Lagrangian mechanics - Euler-Lagrange equations

Last updated: November 14, 2022

We now go back to situations in which we do not need to worry about relativistic effects, that is, to systems whose behavior is well approximated by Newton's laws. However, we are now going to consider a whole new way of describing all the mechanics you know (Newtonian mechanics first, and then relativistic mechanics). This new formalism is very elegant, perhaps one of the most beautiful theories in physics. But, even more important, this formulation of mechanics opens the door to describing pretty much all known physics under a single unifying principle.

Principle of least action

Consider a particle in one dimension traveling from x_1 at t_1 to x_2 at t_2, for example, a stone that we throw upwards and then falls back down. The trajectory will look something like Fig. 1. If I ask you what is the precise shape of this trajectory, your answer will be "we should integrate Newton's equations of motion $F = \frac{dp}{dt}$ and we'll find out". For example, if the particle is going up, and then down, in a gravitational field, the shape will be a parabola, exactly as depicted in Fig. 1.

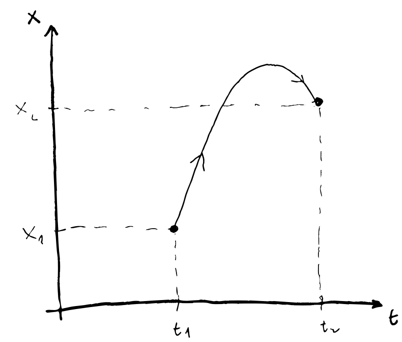

Figure 1: Real trajectory of a particle going up and down in a constant gravitational field.

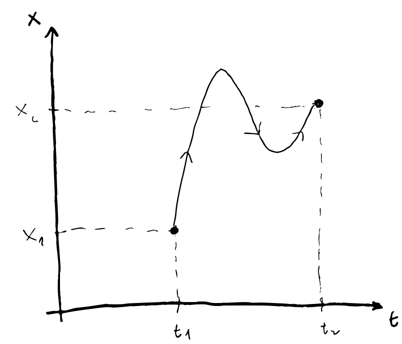

Figure 2: Hypothetical alternative trajectory of a particle going up and down in a constant gravitational field.

But here is another way of answering the question, which is going to occupy us most of the rest of the course. Imagine another trajectory that goes from x_1 at t_1 to x_2 at t_2 as depicted in Fig. 2. Then, and here comes the magic, if you calculate the kinetic energy at every moment on the path, take away the potential energy, and integrate the difference over the whole path, you'll find that the number you get for this imaginary trajectory is larger than for the true trajectory. In other words, the laws of Newton could be stated not in the form $F = ma$ but in the form: **the average kinetic energy less the average potential energy is as little as possible for the path of an object going from one point to another.**

This quantity, the integral of the kinetic energy minus the potential energy, is called the action S

$$S = \int_{t_1}^{t_2} dt \, (T - V) \ . \tag{1}$$

And, as we will see, all of Newtonian mechanics (and, in fact, relativistic mechanics, quantum mechanics, electromagnetism, ...) can be reformulated by saying that the actual trajectories of a system (among all possible trajectories) are those that minimize the action. This is **Hamilton's principle, or principle of least action.**[1]

Unlike Newton's formulation, which is *differential* (it "tells" the particle how to move *at each time*), the principle of least action is *integral* (it makes prescriptions on *the trajectory as a whole*). This may seem a bit strange—how does the particle know what it "needs to do now" in order to minimize the action globally? As we will see next, it turns out that the integral principle of least action implies also some differential principle, the Euler-Lagrange equation. But let's first give some examples and then build our way up to the equation.

[1] Note that we could as well have defined the action as $V - T$; then the action would have to be maximized rather than minimized. In reality, all we are going to impose is that the action is stationary (minimal or maximal) for the true trajectories, so we could as well call it the *principle of stationary action.*

Free particle in one dimension

Consider a free particle moving in one dimension. Since the potential energy is $V(x) = 0$, we only need to minimize the mean (or integrated) kinetic energy $T = mv^2/2$. Since the particle has to go from x_1 at t_1 to x_2 at t_2, the mean velocity will be the same for all trajectories we are interested in. Consider, first, the trajectory with constant velocity $\langle v \rangle = v_0 = (x_2 - x_1)/(t_2 - t_1)$ (Fig. 3).

