Classical Mechanics & Relativity

We're going to spend the next ~ 2 weeks building up all the pieces we need for doing calculations in particle physics.

In particular, we'll ~~~~ discuss

- Relativity (since the universe is fundamentally relativistic, & particle physics involves relativistic particles)

- Quantum Mechanical Perturbation Theory
  → ~~the~~ "Fermi's Golden Rule"
  (how to calculate what you see in a particle physics experiment)

First, we're going to quickly review the Lagrangian & Hamiltonian formulations of classical mechanics

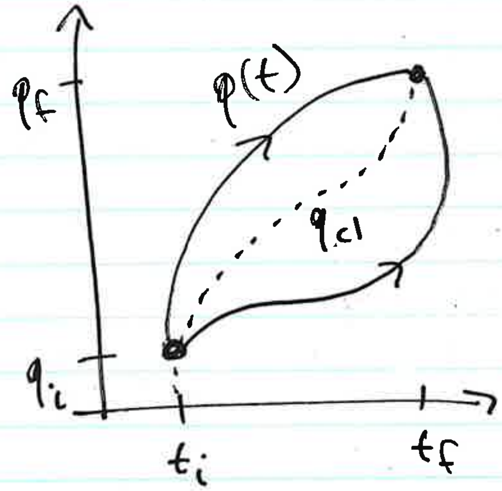
Lagrangian: the starting point of ~~~~ perturbative QFTs, & the language of particle physics.

Hamiltonian: the natural path to Quantum Mechanics.

Lagrangian Mechanics

We wish to ~~describe~~ study a physical system with one "degree of freedom", described by a "generalized coordinate" $q(t)$.

Suppose the particle (dof) is at q_i at $t = t_i$ and q_f at $t = t_f$.



In principle, there are many different "paths" or "trajectories" the particle could take.

How do we determine the actual trajectory? (the "classical trajectory", q_{cl})

→ Principle of least action

Lagrangian: function of the generalized coordinate and the corresponding "generalized velocity":

$$L \equiv L[q(t), \dot{q}(t)]$$

From here, we define the action

$$S = \int_{t_1}^{t_2} L[q(t), \dot{q}(t)]\, dt$$

principle of least action: the classical trajectory is the one that minimizes (or more generally, extremizes) the action.

i.e. we want the trajectory that yields: $\delta S = 0$

More explicitly, consider arbitrary variations of the trajectory, $\delta q(t)$ that leave the endpoints fixed:

$$\delta q_i = \delta q(t_i) = 0, \quad \delta q_f = \delta q(t_f) = 0$$

Then:

$$0 = \delta S = \int_{t_1}^{t_2} \delta L[q(t), \dot{q}(t)]\, dt$$

(chain rule)

$$= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} \right] dt$$

Now, $\delta\dot{q} = \delta\left(\frac{dq}{dt}\right) = \frac{d}{dt}(\delta q)$

Integrating by parts on the second term, we find:

$$0 = \delta S = \underbrace{\frac{\partial L}{\partial \dot{q}}\delta q\Big|_{t_i}^{t_f}}_{\text{vanishes since } \delta q_i = \delta q_f = 0} + \int_{t_i}^{t_f}\left[\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right]\underbrace{\delta q}_{\text{arbitrary}}\,dt$$

$\Rightarrow$ the solution must satisfy:

$$\left\{\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right\}_{q_{cl}(t)} = 0 \qquad t_i \le t \le t_f$$

This is the Euler-Lagrange equation that determines the classical trajectory.

In many physical systems, the Lagrangian is written as the difference of the kinetic & potential energies:

$$L = T - V, \quad T = \tfrac{1}{2}m\dot{q}^2, \quad V \equiv V(q)$$

Plugging this into the E-L equations:

$$\frac{\partial L}{\partial q} = \frac{\partial V}{\partial q} \quad , \quad \frac{\partial L}{\partial \dot{q}} = \frac{\partial T}{\partial \dot{q}} = m\dot{q}$$

$$\Rightarrow \delta S = 0 \iff m\ddot{q} = -\frac{\partial V}{\partial q}$$

This is just Newton's second law, so these formulations of classical mechanics are equivalent.

The Lagrangian formalism is powerful because it generalizes easily to systems with multiple (or infinite!) degrees of freedom, with multiple (infinite) generalized coordinates.

This is eventually what we'll want for <u>fields</u>.

Hamiltonian Formulation

In the Lagrangian formalism, we used q and $\dot{q}$ as independent variables.

