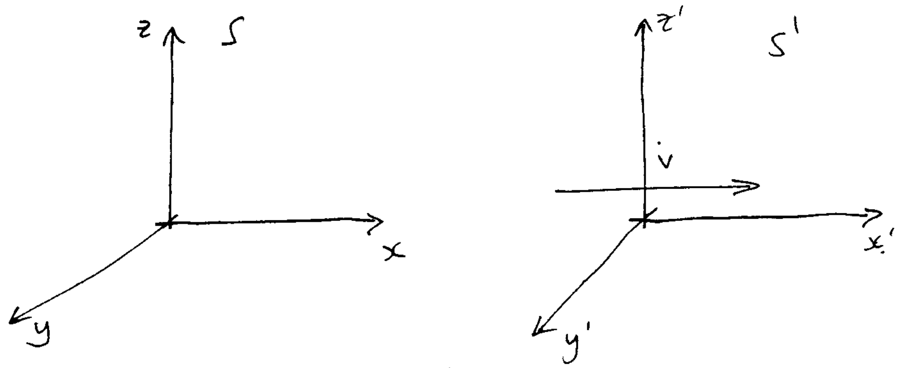
Goldstein Chap 7.

Lagrangian Formulation of Special Relativity

Before we explain how relativistic mechanics can be formulated using a Lagrangian, we need to first define the "4-vector" formalism & coordinate transformation between frames of reference.

As a preliminary, recall that in Newtonian mechanics for two frames of reference S, S' with coordinates (t, x, y, z) & (t', x', y', z'), then if S' is moving along the x-direction with speed v then

$$\begin{aligned} t' &= t \\ x' &= x - vt \\ y' &= y \\ z' &= z \end{aligned}$$



This coordinate transformation is called a Galilean transformation. Notice it can be described by a matrix:

$$\begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix}$$

Since the transformation uses a constant velocity v then we can show that if

$$\vec{F} = \frac{d}{dt}\vec{p} \quad \text{in } S, \text{ then}$$

$$\vec{F}' = \frac{d}{dt'}\vec{p}' \quad \text{in } S'.$$

However, the fundamental problem with the formulation of mechanics in the Newtonian framework is that if $\vec{u} = \frac{d\vec{x}}{dt}$ in S, then in S' we get $\vec{u}' = \vec{u} - \vec{v}$, which if $|\vec{u}| = c$ would violate the constancy of the speed of light.

In relativistic mechanics we employ a 4-dimensional space, called "spacetime." The vectors in spacetime are written $x^\mu \equiv (ct, x, y, z)$ (for example) and have a time component to them. Note that in x^μ we have $\mu = 0, 1, 2, 3$ and greek indices are assumed to run from 0 to 3 in relativistic mechanics. To describe the relative distance between points in spacetime, which we specify as a distance Δs^2 we use the following formula:

$$\begin{aligned}(\Delta s)^2 &= c^2(\text{time interval})^2 - (\text{space interval})^2 \\ &= c^2(\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2\end{aligned}$$

The use of the -ve sign on the spatial components effectively "breaks up" spacetime into separate regions. Consider the following:

Suppose we have two events "separated" by $(\Delta s)^2 = 0$. What does this mean?

If $$0 = c^2(\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2$$
then if we assume (without any real loss of generality) that $\delta y = \Delta z = 0$, then

$$c^2(\Delta t)^2 = (\Delta x)^2$$

$$\Rightarrow \quad c^2 = \left(\frac{\Delta x}{\Delta t}\right)^2$$

So the relative difference in spacetime between these events ends up describing a velocity equal to the speed of light. If the first point was $A \equiv (0,0,0,0)$ & the second point was $B \equiv (c\Delta t, c\Delta t, 0, 0)$ then

$$(\Delta s)^2 = c^2 \Delta t^2 - c^2 \Delta t^2 = 0$$

So the position of B in spacetime describes the position of a photon emitted from (A) at $t=0$. Since $(\delta S)^2 = 0$ we say that line element describing the distance between these two events is "null-like."