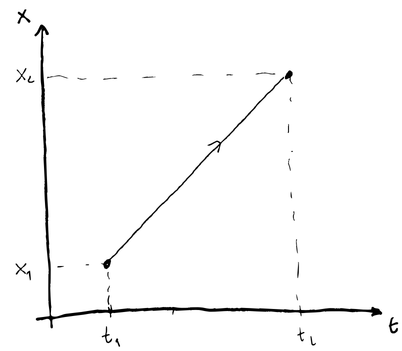

Figure 3: Trajectory of a free particle at constant velocity.

The action for this trajectory is

$$S_{\text{const}} = \int_{t_1}^{t_2} dt\, \frac{1}{2} m\, (v(t))^2 = \frac{t_2 - t_1}{2} m v_0^2\,.$$

Now, any other trajectory can be written as $v(t) = v_0 + \eta(t)$ and, since the mean velocity of this trajectory still has to be v_0,[2]

[2] This is because we still want to go between x_1 at t_1 and x_2 in t_2. In other words, we must have $\eta(t_1) = \eta(t_2) = 0$.

$$v_0 = \langle v(t) \rangle = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} dt\, v(t) = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} dt\, (v_0 + \eta(t)) = v_0 + \langle \eta(t) \rangle\,,$$

then we have that

$$\langle \eta(t) \rangle = 0\,.$$

Therefore

$$\begin{aligned} S_\eta &= \int_{t_1}^{t_2} dt\, \frac{1}{2} m\, (v_0 + \eta(t))^2 = \frac{t_2 - t_1}{2} m \left(v_0^2 + 2 v_o \langle \eta(t) \rangle + \langle \eta(t)^2 \rangle \right) \\ &= S_{\text{const}} + \frac{t_2 - t_1}{2} m \langle \eta(t)^2 \rangle\,, \end{aligned} \qquad (2)$$

and $S_\eta > S_{\text{const}}$ for any $\eta(t) \neq 0$, that is, for any trajectory other than the constant velocity trajectory.

Particle in a potential in one dimension

Now consider a single particle in one dimension with a generic potential $V(x)$. What does the principle of least action imply? Here things are a bit more complicated, since we cannot make a good guess to start (equivalent to "constant velocity" for a free particle).

Here, you may be tempted to say: "But that is not that difficult! Just get the action, calculate the derivative and equate to 0, and you'll have your solution!" But that's not the case because the action is not a "regular" function.[3] However, that is not a bad idea altogether: just as in regular calculus we look for values of x around which changes in f are second order in δx (Fig. 4, first derivative equals 0), here we look for trajectories around which changes in the action S are second order.

[3] We will soon be much more precise about what this means exactly. For now, just note that the action is not a regular function because its value depends, itself, of a whole function!

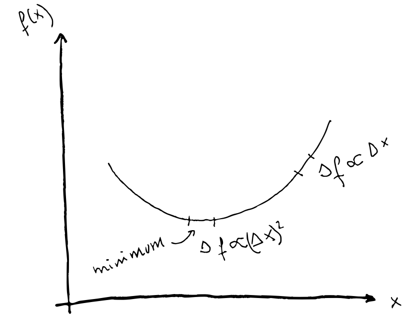

Figure 4: In ordinary calculus, we find minima (or maxima) of functions by identifying the points where increments Δf of the function are proportional to $(\Delta x)^2$ instead of Δx, that is, points where the first derivative vanishes.

So now, let's assume that such trajectory around which changes in the action S are second order is $x_0(t)$, and we build a slightly different trajectory $x_0(t) + \delta x(t)$, where $\delta x(t) \ll x_0(t)$ is a small perturbation of the "true" trajectory (Fig. 5).[4] Then, the change in the action induced by the small perturbation in the trajectory is

$$\begin{aligned} \delta S &= S[x_0 + \delta x] - S[x_0] = \qquad (3)\\ &= \int_{t_1}^{t_2} dt \left[\frac{m}{2} \left(\frac{dx_0}{dt} + \frac{d(\delta x)}{dt} \right)^2 - V(x_0 + \delta x) \right] \\ &- \int_{t_1}^{t_2} dt \left[\frac{m}{2} \left(\frac{dx_0}{dt} \right)^2 - V(x_0) \right] . \qquad (4) \end{aligned}$$

Considering that, at each t,

$$V(x_0 + \delta x) = V(x_0) + \delta x \left. \frac{dV}{dx} \right|_{x_0} + O((\delta x)^2) \, ,$$

to first order in δx we have

$$\delta S = \int_{t_1}^{t_2} dt \left[m \frac{dx_0}{dt} \frac{d(\delta x)}{dt} - \delta x \left. \frac{dV}{dx} \right|_{x_0} \right] + O((\delta x)^2)$$

