

Week ②

PHY 305 Classical Mechanics

Instructor: Sebastian Wüster, IISER Bhopal, 2020

These notes are provided for the students of the class above only. There is no guarantee for correctness, please contact me if you spot a mistake.

2 Lagrangian Mechanics

We had seen that in non-cartesian coordinates systems, that might be forced upon us by symmetry or constraints, Newton's equations can be quite cumbersome. We shall learn Lagrange's equations in this section. Their derivation will be based on what is called a "variational technique". The technique is useful beyond mechanics, also in optics, electro-magnetism, quantum mechanics and quantum field theory, hence we will give a comprehensive introduction in the next section.

2.1 Calculus of Variations

You know how to find an extremal point of a function $y(x)$ using the first derivative $\frac{d}{dx}y(x) \stackrel{!}{=} 0$.

Let us now define a

Functional: $F[y]$ as map of a function $y(x)$ onto a (real or complex) number z .

$$F[y(x)] \rightarrow z \quad (2.1)$$

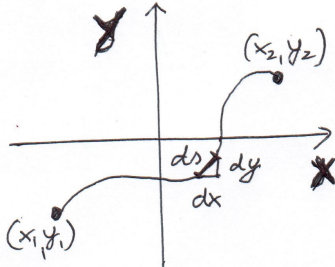
One example would be an integral:

$$F[y(x)] : y(x) \rightarrow z; z = \int_a^b y(x) dx \quad (2.2)$$

In the same way we use the derivative to find a value x_0 for which a function $y(x_0)$ takes an extremal value, we can use the calculus of variations to find a function $y_0(x)$ for which the functional $F[y]$ takes an extremal value.

Before we look at how to do this, let us consider an example to show why one would want to know such a thing:

Example 4, Example Shortest Path between two points: You want to know what is the shortest path in 2D between the points (x_1, y_1) and (x_2, y_2) . Of course you know that is a straight line, but suppose you did not know that? We could try all sorts of complicated path such as shown in the figure.



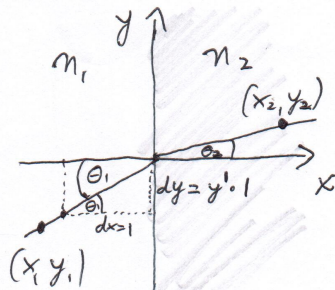
left: (left) 2D path between the two points that clearly is not the shortest.

Since the path is curvy, to find its length we have to integrate up all infinitesimal path segments of length $ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + y'(x)^2}dx$. In the second equality we have used $y'(x) = dy/dx$. The total length of the path, using the line-integral (1.8), is then

$$L = \int_{x_1}^{x_2} ds = \int_{x_1}^{x_2} \sqrt{1 + y'(x)^2} dx. \quad (2.3)$$

Finding the shortest path means minimize L wrt. the values $y(x)$ (and $y'(x)$) at all the points x , so it is much more complicated than finding a function minimum. We will see how to do this further below.

Example 5, Fermat's principle: Path taken by a light ray between two points:



left: (left) Fermat discovered that the path taken by a light ray at an interface with two different refractive indices is such that the light travel time is minimized. See figure)

Similar to the previous example, we can set up the total travel time of the light as

$$T = \int_1^2 dt = \int_1^2 \frac{ds}{v} = \frac{1}{c} \int_1^2 n ds. \quad (2.4)$$

For exercise, we parametrize this curve slightly different than before, using $\mathbf{s}(x) = [x, y(x)]^T$, and thus get

$$T = \frac{1}{c} \int_{x_1}^{x_2} n(x) \sqrt{1 + y'(x)^2} dx, \quad (2.5)$$

where according to the figure of the interface above, the refractive index has the step profile $n(x) = n_1$ for $x < 0$ and $n(x) = n_2$ for $x \geq 0$.

2.2 The Euler-Lagrange equation

In both examples we have just seen, we want to minimize an expression (functional) of the form

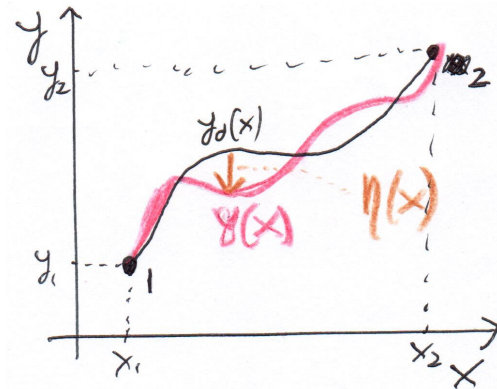
$$S = \int_{x_1}^{x_2} dx f[y(x), y'(x), x] \quad (2.6)$$

- In either example given above, the initial and final points of the path are fixed $y(x_1) = y_1$, $y(x_2) = y_2$, but in between any continuous function $y(x)$ is allowed.