Let's define the associated "generalized momentum"

$$p = \frac{\partial L}{\partial \dot{q}}$$

For systems w/ $L = T - U$, $T = \frac{1}{2} m \dot{q}^2$:

$$p = m\dot{q} \qquad (\text{"mass times velocity"})$$

Hamiltonian formalism: take q, p as the dynamical variables.

Instead of $L[q, \dot{q}]$ we'll define a Hamiltonian:

$H = H[q, p]$.

The Hamiltonian is obtained from the Lagrangian via a "Legendre transformation":

$$H(q, p) = \dot{q} p - L[q, \dot{q}]$$

where it is understood that $\dot{q}$ on the RHS is written in terms of q & p.

Now, treating q, p as independent variables:
consider

$$\frac{\partial H}{\partial p} = \dot{q} + \cancel{p\frac{\partial \dot{q}}{\partial p}} - \cancel{\underbrace{\frac{\partial L}{\partial \dot{q}}}_{= p}\frac{\partial \dot{q}}{\partial p}}$$

$$= \dot{q}.$$

and:

$$\frac{\partial H}{\partial q} = \cancel{p \cdot \frac{\partial \dot{q}}{\partial q}} - \frac{\partial L}{\partial q} - \cancel{\underbrace{\frac{\partial L}{\partial \dot{q}}}_{=p} \cdot \frac{\partial \dot{q}}{\partial q}}$$

$$= -\frac{\partial L}{\partial q} \quad \text{(E-L eqn.)}$$

$$= -\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) = -\dot{p}$$

So we arrive at Hamilton's equations:

$$\frac{\partial H}{\partial p} = \dot{q} \quad , \quad -\frac{\partial H}{\partial q} = \dot{p}$$

→ swapped E-L equations for two first order equations.

For Lagrangians of the form $L = T - V$, $T = \frac{1}{2} m \dot{q}^2$ and $V = V(q)$, we have:

$$p = \frac{\partial L}{\partial \dot{q}} = m\dot{q} \Rightarrow \dot{q} = p/m$$

Then,

$$\begin{aligned} H &= p\dot{q} - L(q, \dot{q}) \\ &= p\dot{q} - \left[\frac{1}{2} m \dot{q}^2 - V(q)\right] \\ &= p\left(\frac{p}{m}\right) - \frac{1}{2} m \left(\frac{p}{m}\right)^2 + V(q) \end{aligned}$$

or,

$$H = \frac{p^2}{2m} + V(q) = T + V$$

→ the Hamiltonian is associated w/ the mechanical energy of the system.

One can go further and define Poisson brackets, etc., as a route to ~~the~~ Hamilton's equations. See e.g., Landau & Lifschitz for a very concise treatment.

Before we switch to relativity:

Warm-up: Rotations & Euclidean space

Intuition: a rotation of an ordinary vector leaves its length invariant.

A rotation in 3-dimensions can be represented by a 3×3 matrix R, acting on an arbitrary vector, $\vec{x}$, which transforms (rotates) to the vector $\vec{x}'$:

$$\vec{x}' = R\vec{x}$$

(invariant!)

That its length is preserved is the statement:

$$\vec{x}' \cdot \vec{x}' = (R\vec{x}) \cdot (R\vec{x}) = \vec{x} \cdot \vec{x}$$

We can write this in different ways, e.g. as a column vector multiplying a row vector:

$$\vec{x} \cdot \vec{y} \equiv \vec{x}^T \vec{y} = \vec{x}^T (\mathbb{1}) \vec{y}$$

(T superscript denotes transpose)

w/ $\mathbb{1} = \text{diag}(1,1,1)$ is the identity matrix.

Thus, we can write

$$\vec{x}' \cdot \vec{x}' = (R\vec{x})^T \mathbb{1} (R\vec{x})$$

$$= \vec{x}^T R^T \mathbb{1} R \vec{x} = \vec{x}^T \mathbb{1} \vec{x}$$

This is true if the matrix R is <u>orthogonal</u>

$$R^T \mathbb{1} R = R^T R = \mathbb{1}.$$

We can also write these expressions in terms of the components x_i of the vector, and R_{ij} of the matrices, with $i = 1, 2, 3$.

For example,

$$\sum_k x_k' x_k' = \sum_i \sum_j \sum_k x_i (R^T)_{ik} R_{kj} x_j$$

$$= \sum_i \sum_j x_i \delta_{ij} x_j \qquad \left(\text{since} \quad \sum_k (R^T)_{ik} R_{kj} = \delta_{ij} \right.$$

$$= \sum_i x_i x_i \qquad \left(\sum_j A_j \delta_{ij} = A_i \right)$$

I will frequently use <u>Einstein summation convention</u> in which the Σ symbol is dropped, and <u>repeated</u> indices are understood to be summed.