Now we can integrate the first term by parts to get

$$\delta S = \int_{t_1}^{t_2} dt \left[-m \frac{d^2 x_0}{dt^2} - \left. \frac{dV}{dx} \right|_{x_0} \right] \delta x + O((\delta x)^2) \, .$$

Now, since this must be 0, we have that

$$m \frac{d^2 x_0}{dt^2} = - \left. \frac{dV}{dx} \right|_{x_0}$$

which is exactly Newton's second law! So, indeed, the trajectory that minimizes the action satisfies Newton's second law, so the physics is exactly the same!

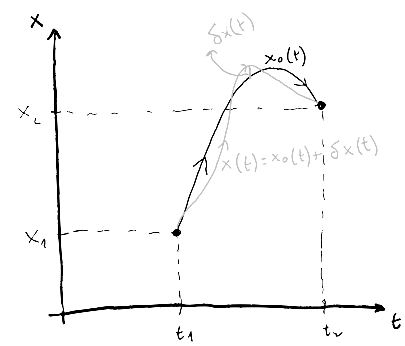

Figure 5: We consider the trajectory $x_0(t)$ that makes the action S stationary, and a slightly different trajectory $x(t) = x_0(t) + \delta x(t)$ around it. Note that $\delta x(t_1) = \delta x(t_2) = 0$.

[4] Note that $\delta x(t)$ is, despite its "weird" name, just a regular function of t. Just call it $\eta(t)$, if you prefer.

Functionals and functional derivatives

Now, let's be a bit more formal about the whole action thing. Consider a function $f : \mathbb{R}^n \to \mathbb{R}$ that maps an n-dimensional space onto the real numbers. Although we are used to functions with finite n, the concept of function still makes sense in the limit $n \to \infty$

$$f(\mathbf{x}) \in \mathbb{R}\,, \quad \text{with} \quad \mathbf{x} = \{x_i,\, i = 1, 2, \ldots\}$$

where i is the index for a discrete infinity of variables. But, instead of having a discrete infinity of variables, we can equally well have a **continuous** infinity of variables $x = \{x(t),\, t \in \mathbb{R}\}$, that is, a function. A function of this continuous infinity of variables (a function of this function $x(t)$) is called a **functional** and denoted $F[x(t)]$ or simply $F[x]$—just to restate, the functional $F[x] \in \mathbb{R}$ maps any function $x(t)$ into a real number. [5]

[5] Note that the functional F depends on the function $x(t)$ but it doesn't depend on the independent variable t; that's just a label, analogous to the discrete label i for the variables x of the function $f(x)$.

Just as we can have functions of many variables, so we can have functionals of many functions

$$F[\mathbf{x}] \in \mathbb{R}\,, \quad \text{with} \quad \mathbf{x} = \{\mathbf{x}(t) \in \mathbb{R}^n, t \in \mathbb{R}\},$$

functionals of a function of many variables

$$F[x] \in \mathbb{R}\,, \quad \text{with} \quad x = \{x(\mathbf{t}) \in \mathbb{R},\, \mathbf{t} \in \mathbb{R}^n\},$$

and functionals of many functions of many variables

$$F[\mathbf{x}] \in \mathbb{R}\,, \quad \text{with} \quad \mathbf{x} = \{\mathbf{x}(\mathbf{t}) \in \mathbb{R}^n,\, \mathbf{t} \in \mathbb{R}^m\},$$

But how can we map a function into a single value? Well, we could take the value of the function at one or a few points t_i—but that would not be very interesting, as that would give us a functional of a finite number of variables, that is, a function! Therefore, we can convince ourselves that functionals naturally involve integrating the function or some suitable transformation of the function. Given these considerations, a quite general form for functionals is

$$F[x] = \int_\alpha^\beta dt\, f(x, x', t)$$

where $x' = dx/dt$. We will focus on this particular form, as this is general enough to deal with a wide range of physical situations. Notice that the integrand has an implicit dependence on t through $x(t)$ and $x'(t)$, as well as an explicit dependence on t. But why not x''? Well, this just expresses the result that, in mechanics, we can predict the evolution of the system by specifying current positions and velocities only, without the need to specify accelerations.

