the "rest mass". Note that the famous \( E = mc^2 \) applies for an object at rest. In general, \( E = \gamma mc^2 \) (\( \gamma c^2 \) of old terminology).
Wave 4-vector $k$

Photons are massless, $m = 0$, and travel along light-like (null) world lines. They do not have a 4-velocity (since they experience zero proper time), but they do have 4-momentum. It is related to their wave 4-vector

\[ k^\mu = \left( \frac{2\pi}{\lambda}, \frac{2\pi}{\lambda} \right) = \frac{2\pi}{\lambda} (1, \vec{k}) = (\omega, \vec{k}) \quad \omega = 2\pi v \]

where $\lambda = \frac{c}{v} = \frac{1}{v}$ is the wavelength, $v$ the frequency, and $\vec{k}$ (unit 3-vector) the direction of propagation of the photon.

\[ k \cdot k = k^\mu k^\mu = \left( \frac{2\pi}{\lambda} \right)^2 (1 + \vec{k} \cdot \vec{k}) = 0 \quad \text{so} \quad k \text{ is light-like (null)} \]

The 4-momentum of a photon is

\[ p^\mu = \hbar k^\mu \quad \Rightarrow \quad E = \hbar \omega = \hbar v, \quad p = \hbar \vec{k} \]

\[ p \cdot p = p^\mu p^\mu = \hbar^2 k^\mu k^\mu = -E^2 + p^2 = 0 \quad \Rightarrow \quad E = p = \hbar v \]
Observations

- Observed particle energy, 3-momentum, photon wavelength, direction etc.

- The time and space components of the 4-vectors in the observer rest frame.

Let $u$ be the observer 4-velocity.

In her rest frame $u^i = 0 \implies u^\mu = (1, 0, 0, 0)$

observer's proper time

Consider a particle w 4-momentum $p^\mu$ and photon w wave 4-vector $k^\mu$

in the observer rest frame

\[
u \cdot p = u^\mu p_\mu = -p^0 = -E
\]

\[
u \cdot k = u^\mu k_\mu = -k^0 = -\frac{2\pi}{\lambda}
\]

\[
E_{\text{obs}} = -\nu_{\text{obs}} \cdot p
\]

\[
\lambda_{\text{obs}} = -\frac{2\pi}{\nu_{\text{obs}} \cdot k}
\]

$\nu_{\text{obs}} \cdot p$ and $\nu_{\text{obs}} \cdot k$ are Lorentz scalars, and can thus be calculated in any frame.
1.8. Maxwell's Equations

The tensors are important in relativity; because a fundamental idea in relativity is that laws of nature can be expressed as tensor equations.

- We already had the law of relativistic dynamics: \[ f^\mu = \frac{dp^\mu}{dt} \]

- Consider new electrodynamics. Classical electrodynamics turned out to be relativistic "as is".

Maxwell eqns

\[ \nabla \cdot \mathbf{E} = 0 \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \]

\[ \nabla \cdot \mathbf{B} = 0 \quad \nabla \times \mathbf{B} = \mathbf{J} + \frac{\partial \mathbf{E}}{\partial t} \]

Lorentz force

\[ \mathbf{F} = \frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \]

Can be written as tensor equations (**)

\[ F^\mu_{\nu} = j^\mu \]

\[ F_{\mu\nu} + F_{\nu\mu} + F_{\gamma\mu\nu} = 0 \]

where \( j^\mu = (\mathbf{S}, \mathbf{J}) \)

\[ [F_{\mu\nu}] = \begin{bmatrix} 0 & -E^1 & -E^2 & -E^3 \\ -E^1 & 0 & B^3 & -B^2 \\ -E^2 & -B^3 & 0 & B^1 \\ -E^3 & B^2 & -B^1 & 0 \end{bmatrix} \]

 antisymmetric

\[ F^{\mu\nu} = \Lambda^{\mu}_{\sigma} \Lambda^{\nu}_{\omega} F^{\sigma\omega} \]

(**) The "covariant" form of Maxwell's equations. Tensor equations are called "covariant", meaning that they are true in any inertial frame (since both sides transform the same in a Lorentz transformation). This use of the word "covariant" has nothing to do with "covariant" as opposed to "contravariant".