What about $(\Delta s)^2 > 0$? In this case take $A = (0,0,0,0)$ again but let $B \equiv (c\Delta t, \frac{1}{2}c\Delta t, 0, 0)$. So while we consider the same time difference as before we use a shorter physical (3d) distance. Thus any signal between these two events does not need to travel at c (can be slower).

$$\Delta s^2 = c^2 \delta t^2 - \frac{1}{4}c^2 \Delta t^2 = \frac{3}{4}c^2(\Delta t)^2 > 0$$

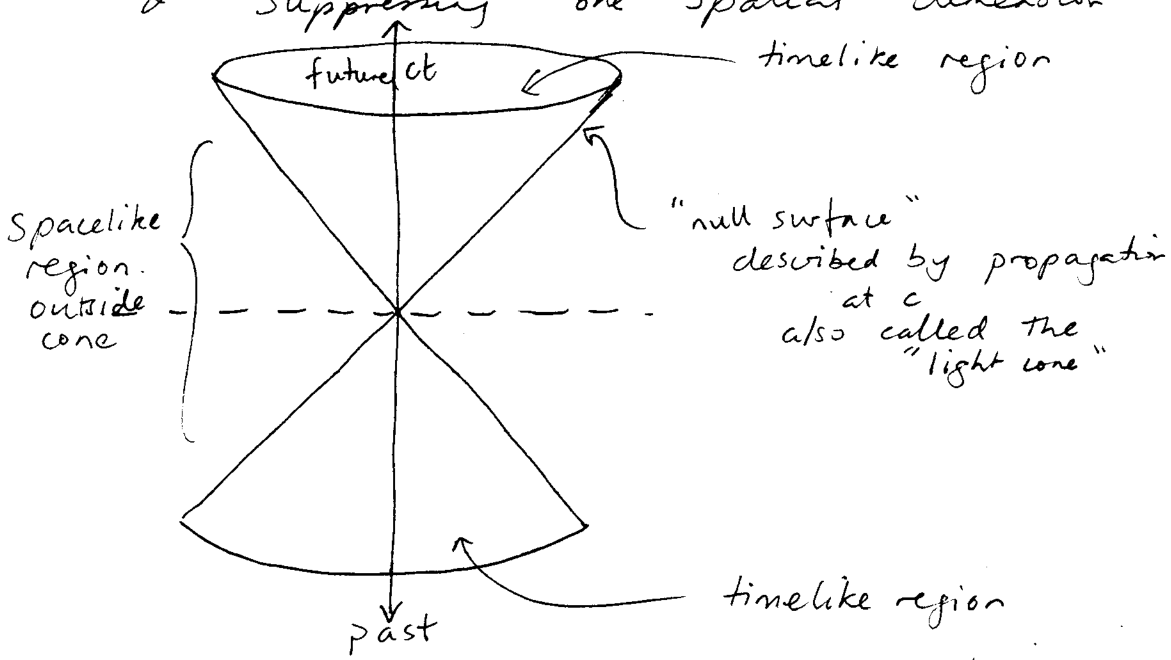
Thus if the relative velocity required to get from (A) to (B) is less than c, then the 4d line element $(\Delta s)^2 > 0$.

Lastly, $(\Delta s)^2 < 0$ corresponds to events that require speeds higher than c to link them (and so you cannot send any information between such events).

Now that we know how the "distance" between events is described in spacetime we can discuss particle movement. Any particle in spacetime will describe a single line, called its "worldline." Along this worldline the relative separation between points on it will be $(ds)^2$. The classification of the worldline is then

particle name	ds	
"photons"	0	"null" – corresponds to speed of light
tardyons	>0	"timelike" – speeds less than c
tachyons	<0	"Spacelike" – speeds greater than c

This separation is traditionally drawn in 2-d & suppressing one spatial dimension as follows:



This 4-d geometry is called "Minkowski space". The geometry builds in the constancy of the speed of light in all inertial reference frames. by requiring that ds^2 is the same for all observers. Consider the following:
If ds^2 were zero in one frame, which we know is the case for something moving at c, but were non-zero in another frame, this would imply a different speed in the other frame, ie c is not constant.
Hence the key part of relativistic dynamics is figuring out the coordinate transformations that preserve ds^2 - we'll come back to this shortly.

However, let us first consider the definition of time. Let "proper time" τ correspond to the time measured by an observer (A) that carries a clock along with their motion. Another observer (B) who is not moving with (A) and is at rest relative to an inertial reference frame is said to measure "laboratory time" t.

(A) does not measure any motion since they are moving with their frame of reference.
(B) assigns coordinates a time associated with the movement of (A), we will assume (A) moves with a velocity $\vec{v}$ as measured by (B).
Then for (B) in a time dt the separation between start & end points is

$$ds_B^2 = c^2 dt^2 - v^2 dt^2$$

(A) also measures a time interval between start & end points of $d\tau$. Thus for (A)

$$ds_A^2 = c^2 d\tau^2$$

then since $ds_A^2 = ds_B^2$ we find

$$c^2 d\tau^2 = c^2 dt^2 - v^2 dt^2 = dt^2(c^2 - v^2)$$

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

This means that $d\tau < dt$, making the moving clock appear to run slower (time dilation).