Now, as in the examples discussed earlier, we are interested in the functions $x(t)$ that minimize $F[x]$. As per the arguments above, this involves identifying the functions that lead to zero variation of the functional, $\delta F[x] = 0$ when we modify $x(t)$ slightly to a nearby function $x(t) + \delta x(t)$.[6] So let's calculate $\delta F[x]$:

[6] Note that $\delta x(t)$ is just a regular function of t, which we could have called for example $\eta(t)$, and not a variation of any functional.

$$\begin{aligned} \delta F[x] &= F[x+\delta x] - F[x] \\ &= \int_\alpha^\beta dt\, f(x+\delta x, x' + (\delta x)', t) - \int_\alpha^\beta dt\, f(x, x', t) \\ &= \int_\alpha^\beta dt\, \left\{ \delta x \frac{\partial f}{\partial x} + (\delta x)' \frac{\partial f}{\partial x'} \right\} + O(2) \end{aligned}$$

where we have expanded f to first order in δx and $(\delta x)'$ and ignored quadratic terms and above. Now, integrating the second term by parts and leaving the first one as is we get

$$\delta F[x] = \left[\delta x \frac{\partial f}{\partial x'} \right]_\alpha^\beta + \int_\alpha^\beta dt\, \left\{ \delta x \left[\frac{\partial f}{\partial x} - \frac{d}{dt} \left(\frac{\partial f}{\partial x'} \right) \right] \right\}$$

Since we are interested in perturbations $x(t) + \delta x(t)$ that start and end at the same points as the original $x(t)$, we must have $\delta x(\alpha) = \delta x(\beta) = 0$, so that the term outside the integral vanishes. Finally, for a functional of the form we have been considering, we can define the **functional derivative** as[7]

$$\frac{\delta F[x]}{\delta x(t)} \equiv \frac{\partial f}{\partial x} - \frac{d}{dt} \left(\frac{\partial f}{\partial x'} \right)$$

and we can write the variation as

$$\delta F[x] = \int_\alpha^\beta dt\, \left\{ \delta x(t) \frac{\delta F[x]}{\delta x(t)} \right\}.$$

Here, we have made the dependency of x on t explicit to highlight the analogy with the result from standard calculus

$$df(\mathbf{x}) = \sum_i dx_i \frac{\partial f}{\partial x_i}.$$

Now, since the perturbation δx we have been considering is arbitrary, the only way to guarantee that $\delta F[x] = 0$ is by satisfying

$$\frac{\delta F[x]}{\delta x(t)} = 0,$$

that is,

$$\frac{\partial f}{\partial x} - \frac{d}{dt} \left(\frac{\partial f}{\partial x'} \right) = 0 \quad , \quad \alpha \leq t \leq \beta.$$

Analogously, when the functional $F[\mathbf{x}]$ depends on multiple functions $\mathbf{x} = \{x_i(t),\, i = 1, 2, \ldots, k\}$, we can derive multiple equations

$$\frac{\partial f}{\partial x_i} - \frac{d}{dt} \left(\frac{\partial f}{\partial x_i'} \right) = 0 \quad , \quad i = 1, 2, \ldots, k \quad \alpha \leq t \leq \beta,$$

These are the so-called **Euler-Lagrange equations**.

For a system with k degrees of freedom, these equations constitute a set of k second-order differencial equations for the functions $x_i(t)$. The general solution contains $2k$ constants; to determine these constants and thereby define uniquely the motion of the system, we need $2k$ initial conditions (for example, the initial positions and velocities for each coordinate).

[7] Note that this is just notation, and the $\delta x(t)$ in the "denominator" does not *cancel out* with any multiplying $\delta x(t)$, which, as we have noted, is just a regular function!

Principle of least action and the Euler-Lagrange equations

We are now in a position to finally state the principle of least action more formally. Consider the action, which is a functional defined as

$$S = \int_{t_1}^{t_2} dt\ \mathcal{L}(\mathbf{x}, \mathbf{x}', t)\ ,$$

with $\mathcal{L}(\mathbf{x}, \mathbf{x}', t) = T - V$ being the so-called **Lagrangian** of the system. According to Hamilton's principle, the true trajectory of the system is the one that makes the action stationary, that is, $\delta S = 0$. Then, according to the functional derivatives defined above, the true trajectories of the system obey the **Euler-Lagrange equations**

$$\boxed{\frac{\partial \mathcal{L}}{\partial x_i} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial x_i'}\right) = 0 \quad , \quad i = 1, 2, \ldots, k} \tag{5}$$

where k is the number of degrees of freedom of the system.[8,9]

A final important consideration. Despite Hamilton's principle being an integral principle, it leads to a system of differential equations. Thus, the "mystery" of how can a particle know, at each step, where to go next so as to globally minimize the action is solved. In fact, there is a simple way to see why this is the case. If we choose to interrupt the trajectory of a particle between t_1 and t_2, the trajectory (up to the interruption) should be the same as when when let the particle run until t_2—in both cases, the action must be minimized. Then, it follows that each "step" of the trajectory must be minimizing the action; minimizing the global action is achieved by minimizing the action at each step.