(*) This can be written as \( F_{\mu\nu} = 0 \), since \( F_{\mu\nu} \) is antisymmetric.
Historical note: In 1905, Einstein did not yet know about tensors. In his famous 1905 paper where he presented SR, he first derived the Lorentz transformation
\[ t' = \gamma (t-vx), \quad y' = y, \]
\[ x' = \gamma (x-vt), \quad z' = z \]
from his first postulate (the invariance of the speed of light), and then found that the \( \vec{E} \) and \( \vec{B} \) fields should transform* as
\[ E'_x = E_x, \quad B'_x = B_x \]
\[ E'_y = \gamma E_y - \gamma v B_z, \quad B'_y = \gamma B_y + \gamma v E_z \]
\[ E'_z = \gamma E_z + \gamma v B_y, \quad B'_z = \gamma B_z - \gamma v E_y \]
for the Maxwell equations to remain valid in all inertial frames, as required by his second postulate (the invariance of the laws of physics).
Therefore Einstein concluded that the fields indeed do transform this way.
The tensor formulation came later.

Einstein's postulates for SR:
1. The speed of light \( c \) is the same in all inertial frames
2. The laws of nature are the same in all inertial frames

*) We proceeded in a different manner. We first formulated the Maxwell eqs as tensor equations (knowing already that tensor equations remain valid in all inertial frames), and then got the transformation law from general tensor transformation properties.

As an exercise, you can deduce Einstein's result, by showing that, indeed, the Maxwell equations remain the same if you do the transformation \((1)+(2)\).
1.9. Energy-Momentum Tensor

- Large number of particles → Describe as a continuum: fluid

  - local fluid velocity
  - average of particle velocities
  - often << individual particle velocities

  \[ \text{mass, energy, momentum} \rightarrow \text{mass density, energy density, momentum density} \]

- BUT density of 4-momentum is NOT a 4-vector,
  - since volume (3-volume) is not Lorentz-invariant:
    - Lorentz contraction in the direction of motion

- Instead we have the energy-momentum tensor \( T^{\mu\nu} \):

  \[ T^{\mu\nu} = \text{flux of 4-momentum} \ p^{\mu} \ \text{across surface of constant} \ x^{\nu} \]

  \[ T^{00} = \text{energy density} \]

  \[ T^{0i} = \text{energy flux} = T^{i0} = \text{momentum density} \]

  \[ T^{jk} = \text{momentum flux}: \]

  \[ j\text{-component of flux at momentum} \ k\text{-component} \]

  \[ = \ k\text{-component} \ - \ j\text{-component} \]

  \( \) (these turn out to be always equal)

  \[ : \ T^{\mu\nu} \text{ is symmetric} \]

- Momentum is transferred by

  a) motion of fluid

  b) forces between fluid elements
- Total energy and momentum

The total energy of a system is 
\[ E = \int T^{00} dV \]

The total momentum of a system is 
\[ p^i = \int T^{i0} dV \]

They depend on the frame they are evaluated in.
Together they form a 4-vector (in special relativity only).

- The energy-momentum tensor satisfies the energy-momentum continuity equation

\[ \frac{\partial T^{\mu \nu}}{\partial x^\mu} = 0 \quad \text{or} \quad \left[ T^{\mu \nu}, \nu \right] = 0 \tag{1} \]

For a closed system this leads to conservation of total energy

\[ \frac{dE}{dt} = \frac{d}{dt} \int T^{00} dV = \int \frac{\partial T^{00}}{\partial t} dV = \int T^{00}_{;0} dV \]

\[ = -\int T^{i0} dV = -\int T^{0i} g_{0i} = 0 \]

where the surface integral vanishes for a closed system, since there is no energy flux through its boundary.

Likewise we get conservation of total momentum

\[ \frac{dp^i}{dt} = 0 \]
In the fluid rest frame:
\[ T^{00} = \text{proper (rest frame) energy density} = \mathcal{E} \]
\[ T^{ii} \text{ describes stress} = \text{force between fluid elements} \]
\[ \text{diagonal terms} = \text{pressure} \quad p_1 = T^{11} \]
\[ (\text{in the three directions}) \quad p_2 = T^{22} \]
\[ p_3 = T^{33} \]
\[ \text{off-diagonal terms} = \text{shear} \]

. Dust: particles at rest wrt each other, no forces between them
\[ \Rightarrow \text{all have same 4-velocity } U^\mu \]
Number-flux 4-vector \[ n^\mu = n U^\mu = n(x^0, \mathbf{x^1}) \]
(4-velocity vector)
\[ \uparrow \text{number density in rest frame} \]
\[ \Rightarrow n^0 = n \gamma = \text{number density in given frame} \]
\[ \text{(since volume element Lorentz-contracted by } \frac{1}{\gamma} \text{)} \]
\[ n^i = n \gamma \mathbf{x}^i = n \mathbf{v} \quad \text{flux of particles in direction } \mathbf{i} \]