What is the form of the transformation that preserves ds^2? These are "Lorentz transformations". The simplest form of interest is the "Lorentz boost" where two inertial reference frames are connected by a velocity "boost" along the x-axis:

$$ct' = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}\left(ct - \frac{v}{c}x\right)$$

$$x' = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}\left(x - \frac{v}{c} \cdot ct\right)$$

$$y' = y$$
$$z' = z$$

This result can be derived quite quickly using something called the k-calculus (see hwk).

Let $\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$ from now on.

Then in matrix form we find

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & -\gamma\frac{v}{c} & 0 & 0 \\ -\gamma\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} \qquad \text{———— (A)}$$

We can also rotate the boost along a arbitrary direction and add a constant displacement. If we write (A) in matrix-vector form then

$$\vec{x}' = L_{Boost}\, \vec{x}$$

and the most general form would be

$$\vec{x}' = L_G\, \vec{x} + \vec{d}$$

where L_G is the matrix associated with the generalized Lorentz transformation along any direction and $\vec{d}$ is the displacement vector. Such transformations are called Poincaré transformations.

Note that we can write the transformation in component form using the summation convention

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu}\, x^{\nu}$$

where $\Lambda^{\mu}{}_{\nu}$ are equivalent to elements of the L matrix.

The component formalism also allows us to express the line element in a compact notation. Since

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

using 4-vectors we could write this as a sum over two sets of dx^μ as follows:

$$ds^2 = \eta_{\alpha\beta}\, dx^\alpha dx^\beta$$

If we define elements of $\eta_{\alpha\beta}$ by

$$\eta_{\alpha\beta} \equiv \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

then it will produce exactly the line element we want. Hence $\eta_{\alpha\beta}$ defines the geometry of the Minkowski spacetime, and plays a very important role. $\eta_{\alpha\beta}$ are actually components of the "metric tensor."

For any two vectors $\vec{u} \equiv u^\mu$, $\vec{v} = v^\nu$ defined by

$$\vec{u} = u^0 \hat{e}_0 + u^1 \hat{e}_1 + u^2 \hat{e}_2 + u^3 \hat{e}_3$$
$$\vec{v} = v^0 \hat{e}_0 + v^1 \hat{e}_1 + v^2 \hat{e}_2 + v^3 \hat{e}_3$$

then η defines a map/function which has two vector inputs & a real valued function output:

$$\eta(\vec{u}, \vec{v}) = u^0 v^0 - u^1 v^1 - u^2 v^2 - u^3 v^3$$

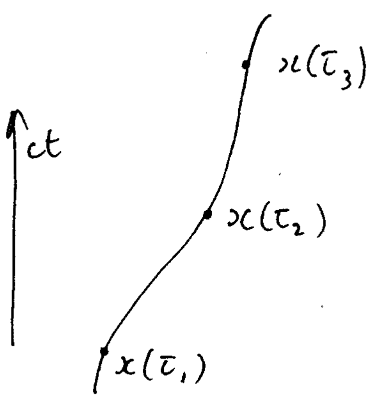
and we can connect to $\eta_{\alpha\beta}$ by noting

$$\eta(\hat{e}_\alpha, \hat{e}_\beta) = \eta_{\alpha\beta}$$

In General Relativity $\eta_{\alpha\beta}$ becomes a function of space-time coordinates & describes the "geometry of gravity."

4-vector Formalism & the Relativistic Lagrangian

We now have all the components we need for discussing geometry & coordinate transformations in relativistic systems. However, we have not yet defined how we measure 4-d velocity in spacetime.



Consider a worldline in spacetime (ignoring y, z coordinates in the diagram).

This line can be parameterized in the laboratory frame using the world time (proper time) along the curve, τ.

$$\text{ie} \quad x^\mu \equiv (ct(\tau), x(\tau), y(\tau), z(\tau))$$

We then define the 4-velocity u^μ by

$$u^\mu = \frac{dx^\mu}{d\tau} \equiv \left(c\frac{dt}{d\tau}, \frac{dx}{d\tau}, \frac{dy}{d\tau}, \frac{dz}{d\tau}\right)$$

Since we showed earlier that $dt = \dfrac{d\tau}{\sqrt{1-v^2/c^2}}$

$$\Rightarrow \quad \frac{dt}{d\tau} = \gamma \qquad \left(\text{Remember } \gamma \overset{\text{def}}{=} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\right)$$

and

$$\frac{d}{d\tau} = \gamma \frac{d}{dt}$$

thus

$$u^\mu \equiv \left(c\gamma, \gamma\frac{dx}{dt}, \gamma\frac{dy}{dt}, \gamma\frac{dz}{dt}\right)$$

$$= (c\gamma, \gamma\vec{v}) \quad \text{where} \quad \vec{v} = \frac{d\vec{x}}{dt}$$