[8] For example, for a system of N particles in 3 dimensions, $k = 3N$.

[9] EXERCISE: Derive, this time using the Euler-Lagrange equations, the equations of motion for a free particle and for a particle in a potential $V(x)$, both in 1D.

Euler-Lagrange equations in arbitrary coordinates

There is nothing in the formulation of the principle of least action that makes reference to one particular coordinate system; it would be really bad news if the Euler-Lagrange equations only held in some coordinate systems but not in others (for example, if trajectories minimized the action in Cartesian coordinates but not in, say, polar coordinates). However, so far we have only been using Cartesian coordinates (x, y, z). Next, we show that, if the Euler-Lagrange equations hold in one coordinate system (and we know they do in Cartesian coordinates), then they **hold in *any* coordinate system**.

Indeed, consider the **generalized coordinates** $\mathbf{q} = \mathbf{q}(\mathbf{x}; t)$.[10] Conversely, we can express the old coordinates in terms of the new ones $\mathbf{x} = \mathbf{x}(\mathbf{q}; t)$. We aim to show that

$$\frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial q_i'}\right) = 0\ ;$$

[10] Note that the change of coordinates cannot depend on the velocities $\mathbf{x}'$. Otherwise, each point in space would not be uniquely mapped for a given t.

let's start with the second term. Since $\partial x_j/\partial q'_i = 0$ because $\mathbf{x}$ only depends explicitly of $\mathbf{q}$ (and perhaps t, but not $\mathbf{q}'$), we have that

$$\frac{\partial \mathcal{L}}{\partial q'_i} = \sum_j \frac{\partial \mathcal{L}}{\partial x'_j}\frac{\partial x'_j}{\partial q'_i} = \sum_j \frac{\partial \mathcal{L}}{\partial x'_j}\frac{\partial x_j}{\partial q_i} \,,$$

where we have used the fact that $\partial x'_j/\partial q'_i = \partial x_j/\partial q_i$.[11] Now we take the total derivative with respect to t

[11] Indeed, $x'_j = dx_j/dt = \sum_k \partial x_j/\partial q_k \cdot q'_k + \partial x_j/\partial t$. Therefore, $\partial x'_j/\partial q'_i = \partial x_j/\partial q_i$ $\Box$.

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial q'_i}\right) = \sum_j \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial x'_j}\right)\frac{\partial x_j}{\partial q_i} + \sum_j \frac{\partial \mathcal{L}}{\partial x'_j}\frac{d}{dt}\left(\frac{\partial x_j}{\partial q_i}\right) \,.$$

Finally, we use the Euler-Lagrange equation on the first term, and switch the order of the derivatives in the second[12] to get

[12] Indeed, $d/dt\left(\partial x_j/\partial q_i\right) = \sum_k \partial/\partial q_k \left(\partial x_j/\partial q_i\right) q'_k + \partial/\partial t \left(\partial x_j/\partial q_i\right) = \partial/\partial q_i \left(\sum_k \partial x_j/\partial q_k\, q'_k + \partial x_i/\partial t\right) = \partial x'_i/\partial q_i$ $\Box$.

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial q'_i}\right) = \sum_j \frac{\partial \mathcal{L}}{\partial x_j}\frac{\partial x_j}{\partial q_i} + \sum_j \frac{\partial \mathcal{L}}{\partial x'_j}\frac{\partial x'_j}{\partial q_i} = \frac{\partial \mathcal{L}}{\partial q_i} \quad \Box \,.$$

Solving problems using the Lagrangian formalism

So far, we have shown that the principle of least action and the ensuing Euler-Lagrange equations lead to the same physics as Newton's laws. The **general workflow for solving mechanics problems using the Lagrangian formalism** is always the same:

1. Write the Lagrangian in a coordinate system that is convenient (typically, choosing an inertial frame and using Cartesian coordinates).

2. Identify the most convenient coordinates to describe the system and rewrite the Lagrangian using the new generalized coordinates. In general, if the system consists of N masses in 3D, and there are k geometric constraints, we will need $3N - k$ coordinates to describe the system.

