Massachusetts Institute of Technology
Department of Physics
8.033 Fall 2024

Lecture 16
The calculus of variations, the principle of maximum aging, and the motion of bodies in spacetime

16.1 Our goal

In Lecture 17, we will introduce a hypothesis which will allow us to formulate how a body moves in any spacetime, not just the spacetimes of special relativity. The key idea is to use the principle we recently discussed which argued that unaccelerated motion means bodies move along trajectories of "maximal aging." Any acceleration slows down their own clocks such that the proper time they accumulate during their motion is less than it would be on the non-accelerated path. The goal of today's lecture is to develop tools that allow us to use this principle.

16.2 Euler's equation

Imagine a function f that depends on time, on a position variable x, and on the derivative $\dot{x} \equiv dx/dt$. Suppose that f encodes something important about our physical situation. Suppose that at time t_i we are at position x_i, and that at time t_f we are at position x_f. Subject to the boundary condition that we must start at the event (t_i, x_i) and we must end at (t_f, x_f), we are free to take any trajectory $x(t)$ that connects these two events.

Imagine that the trajectory we actually take is the one that gives us the *extremum* of

$$J = \int_{t_i}^{t_f} f\left[x(t), \dot{x}(t); t\right]\, dt\ . \tag{16.1}$$

For example, f might tell us about the rate at which we age along the trajectory, and J could be the accumulated aging we experience. Of the infinite number of ways that we can connect (t_i, x_i) to (t_f, x_f), how do we find that one that extremizes J?

To proceed, we imagine that there exists some $x_e(t)$ which gives us this extremum. We do not know $x_e(t)$, so our current guess deviates from this correct choice. We parameterize how our current guess deviates from the correct trajectory as follows:

$$x(t) \equiv x(t;\alpha) = x_e(t) + \alpha A(t)\ . \tag{16.2}$$

The function $A(t)$ is totally arbitrary, except that we require it to vanish at the endpoints: $A(t_i) = A(t_f) = 0$; otherwise, our trajectory would not meet the boundary condition. The parameter α allows us to control how the variation $A(t)$ enters into our path $x(t;\alpha)$.

Our basic idea is to ask how the integral J behaves when we are in the vicinity of the extremum. We know that ordinary functions are flat — they have zero first derivative — when we are at an extremum. Let us put

$$J(\alpha) = \int_{t_i}^{t_f} f\left[x(t;\alpha), \dot{x}(t;\alpha); t\right]\, dt\ . \tag{16.3}$$

We've now made the integral J a function of the parameter α. We know that $\alpha = 0$ corresponds to the extremum of J by its definition. However, this isn't useful for us, since we don't know what $x(t)$ this corresponds to. However, because $\alpha = 0$ corresponds to an extremum, we also know that $(\partial J/\partial\alpha)_{\alpha=0} = 0$; in essence, we're taking advantage of the fact that the *shape* of $J(\alpha)$ has a particular form as we approach this extremum.

Let's take a look at the derivative of J with respect to α:

$$\frac{\partial J}{\partial \alpha} = \int_{t_i}^{t_f} \left[\frac{\partial f}{\partial x}\frac{\partial x}{\partial \alpha} + \frac{\partial f}{\partial \dot{x}}\frac{\partial \dot{x}}{\partial \alpha}\right] dt \; . \tag{16.4}$$

Using Eq. (16.2), we have

$$\frac{\partial x}{\partial \alpha} = A(t) \; , \qquad \frac{\partial \dot{x}}{\partial \alpha} = \frac{dA}{dt} \; . \tag{16.5}$$

Plugging this in, we have

$$\frac{\partial J}{\partial \alpha} = \int_{t_i}^{t_f} \left[\frac{\partial f}{\partial x}A(t) + \frac{\partial f}{\partial \dot{x}}\frac{dA}{dt}\right] dt \; . \tag{16.6}$$

The last term on the right-hand side of (16.6) can be rearranged in a really useful way using integration by parts:

$$\begin{aligned}\int_{t_i}^{t_f} \frac{\partial f}{\partial \dot{x}}\frac{dA}{dt}\, dt &= A(t)\frac{\partial f}{\partial \dot{x}}\bigg|_{t_i}^{t_f} - \int_{t_i}^{t_f} A(t)\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}}\right) dt \\ &= -\int_{t_i}^{t_f} A(t)\frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}}\right) dt \; . \end{aligned} \tag{16.7}$$

To get the final expression, we used the fact that $A(t_i) = A(t_f) = 0$. Using this we have

$$\frac{\partial J}{\partial \alpha} = \int_{t_i}^{t_f} A(t)\left[\frac{\partial f}{\partial x} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}}\right)\right] dt = 0 \quad \text{at an extremum of } J. \tag{16.8}$$

The function $A(t)$ is totally arbitrary, aside from the boundary condition that it vanish at t_i and t_f. We require $\partial J/\partial\alpha = 0$ for all $A(t)$; for this to occur, the quantity inside square brackets must vanish:

$$\frac{\partial f}{\partial x} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}}\right) = 0 \; . \tag{16.9}$$

This is known as *Euler's equation*, and was first derived by the Swiss polymath Leonhard Euler. Properly applied, it yields a differential equation which allows us to find the trajectory $x(t)$ which extremizes the integral J.