In these expressions we have emphasized the (usually implicit) δ_{ij} — it lets us take two (of the same, in this case) vectors and returns a distance.

The Kronecker δ-function is the metric tensor for Euclidean space.

Lorentz transformations & Minkowski spacetime

In special relativity, we treat space and time on a (more) equal footing, and work in 4-dimensional spacetime or Minkowski space

Define a 4-vector that groups spatial/time coords:

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, \vec{x})$$

Greek indices (from the middle of the alphabet) $\mu, \nu, \ldots$ take values 0, 1, 2, 3.

We'll try to reserve Latin indices $(i, j, \ldots)$ to take values 1, 2, 3 for the spatial components.

A point x^μ defines an "event" in spacetime.

In analogy to our discussion of rotations, we can define a scalar product of four-vectors x^μ, y^μ

$$x \cdot y \equiv x^T g\, y$$

$$= \sum_{\mu,\nu} x^\mu g_{\mu\nu} y^\nu = g_{\mu\nu} x^\mu y^\nu$$

Where I have introduced the Minkowski Metric

$$g_{\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

(This is the "mostly minus" or "West coast" or "particle physics" (as opposed to relativists) convention)

Writing things out term-by-term,

$$x \cdot y = x^0 y^0 - (x^1 y^1 + x^2 y^2 + x^3 y^3)$$

$$= x^0 y^0 - \vec{x} \cdot \vec{y}$$

In contrast to Euclidean space, it's useful in Minkowski space to distinguish between "contravariant" vectors with upper indices and "covariant" vectors with lower indices.

We can always go from one to another by contracting with the metric:

$$x_\mu \equiv g_{\mu\nu} x^\nu$$

Component-wise, $x_\mu = (x^0, -x^1, -x^2, -x^3)$

The metric $g_{\mu\nu}$ has an <u>inverse</u> $g^{\mu\nu}$ which satisfies

$$g\, g^{-1} = \mathbb{1}$$

or,

$$g_{\mu\nu}\, g^{\nu\rho} = \delta_\mu^\rho$$

In flat space of course, the components are the same:

$$g^{\mu\nu} = \mathrm{diag}(1,-1,-1,-1)$$

We then have a bunch of relations like:

$$x \cdot y = g_{\mu\nu} x^\mu y^\nu = x_\mu y^\mu = x^\mu y_\mu = g^{\mu\nu} x_\mu y_\nu$$

Now we can define more general Lorentz tensors, with a mix of upper & lower indices.

For instance, we might have an object $T^{\mu\nu}{}_{\rho\sigma}$ which satisfies

$$T^{\mu\nu\lambda}{}_{\sigma} = g^{\lambda\rho} T^{\mu\nu}{}_{\rho\sigma}$$

note the order here, as well as position matters!

(e.g., don't write $T^{\mu\nu\lambda}_{\sigma}$ unless it has some special symmetry properties)

It's often useful in manipulations to use our freedom to relabel dummy indices — if an index is summed over all its components, it doesn't matter what we call it.

As an example, consider a symmetric tensor $S_{\mu\nu}$ satisfying $S_{\mu\nu} = + S_{\nu\mu}$
and an anti-symmetric tensor $A^{\mu\nu} = -A^{\nu\mu}$.

Then,

$$S_{\mu\nu} A^{\mu\nu} = \frac{1}{2} S_{\mu\nu} A^{\mu\nu} + \frac{1}{2} S_{\mu\nu} A^{\mu\nu}$$

$$= \frac{1}{2} S_{\mu\nu} A^{\mu\nu} + \frac{1}{2} (S_{\nu\mu})(-A^{\nu\mu})$$

(using symm./anti-symm.)

$$= \frac{1}{2} S_{\mu\nu} A^{\mu\nu} - \frac{1}{2} S_{\mu\nu} A^{\mu\nu}$$

(just relabelling $\mu \leftrightarrow \nu$)

$$= 0 \ !$$

We'll perform such manipulations repeatedly in the context of relativistic QFT.

Just as rotations preserve the length of a Euclidean vector, we can define <u>Lorentz transformations</u> as "rotations" in Minkowski spacetime, which preserve the scalar product.