3. Use the Euler-Lagrange equations to obtain the equations of motion.

But, given that this leads to the same physics as Newton's second law, what is the advantage of using the Lagrangian formalism?

The first advantage is that all of the physics of a system (all the parameters such as masses, and all the equations of motion) is packed into a single function—the Lagrangian. A single function specifying the behavior of any number of degrees of freedom, not bad! And there is more, it turns out that this is true not only for mechanics—all the physics we know can be formulated in terms of Lagrangians, including, for example, Maxwell's theory of electrodynamics, Einstein's relativity, or the standard model of particle physics (Fig. 6).

$$
\begin{gathered}
\mathcal{L}_{SM} = -\tfrac{1}{2}\partial_\nu g^a_\mu \partial_\nu g^a_\mu - g_s f^{abc}\partial_\mu g^a_\nu g^b_\mu g^c_\nu - \tfrac{1}{4}g_s^2 f^{abc}f^{ade} g^b_\mu g^c_\nu g^d_\mu g^e_\nu - \partial_\nu W^+_\mu \partial_\nu W^-_\mu - \\
M^2 W^+_\mu W^-_\mu - \tfrac{1}{2}\partial_\nu Z^0_\mu \partial_\nu Z^0_\mu - \tfrac{1}{2c_w^2} M^2 Z^0_\mu Z^0_\mu - \tfrac{1}{2}\partial_\mu A_\nu \partial_\mu A_\nu - igc_w(\partial_\nu Z^0_\mu (W^+_\mu W^-_\nu - \\
W^+_\nu W^-_\mu) - Z^0_\nu (W^+_\mu \partial_\nu W^-_\mu - W^-_\mu \partial_\nu W^+_\mu) + Z^0_\mu (W^+_\nu \partial_\nu W^-_\mu - W^-_\nu \partial_\nu W^+_\mu)) - \\
igs_w(\partial_\nu A_\mu (W^+_\mu W^-_\nu - W^+_\nu W^-_\mu) - A_\nu (W^+_\mu \partial_\nu W^-_\mu - W^-_\mu \partial_\nu W^+_\mu) + A_\mu (W^+_\nu \partial_\nu W^-_\mu - \\
W^-_\nu \partial_\nu W^+_\mu)) - \tfrac{1}{2}g^2 W^+_\mu W^-_\mu W^+_\nu W^-_\nu + \tfrac{1}{2}g^2 W^+_\mu W^-_\nu W^+_\mu W^-_\nu + g^2 c_w^2 (Z^0_\mu W^+_\mu Z^0_\nu W^-_\nu - \\
Z^0_\mu Z^0_\mu W^+_\nu W^-_\nu) + g^2 s_w^2 (A_\mu W^+_\mu A_\nu W^-_\nu - A_\mu A_\mu W^+_\nu W^-_\nu) + g^2 s_w c_w (A_\mu Z^0_\nu (W^+_\mu W^-_\nu - \\
W^+_\nu W^-_\mu) - 2A_\mu Z^0_\mu W^+_\nu W^-_\nu) - \tfrac{1}{2}\partial_\mu H \partial_\mu H - 2M^2 \alpha_h H^2 - \partial_\mu \phi^+ \partial_\mu \phi^- - \tfrac{1}{2}\partial_\mu \phi^0 \partial_\mu \phi^0 - \\
\beta_h \left(\tfrac{2M^2}{g^2} + \tfrac{2M}{g}H + \tfrac{1}{2}(H^2 + \phi^0\phi^0 + 2\phi^+\phi^-)\right) + \tfrac{2M^4}{g^2}\alpha_h - \\
g\alpha_h M\left(H^3 + H\phi^0\phi^0 + 2H\phi^+\phi^-\right) - \\
\tfrac{1}{8}g^2\alpha_h\left(H^4 + (\phi^0)^4 + 4(\phi^+\phi^-)^2 + 4(\phi^0)^2\phi^+\phi^- + 4H^2\phi^+\phi^- + 2(\phi^0)^2 H^2\right) - \\