For simplicity, we did this for a function of just one variable. However, we could have imagined a trajectory in all three spatial directions. With a little more effort, it's not too hard to show that more general version of Eq. (16.9) is just the trio of equations

$$\begin{aligned}\frac{\partial f}{\partial x} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}}\right) &= 0 \; , \\ \frac{\partial f}{\partial y} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{y}}\right) &= 0 \; , \\ \frac{\partial f}{\partial z} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{z}}\right) &= 0 \; . \end{aligned} \tag{16.10}$$

Those of you who have studied Lagrangian mechanics (which we discuss briefly at the end of these notes) presumably have already encountered equations of this form.

16.3 An example: The brachistochrone ("shortest time")

A bead starts from rest at $(x_i, y_i) = (0, 0)$ and slides without friction down a wire, reaching (x_f, y_f). What shape should the wire have in order for the bead to reach (x_f, y_f) in as little time as possible?

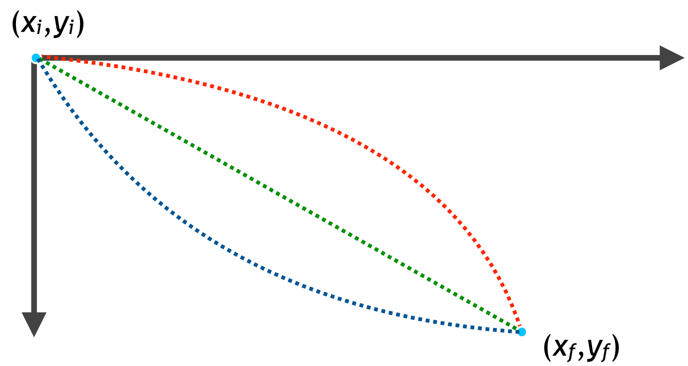

Figure 1: Three plausible paths for the brachistochrone connecting $(0, 0)$ and (x_f, y_f).

To figure this out, apply the Euler equation to minimize the total travel time of the bead as it slides along the wire. Think of the integral we are minimizing as T, defined by

$$T = \int_{\text{initial position}}^{\text{final position}} \frac{ds}{v} , \tag{16.11}$$

where ds is the differential of path length along the wire, and v is its speed. For the path length, note that the bead moves in both x and y, so

$$ds = \sqrt{dx^2 + dy^2} = dy\sqrt{1 + (x')^2} , \qquad \text{where} \quad x' \equiv \frac{dx}{dy} . \tag{16.12}$$

For the speed v, since the bead starts from rest, it only gets speed from falling a distance y:

$$\frac{1}{2}mv^2 = mgy \qquad \longrightarrow \qquad v = \sqrt{2gy} . \tag{16.13}$$

The equation we wish to minimize is thus

$$T = \int_0^{y_f} \sqrt{\frac{1 + (x')^2}{2gy}} dy . \tag{16.14}$$

This is perfectly set up for us to apply the Euler equation provided we make a few adjustments: we put

$$f = \sqrt{\frac{1 + (x')^2}{2gy}} ; \tag{16.15}$$

we change the integration variable from t to y, and replace $\dot{x}$ with x'. Our slightly tweaked Euler equation is thus

$$\frac{\partial f}{\partial x} - \frac{d}{dy}\left(\frac{\partial f}{\partial x'}\right) = 0 . \tag{16.16}$$

Let's evaluate these terms:

$$\frac{\partial f}{\partial x} = 0 \,, \qquad \frac{\partial f}{\partial x'} = \left(\frac{1}{\sqrt{2gy}}\right)\left(\frac{x'}{\sqrt{1+(x')^2}}\right) . \tag{16.17}$$

Plugging these into the Euler equation yields

$$\frac{d}{dy}\left(\frac{1}{\sqrt{2gy}}\right)\left(\frac{x'}{\sqrt{1+(x')^2}}\right) = 0 . \tag{16.18}$$

We can immediately integrate this up to find

$$\left(\frac{1}{\sqrt{2gy}}\right)\left(\frac{x'}{\sqrt{1+(x')^2}}\right) = \text{constant} . \tag{16.19}$$