For a function $f(x)$ to have a local minimum at a point x_0 , we need that $f(x) > f(x_0)$ for all x that are sufficiently close to x_0 . If we adapt this definition to to a functional, then for a functional $S[y(x)]$ to be “minimal” near a curve $y_0(x)$, we need that $S[y(x)] > S[y_0(x)]$ for all curves that are sufficiently close to $y_0(x)$. We can write these as:

$$y(x) = y_0(x) + \eta(x), \quad (2.7)$$

where $y_0(x)$ is the curve we want to find, that provides a minimum of S in (2.6) and $\eta(x)$ is some small deviation from it as shown in the figure below. To make sure that $y(x_1) = y_1$ and $y(x_2) = y_2$ regardless of $\eta(x)$ we allow only $\eta(x_1) = \eta(x_2) = 0$.



left: (left) A curve $y_0(x)$ (black) minimizes a functional S as in (2.6), if for all nearby curves $y(x)$ (pink), which are shifted from $y_0(x)$ by an offset $\eta(x)$ (brown), the functional value is larger: $S[y] > S[y_0]$.

Let us use the following trick to reduce the problem to minimisation in terms of just one parameter, which we already know how to solve. For that we write

$$y(x) = y_0(x) + \alpha\eta(x). \quad (2.8)$$

instead of (2.7). If $y_0(x)$ minimizes S , then S is minimal for $\alpha = 0$, hence we know $\frac{d}{d\alpha}S(\alpha) = 0$, where S depends on α via $y(x)$ and $y'(x)$ using (2.8): $S(\alpha) = \int_{x_1}^{x_2} dx f[y_0(x) + \alpha\eta(x), y_0'(x) + \alpha\eta'(x), x]$. We find

$$\begin{aligned} \frac{d}{d\alpha}S(\alpha) &= \int_{x_1}^{x_2} dx \frac{d}{d\alpha} f[y(x), y'(x), x] \\ &= \int_{x_1}^{x_2} dx \left[\frac{\partial}{\partial y(x)} f[y(x), y'(x), x] \frac{dy(x)}{d\alpha} + \frac{\partial}{\partial y'(x)} f[y(x), y'(x), x] \frac{dy'(x)}{d\alpha} \right] \\ &= \int_{x_1}^{x_2} dx \left[\frac{\partial}{\partial y(x)} f[y(x), y'(x), x] \eta(x) + \frac{\partial}{\partial y'(x)} f[y(x), y'(x), x] \eta'(x) \right]. \end{aligned} \quad (2.9)$$

We now apply integration by parts to the last term, to move the derivative from $\eta(x)$ onto the rest of the term:

$$\int_{x_1}^{x_2} dx \frac{\partial}{\partial y'(x)} f[y, y', x] \eta'(x) = \left[\frac{\partial}{\partial y'(x)} f[y, y', x] \eta(x) \right]_{x_1}^{x_2} - \int_{x_1}^{x_2} dx \frac{d}{dx} \left\{ \frac{\partial}{\partial y'(x)} f[y, y', x] \right\} \eta(x). \quad (2.10)$$

The first (boundary) term vanishes due to $\eta(x_1) = \eta(x_2) = 0$ and thus (2.9) becomes

$$\frac{d}{d\alpha} S(\alpha) = \int_{x_1}^{x_2} dx \eta(x) \left[\frac{\partial}{\partial y(x)} f[y, y', x] - \frac{d}{dx} \left\{ \frac{\partial}{\partial y'(x)} f[y, y', x] \right\} \right] \stackrel{!}{=} 0 \quad (2.11)$$

Now comes a tricky argument: Since (2.11) must be true for any choice of the function $\eta(x)$, we require the argument multiplying $\eta(x)$ to vanish itself², which leads us to the

Euler-Lagrange equation Any path $y(x)$ that yields an extremum of the general functional (2.6) must fulfill

$$\frac{\partial}{\partial y} f - \frac{d}{dx} \frac{\partial}{\partial y'} f = 0. \quad (2.12)$$

- If the original functional (2.6) depends on more than one function $S = \int_{x_1}^{x_2} dx f[y(x), y'(x), z(x), z'(x), x]$, one can show that we get a separate Euler Lagrange equation for each: $\frac{\partial}{\partial y} f - \frac{d}{dx} \frac{\partial}{\partial y'} f = 0$ and $\frac{\partial}{\partial z} f - \frac{d}{dx} \frac{\partial}{\partial z'} f = 0$. Importantly this requires that $y(x), z(x)$ can be varied independently from one another.

We can now use this result to solve the two exemplary problems we had introduced above:

Example 6, Shortest Path between two points, solution: Taking the expression for length (2.4) as the functional to be minimized in (2.6), we identify $f(y, y', x) = \sqrt{1 + y'^2}$. Evaluating the partial derivatives of the Euler-Lagrange equation we thus find $\frac{\partial}{\partial y} f = 0$ and $\frac{\partial}{\partial y'} f = \frac{y'}{\sqrt{1 + y'^2}}$. Assembling (2.12) we have

$$\frac{d}{dx} \frac{\partial}{\partial y'} f = 0, \quad (2.13)$$

which implies that $\frac{\partial}{\partial y'} f = C = \text{const}$, hence $y'(x)^2 = C^2(1 + y'(x)^2) \Rightarrow y'(x) = \text{const} = m$. Of course you know $y' = \text{const}$ implies a straight line, or you can integrate the equation to explicitly find $y(x) = mx + y(0)$.