If we apply the metric tensor to the velocity vector then

$$\begin{aligned}\eta_{\alpha\beta}u^\alpha u^\beta &= u^0u^0 - u^1u^1 - u^2u^2 - u^3u^3 \\ &= \gamma^2(c^2 - v^2) \\ &= c^2\end{aligned}$$

Since the metric tensor defines the length of vectors, we see that the square of the magnitude of the 4-velocity is c^2.

Note thus far we have written vectors with upper indices. The metric tensor actually defines an operation that lowers indices:

$$\eta_{\alpha\beta} v^\beta = v_\alpha \qquad \text{for a general vector } v^\alpha$$

this is very similarly to what we saw when the summation convention was introduced, ie

$$x_j = \delta_{ji} x_i \qquad \text{for example.}$$

Hence whenever we see a lowered index we can equivalently write $u_\alpha = \eta_{\alpha\beta}u^\beta$. Hence $\eta_{\alpha\beta}u^\alpha u^\beta \equiv u_\beta u^\beta$.

We now define the 4-momentum in the logical way. If m_0 is the rest mass of a particle the the 4-momentum is

$$p^\mu \overset{\text{def}}{\equiv} m_0 u^\mu \qquad \text{where } u^\mu = \text{4-velocity}$$

We can again look at square of this 4-vector

$$p_\mu p^\mu = \eta_{\nu\mu} p^\nu p^\mu = m_0^2 u_\mu u^\mu = m_0^2 c^2$$

Writing out components explicitly:

$$p_\mu p^\mu = m_0^2 c^2 \gamma^2 - m_0^2 \gamma^2 \vec{v}^{\,2} = m_0^2 c^2$$

if we define the 3-vector $\vec{p} = m_0 \gamma \vec{v}$ ("relativistic 3-momentum") then

$$m_0^2 c^2 + \vec{p}^{\,2} = m_0^2 \gamma^2 c^2 \stackrel{\text{def}}{=} \frac{E^2_{total}}{c^2} \quad \text{ie} \quad E_{total} \stackrel{\text{def}}{=} m_0 c^2 \gamma$$

E then corresponds to the total relativistic energy which is the sum of the kinetic plus rest mass energy.

$$\therefore T = E_{kinetic} = E_{total} - E_{rest\ mass}$$
$$= m_0 c^2 \gamma - m_0 c^2 = m_0 c^2 (\gamma - 1)$$

substituting from definition of E_{total}

$$T = \sqrt{(m_0 c^2)^2 + \vec{p}^{\,2} c^2} - m_0 c^2$$

If we let $\beta = \frac{v}{c}$ then

$$T = m_0 c^2 \left(\frac{1}{1-\beta^2} \right)^{1/2} - m_0 c^2$$

Expanding $(1-\beta^2)^{-1/2}$ as a Taylor series we get

$$T = \frac{1}{2} m_0 v^2 + \frac{3}{8} m_0 \frac{v^4}{c^2} + \dots$$

The first term is the familiar Newtonian definition of K.E., while the second is called the first relativistic correction (There is of course an infinite series of corrections).

Also note that a logical consequence of defining $\vec{p} = m_0 \gamma \vec{v}$ is that a particle with finite rest mass m_0 can never achieve a velocity of c. This is because as $v \to c$ γ becomes infinite & hence the momentum would be infinite.

What about forces? The 4-vector formulation of Newton's 2nd Law is also straightforward:

$$F^\mu = \frac{dp^\mu}{d\tau} = m_0 \frac{du^\mu}{d\tau} + \frac{dm_0}{d\tau} u^\mu$$
$$= m_0 a^\mu + \frac{dm_0}{d\tau} u^\mu$$

where a^μ is the 4-acceleration & $a^\mu = \frac{du^\mu}{d\tau}$.

The 4-vector formulation of SR is not only notationally simple, it also has some extremely useful mathematical properties. Recall that for a Lorentz transformation we defined the vector in the transformed coordinate by

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu$$

Lets look at how the 4-vector equations appear in transformed variables. For example, let's consider

$$F^\mu = \frac{dp^\mu}{d\tau} \qquad \text{—— (A)}$$

Since $\quad p^\mu = m_0 u^\mu = m_0 \frac{dx^\mu}{d\tau}$

we apply the Lorentz transformation to x^μ

ie $\quad x'^\mu = \Lambda^\mu{}_\nu x^\nu \quad$ Then

$$p'^\mu = m_0 u'^\mu = m_0 \frac{d}{d\tau}\left(\Lambda^\mu{}_\nu x^\nu\right)$$

What happens to $d/d\tau$? This is the <u>same</u> in all coordinate systems. This follows immediately because we defined the world time (proper time) in terms of the line element ds which is preserved by Lorentz transformations.