Let's set this constant to $1/\sqrt{4gA}$, where A is another constant[1] with the dimensions of length. Squaring both sides of Eq. (16.19), we find

$$\frac{(x')^2}{2gy\,(1+(x')^2)} = \frac{1}{4gA} \,, \tag{16.20}$$

which we can manipulate into

$$\begin{aligned}\left(\frac{dx}{dy}\right)^2 &= \frac{y/(2A)}{1-y/(2A)} \\ &= \frac{y^2}{2Ay-y^2} .\end{aligned} \tag{16.21}$$

We thus at last have our equation governing x as a function of y:

$$x(y) = \int_0^y \frac{y\,dy}{\sqrt{2Ay-y^2}} . \tag{16.22}$$

To wrap this up, we change variables: define $y = A(1-\cos\theta)$, $dy = A\sin\theta\,d\theta$. It's not too hard to show that $2Ay - y^2 = A^2\sin^2\theta$; our equation for x becomes

$$x = \int_0^\theta A(1-\cos\theta)\,d\theta = A(\theta-\sin\theta) . \tag{16.23}$$

The full solution for the brachistotrone is then given by

$$\begin{aligned}x &= A(\theta-\sin\theta) \\ y &= A(1-\cos\theta) .\end{aligned} \tag{16.24}$$

The bead's motion goes over the range $0 \le \theta \le \theta_{\rm max}$; both the constant A and the angle $\theta_{\rm max}$ can be found by solving $x(\theta_{\rm max}) = x_f$, $y(\theta_{\rm max}) = y_f$.

[1]This is one of those places where I get to invoke instructors' privilege and cheat a little bit. If you were doing this problem by yourself, you'd probably just set the right hand side to something like C, and hope that C's role is explained later in the calculation. Doing so, you would eventually find that C shows up as $1/4gC^2$ in the analysis. Since I've already done the calculation, I'm using the fact that I know this in advance to streamline things here.

16.4 Maximal aging in special relativity

Let's use the calculus of variations to see what kind of motion results in "maximal aging" on an observer's trajectory in special relativity. We start with the fact that, for an observer moving on a timelike trajectory,

$$d\tau^2 = dt^2 - \left(dx^2 + dy^2 + dz^2\right)/c^2 \; . \tag{16.25}$$

For simplicity, let's restrict ourselves to one spatial dimension for now, setting $dy = dz = 0$.

Now think about the many paths which can connect event A to event B. Our goal is to compute the accumulated τ along those paths:

$$\tau_{A\to B} = \int_A^B dt\sqrt{1 - \frac{1}{c^2}\left(\frac{dx}{dt}\right)^2} \equiv \int_A^B dt\sqrt{1 - \frac{\dot{x}^2}{c^2}} \; . \tag{16.26}$$

We've introduced $\dot{x} = dx/dt$ for notational convenience. Notice that the integrand looks like $dt/\gamma(\dot{x})$ — a form that hopefully makes a lot of sense.

Let's now think about how to extremize $\tau_{A\to B}$ by putting $J \to \tau$, and setting $f(x, \dot{x}) = \sqrt{1 - \dot{x}^2/c^2}$. Re-stating Euler's equation,

$$\frac{\partial f}{\partial x} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}}\right) = 0 \; , \tag{16.27}$$

with

$$\frac{\partial f}{\partial x} = 0 \; , \qquad \frac{\partial f}{\partial \dot{x}} = -\frac{\dot{x}/c^2}{\sqrt{1 - \dot{x}^2/c^2}} \; . \tag{16.28}$$

Plugging these into (16.27), we find

$$\frac{d}{dt}\left(\frac{\dot{x}}{\sqrt{1 - \dot{x}^2/c^2}}\right) = 0 \tag{16.29}$$

whose solution is

$$\dot{x} = \text{constant} \; . \tag{16.30}$$

If you include y and z in your analysis, you'll likewise conclude that $\dot{y}$ and $\dot{z}$ must be constants in order to follow the trajectory that maximizes the accumulated proper time from A to B. A trajectory with $\dot{x}$, $\dot{y}$, and $\dot{z}$ all constants is nothing more than an inertial trajectory in spacetime. **The unaccelerated trajectory is the one which maximizes an observer's accumulated proper time as they move through spacetime: It is the trajectory of maximal aging.**

Important side issue: strictly speaking, the calculation we just did tells us that the unaccelerated trajectory represents an *extremum* of accumulated proper time, which can be either a maximum or a minimum. How do we know this extremum is a maximum and not a minimum? In this particular case, it is because we know that the minimum aging trajectory is the trajectory along which $\tau_{A\to B} = 0$, and in fact that there are an infinite number of such trajectories, all with crazy — essentially unphysical — accelerations. In general, knowing whether the extremum is a minimum or a maximum requires you to think a little bit about the physics of your situation. You will find that the outcome of the Euler equation calculation picks a unique extremum; the opposite extremum tends to not be uniquely specified.