In matrix form:

$$x' = \Lambda x$$

or in components,

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

note the index structure <u>here!</u>

These are defined to leave the scalar product of four-vectors invariant:

$$x' \cdot y' = (x')^T g\, y' = x^T \Lambda^T g \Lambda y$$
$$= x^T g\, y$$

the <u>metric</u>

The matrix Λ must then satisfy:

$$\Lambda^T g \Lambda = g$$

This is the analogue of the orthogonality relation, $R^T \mathbb{1} R = \mathbb{1}$.

Note that the orthoognality condition implies that g is invariant under ~~metric~~ Lorentz transformations. In components:

$$\Lambda^T g \Lambda = g \Rightarrow (\Lambda^T)_\rho{}^\mu g_{\mu\nu} \Lambda^\nu{}_\sigma = g_{\rho\sigma}$$

or, $$g_{\mu\nu} \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma = g_{\rho\sigma} \qquad (*)$$

Note that we've used $\Lambda^\mu{}_\rho = (\Lambda^T)_\rho{}^\mu$ and reversed the index structure.

This is done so that matrix multiplication follows naturally from adjacent up/down indices.

Any Lorentz transformation Λ has an <u>inverse</u> transformation, Λ^{-1}.

A convenient way to derive its form is to multiply $(*)$ above by the inverse metric:

$$\underbrace{g^{\lambda\rho} [g_{\mu\nu} \Lambda^\mu{}_\rho}\, \Lambda^\nu{}_\sigma] = g^{\lambda\rho} g_{\rho\sigma} = \delta^\lambda{}_\sigma$$

this quantity $\times \Lambda^\nu{}_\sigma$ is the identity, so by definition it is the <u>inverse matrix</u>:

$$(\Lambda^{-1})^\lambda{}_\nu \Lambda^\nu{}_\sigma = \delta^\lambda{}_\sigma$$

where

$$(\Lambda^{-1})^{\lambda}{}_{\nu} \equiv \Lambda_{\nu}{}^{\lambda} = g_{\nu\mu} g^{\lambda\rho} \Lambda^{\mu}{}_{\rho}$$

used symmetry of the metric tensor.

Here I've introduced a convenient shorthand, where Λ with the "wrong" index structure in components represents the components of the inverse metric.

Note that Λ^{-1} has the same structure as Λ^T

Note that a (covariant) vector ~~transforms~~ w/ a lower index transforms by the action of Λ^{-1}, rather than Λ:

$$\begin{aligned} x'_{\nu} &= g_{\nu\mu} x'^{\mu} = g_{\nu\mu} (\Lambda^{\mu}{}_{\rho} x^{\rho}) \\ &= g_{\nu\mu} \Lambda^{\mu}{}_{\rho} g^{\rho\lambda} x_{\lambda} = [g_{\nu\mu} \Lambda^{\mu}{}_{\rho} g^{\rho\lambda}] x_{\lambda} \\ &= \Lambda_{\nu}{}^{\lambda} x_{\lambda} = x_{\lambda} (\Lambda^{-1})^{\lambda}{}_{\nu} \end{aligned}$$

the reversed index structure appears naturally here.

The inverse metric can be used to write the unprimed coordinate vectors in terms of the primed ones (that's the point of an inverse):

$$x^\mu = (\Lambda^{-1})^\mu{}_\nu \, x'^\nu$$

$$x_\mu = x_\nu \Lambda^\nu{}_\mu$$

Now we'll consider the explicit form of some particular transformations:

Spatial rotations:

An obvious subgroup of the set of Lorentz transformations is the set of rotations of the three spatial coordinates.

E.g., a rotation about the $z = x^3$ axis mixes the $x^1 = x$ and $x^2 = y$ coordinates, so it has the form:

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \Lambda(R) \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & b & 0 \\ 0 & c & d & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

To derive a more explicit form, use the orthogonality relation: $\Lambda^T g \Lambda = g$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & a & c & 0 \\ 0 & b & d & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & b & 0 \\ 0 & c & d & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -(a^2+c^2) & -(ab+cd) & 0 \\ 0 & -(ab+cd) & -(b^2+d^2) & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \underset{\text{demand}}{\overset{\uparrow}{=}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

So we have the requirements:

$$a^2 + c^2 = 1, \quad b^2 + d^2 = 1, \quad ab + cd = 0$$

These are easily satisfied by choosing

$$a = \cos\theta, \quad b = -\sin\theta, \quad c = \sin\theta, \quad d = \cos\theta$$

So the Lorentz transformation for this particular spatial rotation has the form:

$$\Lambda^\mu{}_\nu(R(\theta)) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$