Suppose all particles have same mass \( m \)
\[ \Rightarrow \text{in rest frame, energy density is } \mathcal{E} = mn \]
\[ \text{in some other frame, energy density is } E^0 = mn^0 = P^0 \]
\[ = \text{0-component of } \mathbf{p} \quad \text{both 0-component of a 4-vector} \]
Define energy-momentum tensor for dust:
\[ T^{\mu\nu}_{\text{dust}} = \rho \gamma \gamma^\mu \gamma^\nu = \gamma \gamma^\mu \gamma^\nu \]
\[ \gamma \gamma^\mu \gamma^\nu \quad \text{proper energy density} \]
In rest frame \[ T^{0i}_{\text{dust}} = T^{ij}_{\text{dust}} = 0 \quad \text{dust has no pressure, no shear.} \]
Perfect fluid: frictionless, i.e. cannot support any shear

\[ \Rightarrow \text{pressure is isotropic, } p_1 = p_2 = p_3 \]

\[ \Rightarrow \text{in rest frame } T^{\mu \nu} = \begin{bmatrix} g & \rho \\ \rho & \rho & \rho \end{bmatrix} \]

General expression for arbitrary inertial frame:

Available: two scalars \( g, \rho \) (proper energy density, pressure)

- fluid 4-velocity \( u^\mu \)
- spacetime metric \( g^{\mu \nu} \)

\[ \Rightarrow \text{can construct two independent } (g_0) \text{ tensors: } u^\mu u_\mu \text{ and } g^{\mu \nu} \]

\[ \Rightarrow T^{\mu \nu} = A u^\mu u_\nu + B g^{\mu \nu} \]

\[ \Rightarrow \text{in rest frame } T^{\mu \nu} = \begin{bmatrix} A-B \\ B \\ B \end{bmatrix} \]

\[ \Rightarrow \begin{cases} A-B = g \\ B = \rho \end{cases} \Rightarrow A = g + \rho \]

\[ \therefore T^{\mu \nu} = (g + \rho) u^\mu u_\nu + \rho g^{\mu \nu} \] (2)

- The nature of the perfect fluid is determined by its equation of state \( p = p(g) \)
  - dust \( p = 0 \) (particles whose thermal velocities \( v << 1 \))
  - radiation \( p = \frac{1}{3} g \) (massless particles, e.g. isothermal photon gas)
Apply (1) $T_{\mu \nu, \mu} = T_{\mu \nu, \mu} = 0$ to a perfect fluid.

Note first $u^\nu u_\nu = -1 \Rightarrow 0 = (u^\nu u_\nu)_\mu = u^\nu u_{\nu, \mu} + u_{\nu} u_{\nu, \mu} = 2 u^\nu u_{\nu, \mu}$

$\Rightarrow u_{\nu} u_{\nu, \mu} = 0 \quad (3)$

whereas

$u_{\nu} u_{\nu, \mu} = \frac{\partial u^\nu}{\partial x^\mu} \frac{dx^\mu}{dt} = \frac{du^\nu}{dt} = u^\nu$ is the $u$-acceleration.

From (2) $T_{\mu \nu} = g_{\mu \nu} u^\nu + p u^\mu u_{\nu} + p u_{\mu} u_{\nu}$

$T_{\mu \nu} = (g_{\mu \nu})_{, \mu} u^\nu + g_{\mu \nu} u_{\mu, \nu} + p u_{\mu, \nu} u^\nu + p u_{\mu} u_{, \nu} u^\nu + p u_{\nu} u_{, \mu} u^\nu + p_{, \mu} y_{\nu} = 0 \quad (4)$

by Eq. (1).

Contract with $u^\nu$ to get a scalar equation

$u^\nu T_{\mu \nu, \mu} = -(g_{\mu \nu})_{, \mu} u^\nu + 0 - p_{, \mu} u^\nu - p u_{\mu, \nu} + 0 + p_{, \mu} u_{\nu} = 0$

$\Rightarrow (g_{\mu \nu})_{, \mu} + p u_{\mu, \nu} = 0 \quad (5)$\hspace{1cm} Energy continuity

Subtract $u^\nu (5)$ from (4):

$g_{\mu \nu} u^\nu u_{\mu, \nu} + p_{, \mu} u^\nu u_{\nu} + p u_{\nu} u_{, \mu} u_{\nu} + p_{, \mu} y_{\nu} = 0$

$\Rightarrow (g+p) u^\nu u_{\mu, \nu} + (y_{\mu} + u_{\mu} u^\nu) p_{, \mu} = 0 \quad (6)$

$\alpha^\nu$ projection to component orthogonal to $u^\mu$

Pressure gradient accelerates fluid

Eqs. (5) and (6) are the eqs. of special relativistic hydrodynamics (for a perfect fluid)