16.5 Lagrangian mechanics and relativity

Independent of whether you continue to study relativity into the future or not, the calculus of variations and the Euler equations are likely to be important for you as long as you remain a physics student. The reason is that ordinary mechanics can be formulated in a way that uses these tools. The basic idea works as follows:

- Suppose a body moves from event (t_i, x_i, y_i, z_i) to event (t_f, x_f, y_f, z_f).
- Consider *every possible trajectory* that connects these events. For every point along those trajectories, compute the body's kinetic energy K and its potential energy U.
- Define the *Lagrangian* L as the difference in these quantities: $L \equiv K - U$.
- Define the *action* S as the time integral of L: $S \equiv \int_{t_i}^{t_f} L\,dt$.

A remarkable result, which is discussed in great detail in the IAP course 8.223 and the advanced mechanics course 8.09, is that Newtonian mechanics is equivalent to the path of *least action*, and can be found by applying the Euler equations (often called the Euler-Lagrange equations in this context) to L:

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = 0\ , \quad \frac{\partial L}{\partial y} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right) = 0\ , \quad \frac{\partial L}{\partial z} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{z}}\right) = 0\ . \tag{16.31}$$

These equations work in other coordinate systems too — you can replace the Cartesian set (x, y, z) with cylindrical coordinates (r, ϕ, z) or spherical ones (r, θ, ϕ) or really bizarre ones that just happen to be adapted to the geometry of your problem.

A Lagrangian formulation of mechanics is often much easier to work with than the $\mathbf{F} = m\mathbf{a}$ based techniques you learned in 8.01/8.012, particularly if the problem is subject to constraints. What makes them particularly nice to work with is that ultimately one need only compute a single scalar quantity, L, rather than work with vector-valued forces or torques. It is also worth noting that the classical action is intimately connected to the phase of a quantum wavefunction. The trajectory of "least action" is closely related to the phase corresponding to the most likely outcome of a quantum process in the classical limit.

In another few lectures, we will start working with general spacetime metrics, for which $ds^2 = g_{\alpha\beta}dx^\alpha dx^\beta$, and for which the metric $g_{\alpha\beta}$ will be a function of the different coordinates. However, it will remain the case that for a timelike observer, $c^2 d\tau^2 = -g_{\alpha\beta}dx^\alpha dx^\beta$. Let's use this to define the Lagrangian-like quantity that we will want to use to describe motion in general spacetimes. Consider motion that begins at event A and ends at event B. The proper time accumulated along a trajectory between these events is

$$c\Delta\tau = \int_A^B (f)^{1/2} d\tau\ , \tag{16.32}$$

where

$$f = -g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} \equiv -g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu\ . \tag{16.33}$$

Requiring that the trajectory through spacetime between events A and B be an extremum leads us to the following 4 Euler equations (one for each value of the index α):

$$\frac{\partial(f)^{1/2}}{\partial x^\alpha} - \frac{d}{d\tau}\left(\frac{\partial(f)^{1/2}}{\partial \dot{x}^\alpha}\right) = 0\ . \tag{16.34}$$

This can be simplified a bit more. First, note that

$$\frac{\partial(f)^{1/2}}{\partial x^\alpha} = \frac{1}{2(f)^{1/2}}\frac{\partial f}{\partial x^\alpha} , \qquad \frac{\partial(f)^{1/2}}{\partial \dot{x}^\alpha} = \frac{1}{2(f)^{1/2}}\frac{\partial f}{\partial \dot{x}^\alpha} . \tag{16.35}$$

Second, note that $df/d\tau = 0$: f is nothing more than $g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = \vec{u}\cdot\vec{u} = -c^2$, and the *total* derivative of this quantity with respect to proper time is zero. (Its *partial* derivatives are not zero: if we vary a particular value of x^α or a particular value of $\dot{x}^\alpha$ while holding all other quantities constant, we push f away from the value it "should" have.) This allows us out to clear out an overall factor of $2(f)^{1/2}$, and the Euler equations (16.34) then become

$$\frac{\partial f}{\partial x^\alpha} - \frac{d}{d\tau}\left(\frac{\partial f}{\partial \dot{x}^\alpha}\right) = 0 . \tag{16.36}$$

This tells us that $f = -g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$ plays a role in relativistic mechanics exactly like the Lagrangian of ordinary classical mechanics. It is traditional to multiply this by a factor of $-1/2$ — after all, the extremum of $-1/2$ times a function occurs at the same place as the extremum of that function. We then define the relativistic Lagrangian as

$$L = \frac{1}{2}g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu . \tag{16.37}$$

(Strictly speaking, this is a "Lagrangian per unit rest mass" for a body moving through spacetime.) The motion of a body in the spacetime $g_{\mu\nu}$ can then be found by applying the Euler-Lagrange equations to this L.

After a bit of discussion about how to get the spacetime metric $g_{\mu\nu}$, we will use this "relativistic Lagrangian" quite a bit in the last few weeks of this course.