²The statement is: If $\int_{x_1}^{x_2} dx \eta(x)g(x) = 0$ for all choices of continuous functions $\eta(x)$, then $g(x) \equiv 0$. Suppose the latter was not the case and zoom in onto some region of x for which $g(x) > 0$. Then pick a function $\eta(x)$ that also is $\eta(x) > 0$ in that same region, but goes to 0 elsewhere. We then have $\int_{x_1}^{x_2} dx \eta(x)g(x) > 0$ in contradiction to the starting assumption.

Example 7, Fermat's principle, solution: Following the same path as in the previous example, we identify $f(y, y', x) = n(x)\sqrt{1 + y'(x)^2}$. It still does not depend on y (hence $\frac{\partial}{\partial y} f = 0$), but now depends on x in addition to y' . We find $\frac{\partial}{\partial y'} f = \frac{n(x)y'}{\sqrt{1+y'^2}}$ and through the same arguments as before

$$\frac{n(x)y'}{\sqrt{1 + y'^2}} = C = \text{const.} \quad (2.14)$$

We can solve this separately at the left-hand-side and rhs of the interface, to give $n_1 \frac{y'_1}{\sqrt{1+y_1'^2}} = C = n_2 \frac{y'_2}{\sqrt{1+y_2'^2}}$, where $y'_{1,2}$ are the slopes in region 1 and 2. You can show from this that on either side the light follows a straight line (as in the previous example), but using $\sin(\theta_1) = \frac{y'_1}{\sqrt{1+y_1'^2}}$ and $\sin(\theta_2) = \frac{y'_2}{\sqrt{1+y_2'^2}}$ you also immediately recover Snell's refraction law.

Additional examples for the application of the variational principle can be found in the text-books and our assignments.

2.3 Hamilton's principle

Our motivation for section 2.1 was to apply variational calculus to mechanics, which we shall do now. We begin with a simple case: The motion of one unconstrained particle in 1D, subject to a force. For that particle the kinetic energy T and potential energy V are defined in (1.11) and (1.13).

With these, we first define the:

Lagrangian (or Lagrangian function) as

$$\mathcal{L}(x, \dot{x}, t) = T(\dot{x}) - V(x, t) = \frac{1}{2}m\dot{x}^2 - V(x, t). \quad (2.15)$$

- \mathcal{L} has dimensions of energy and the same constituents as the total energy, but it is not the total energy which would be $E = T + V$.
- The symbol \mathcal{L} is chosen to avoid confusion with a length or angular momentum.³
- Unlike our discussion near Eq. (1.13), we have allowed for a time-dependent potential.

From the Lagrangian we then define the

³Note that later quantum field theory texts tend to use \mathcal{L} for a Lagrange density, we will not consider that here.

Action (or action integral) as

$$S = \int_{t_1}^{t_2} \mathcal{L}(x(t), \dot{x}(t), t) dt \quad (2.16)$$

- An action has units “energy×time”. h is also called “quantum of the action”.
- Note that this falls precisely into definition (2.6) for a functional. That definition was just a mathematical one, so we can rename all the variables and it does not matter which physical dimensions they have. So to get from (2.6) to (2.16) $x \rightarrow t$, $y(x) \rightarrow x(t)$, $\frac{d}{dx}y(x) \rightarrow \frac{d}{dt}x(t)$.

With the action we are finally in a position to formulate

Hamilton’s principle The path that a particle takes between times t_1 and t_2 (and locations $\mathbf{x}(t_1)$ and $\mathbf{x}(t_2)$) is such that the action (2.16) is stationary / extremal along the chosen path. This is also called the principle of least action.

To see why this makes sense, let us now apply the Euler-Lagrange equations (2.12) to this mechanical problem. We find

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x} &= \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}} \Leftrightarrow \\ -\frac{\partial V(x, t)}{\partial x} &= \frac{d}{dt} [m\dot{x}(t)] = \dot{p}(t) \end{aligned} \quad (2.17)$$

- We recovered Newton’s equation (1.5). We thus notice that we can base mechanics on the statement ”Hamilton’s principle is valid”, instead of ”Newton’s equations are valid”.
- Note that the correct use of partial derivatives is crucial: When evaluating $\frac{\partial}{\partial x}$, we treat x as independent of \dot{x} and the reverse.
- We could have demonstrated Hamilton’s principle in 3D as well (see book by Taylor), but will rather do ND directly in the next section.
- In fact, while we have introduced it here for a very minimal example, the least action principle is unbelievably powerful. We have found that all the fundamental laws of physics can be formulated in terms of an action principle, where the Lagrangian (density) for the “universe” looks like

$$\mathcal{L} = \sqrt{g}(R - \frac{1}{2}F_{\mu\nu}F^{\mu\nu} + \bar{\Psi}\not{D}\Psi), \quad (2.18)$$

where the three terms in brackets relate to (in convoluted ways) to the gravitational field, all other force fields (electro-magnetic, strong, weak) and all matter fields (quarks, leptons). *[For the idea to this motivation I acknowledge lecture notes by Prof. David Tong that I found online].*