gMW^+_\mu W^-_\mu H - \tfrac{1}{2}g\tfrac{M}{c_w^2} Z^0_\mu Z^0_\mu H - \\
\tfrac{1}{2}ig\left(W^+_\mu(\phi^0\partial_\mu\phi^- - \phi^-\partial_\mu\phi^0) - W^-_\mu(\phi^0\partial_\mu\phi^+ - \phi^+\partial_\mu\phi^0)\right) + \\
\tfrac{1}{2}g\left(W^+_\mu(H\partial_\mu\phi^- - \phi^-\partial_\mu H) + W^-_\mu(H\partial_\mu\phi^+ - \phi^+\partial_\mu H)\right) + \tfrac{1}{2}g\tfrac{1}{c_w}(Z^0_\mu(H\partial_\mu\phi^0 - \phi^0\partial_\mu H) + \\
M\left(\tfrac{1}{c_w}Z^0_\mu\partial_\mu\phi^0 + W^+_\mu\partial_\mu\phi^- + W^-_\mu\partial_\mu\phi^+\right) - ig\tfrac{s_w^2}{c_w}MZ^0_\mu(W^+_\mu\phi^- - W^-_\mu\phi^+) + igs_w MA_\mu(W^+_\mu\phi^- - \\
W^-_\mu\phi^+) - ig\tfrac{1-2c_w^2}{2c_w}Z^0_\mu(\phi^+\partial_\mu\phi^- - \phi^-\partial_\mu\phi^+) + igs_w A_\mu(\phi^+\partial_\mu\phi^- - \phi^-\partial_\mu\phi^+) - \\
\tfrac{1}{4}g^2 W^+_\mu W^-_\mu\left(H^2 + (\phi^0)^2 + 2\phi^+\phi^-\right) - \tfrac{1}{8}g^2\tfrac{1}{c_w^2}Z^0_\mu Z^0_\mu\left(H^2 + (\phi^0)^2 + 2(2s_w^2 - 1)^2\phi^+\phi^-\right) - \\
\tfrac{1}{2}g^2\tfrac{s_w^2}{c_w}Z^0_\mu\phi^0(W^+_\mu\phi^- + W^-_\mu\phi^+) - \tfrac{1}{2}ig^2\tfrac{s_w^2}{c_w}Z^0_\mu H(W^+_\mu\phi^- - W^-_\mu\phi^+) + \tfrac{1}{2}g^2 s_w A_\mu\phi^0(W^+_\mu\phi^- + \\
W^-_\mu\phi^+) + \tfrac{1}{2}ig^2 s_w A_\mu H(W^+_\mu\phi^- - W^-_\mu\phi^+) - g^2\tfrac{s_w}{c_w}(2c_w^2 - 1)Z^0_\mu A_\mu\phi^+\phi^- - \\
g^2 s_w^2 A_\mu A_\mu\phi^+\phi^- + \tfrac{1}{2}ig_s\lambda^a_{ij}(\bar{q}^\sigma_i\gamma^\mu q^\sigma_j)g^a_\mu - \bar{e}^\lambda(\gamma\partial + m^\lambda_e)e^\lambda - \bar{\nu}^\lambda(\gamma\partial + m^\lambda_\nu)\nu^\lambda - \bar{u}^\lambda_j(\gamma\partial + \\
m^\lambda_u)u^\lambda_j - \bar{d}^\lambda_j(\gamma\partial + m^\lambda_d)d^\lambda_j + igs_w A_\mu\left(-(\bar{e}^\lambda\gamma^\mu e^\lambda) + \tfrac{2}{3}(\bar{u}^\lambda_j\gamma^\mu u^\lambda_j) - \tfrac{1}{3}(\bar{d}^\lambda_j\gamma^\mu d^\lambda_j)\right) + \\
\tfrac{ig}{4c_w}Z^0_\mu\{(\bar{\nu}^\lambda\gamma^\mu(1+\gamma^5)\nu^\lambda) + (\bar{e}^\lambda\gamma^\mu(4s_w^2 - 1 - \gamma^5)e^\lambda) + (\bar{d}^\lambda_j\gamma^\mu(\tfrac{4}{3}s_w^2 - 1 - \gamma^5)d^\lambda_j) + \\
(\bar{u}^\lambda_j\gamma^\mu(1 - \tfrac{8}{3}s_w^2 + \gamma^5)u^\lambda_j)\} + \tfrac{ig}{2\sqrt{2}}W^+_\mu\left((\bar{\nu}^\lambda\gamma^\mu(1+\gamma^5){U^{lep}}_{\lambda\kappa}e^\kappa) + (\bar{u}^\lambda_j\gamma^\mu(1+\gamma^5)C_{\lambda\kappa}d^\kappa_j)\right) + \\