The velocity connecting the two coordinate frames is a constant. (<u>note</u> don't confuse this with the velocity of the particle we are measuring). Thus

$$\frac{d}{d\tau}\left(\Lambda^\mu{}_\nu x^\nu\right) = \Lambda^\mu{}_\nu \frac{dx^\nu}{d\tau}$$

since the $\Lambda^\mu{}_\nu$ are only functions of the velocity connecting the two frames, or they are constant numbers.

$$\therefore \quad p'^\mu = m_0 u'^\mu = \Lambda^\mu{}_\nu m_0 u^\nu = \Lambda^\mu{}_\nu p^\nu$$

Thus for (A) we have in transformed coordinates

$$F'^\mu = \frac{d}{d\tau}\left(\Lambda^\mu{}_\nu p^\nu\right) = \Lambda^\mu{}_\nu \frac{dp^\nu}{d\tau} = \Lambda^\mu{}_\nu F^\nu$$

thus if we begin from

$$F^\mu = \frac{dp^\mu}{d\tau}$$

$$\Rightarrow \quad \Lambda^\nu{}_\mu F^\mu = \Lambda^\nu{}_\mu \frac{dp^\mu}{d\tau}$$

and thus

$$F'^{\gamma} = \frac{dp'^{\gamma}}{d\tau}$$

hence we say that the equation is invariant under the Lorentz transformation.

We now look at how we would form a relativistic Lagrangian. In the non-relativistic case we were able to derive the E-L equations by considering either Hamilton's Principle or by using D'Alembert's Principle. There is no analog of D'Alembert's Principle for S.R. though.
Notice that using a variational principle is also somewhat troubling. Recall that we considered the action integral

$$I = \int_{t_1}^{t_2} L \, dt,$$

in S.R. this doesn't seem natural since t will be different for different observers. However, we can still try deriving equations of motion from the E-L equations & see what we get. If

$$L = -m_0 c^2 \sqrt{1 - \frac{v^2}{c^2}} - V$$

where v is the speed of the particle & V is a potential term, then the E-L equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v^i}\right) - \frac{\partial L}{\partial x^i} = 0 \qquad (v^i = \dot{x}^i)$$

gives $$\frac{d}{dt}\frac{mv^i}{\sqrt{1-\frac{v^2}{c^2}}} = -\frac{\partial V}{\partial x^i}$$

which is actually a correct S.R. expression & can be used to solve many problems. However, note that (1) The Lagrangian is no longer $T-V$ and (2) this is not a covariant expression. If we change our frame of reference then the velocity v will change.

Hence let us look for a Lagrangian that is invariant under Lorentz transformations.

We noted that using t is the integration variable in a variational principle does not seem natural (it is now one of the coordinate in spacetime). In fact, the leap to using the proper time is the integration variable is the correct way to proceed.

In this case the E-L equations will have a form
$$\frac{\partial L}{\partial x^\mu} - \frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{x}^\mu}\right) = 0 \quad \text{where } \dot{x}^\mu = \frac{dx^\mu}{d\tau}$$

Hence we wish to find an L that produces conservation of 4-momentum for a free particle.

Consider $$L = \frac{1}{2} m_0 u_\mu u^\mu = \frac{1}{2} m_0 \eta_{\mu\nu} u^\mu u^\nu$$

Applying the E-L equations, we first calculate derivatives,

$$\frac{\partial L}{\partial x^\nu} = \frac{\partial}{\partial x^\nu}\left(q u^\mu A_\mu\right)$$

$$\frac{\partial L}{\partial \dot{x}^\nu} = m_0 u_\nu + q A_\nu$$

then substitute into the E-L equations:

$$\frac{\partial}{\partial x^\nu}\left(q u^\mu A_\mu\right) - \frac{d}{d\tau}\left(m_0 u_\nu\right) - q\frac{dA_\nu}{d\tau} = 0$$

$$\Rightarrow \qquad \frac{dp_\nu}{d\tau} = -q\frac{dA_\nu}{d\tau} + \frac{\partial}{\partial x^\nu}\left(q u^\mu A_\mu\right)$$

This ends our brief sojourn into relativistic Lagrangians!