\tfrac{ig}{2\sqrt{2}}W^-_\mu\left((\bar{e}^\kappa {U^{lep}}^\dagger_{\kappa\lambda}\gamma^\mu(1+\gamma^5)\nu^\lambda) + (\bar{d}^\kappa_j C^\dagger_{\kappa\lambda}\gamma^\mu(1+\gamma^5)u^\lambda_j)\right) + \\
\tfrac{ig}{2M\sqrt{2}}\phi^+\left(-m^\kappa_e(\bar{\nu}^\lambda {U^{lep}}_{\lambda\kappa}(1-\gamma^5)e^\kappa) + m^\lambda_\nu(\bar{\nu}^\lambda {U^{lep}}_{\lambda\kappa}(1+\gamma^5)e^\kappa)\right) + \\
\tfrac{ig}{2M\sqrt{2}}\phi^-\left(m^\lambda_e(\bar{e}^\lambda {U^{lep}}^\dagger_{\lambda\kappa}(1+\gamma^5)\nu^\kappa) - m^\kappa_\nu(\bar{e}^\lambda {U^{lep}}^\dagger_{\lambda\kappa}(1-\gamma^5)\nu^\kappa)\right) - \tfrac{g}{2}\tfrac{m^\lambda_\nu}{M}H(\bar{\nu}^\lambda\nu^\lambda) - \\
\tfrac{g}{2}\tfrac{m^\lambda_e}{M}H(\bar{e}^\lambda e^\lambda) + \tfrac{ig}{2}\tfrac{m^\lambda_\nu}{M}\phi^0(\bar{\nu}^\lambda\gamma^5\nu^\lambda) - \tfrac{ig}{2}\tfrac{m^\lambda_e}{M}\phi^0(\bar{e}^\lambda\gamma^5 e^\lambda) - \tfrac{1}{4}\bar{\nu}_\lambda M^R_{\lambda\kappa}(1-\gamma_5)\hat{\nu}_\kappa - \\
\tfrac{1}{4}\bar{\nu}_\lambda M^R_{\lambda\kappa}(1-\gamma_5)\hat{\nu}_\kappa + \tfrac{ig}{2M\sqrt{2}}\phi^+\left(-m^\kappa_d(\bar{u}^\lambda_j C_{\lambda\kappa}(1-\gamma^5)d^\kappa_j) + m^\lambda_u(\bar{u}^\lambda_j C_{\lambda\kappa}(1+\gamma^5)d^\kappa_j)\right) + \\
\tfrac{ig}{2M\sqrt{2}}\phi^-\left(m^\lambda_d(\bar{d}^\lambda_j C^\dagger_{\lambda\kappa}(1+\gamma^5)u^\kappa_j) - m^\kappa_u(\bar{d}^\lambda_j C^\dagger_{\lambda\kappa}(1-\gamma^5)u^\kappa_j)\right) - \tfrac{g}{2}\tfrac{m^\lambda_u}{M}H(\bar{u}^\lambda_j u^\lambda_j) - \\
\tfrac{g}{2}\tfrac{m^\lambda_d}{M}H(\bar{d}^\lambda_j d^\lambda_j) + \tfrac{ig}{2}\tfrac{m^\lambda_u}{M}\phi^0(\bar{u}^\lambda_j\gamma^5 u^\lambda_j) - \tfrac{ig}{2}\tfrac{m^\lambda_d}{M}\phi^0(\bar{d}^\lambda_j\gamma^5 d^\lambda_j) + \bar{G}^a\partial^2 G^a + g_s f^{abc}\partial_\mu\bar{G}^a G^b g^c_\mu + \\
\bar{X}^+(\partial^2 - M^2)X^+ + \bar{X}^-(\partial^2 - M^2)X^- + \bar{X}^0(\partial^2 - \tfrac{M^2}{c_w^2})X^0 + \bar{Y}\partial^2 Y + igc_w W^+_\mu(\partial_\mu\bar{X}^0 X^- - \\
\partial_\mu\bar{X}^+ X^0) + igs_w W^+_\mu(\partial_\mu\bar{Y}X^- - \partial_\mu\bar{X}^+ Y) + igc_w W^-_\mu(\partial_\mu\bar{X}^- X^0 - \\
\partial_\mu\bar{X}^0 X^+) + igs_w W^-_\mu(\partial_\mu\bar{X}^- Y - \partial_\mu\bar{Y}X^+) + igc_w Z^0_\mu(\partial_\mu\bar{X}^+ X^+ - \\
\partial_\mu\bar{X}^- X^-) + igs_w A_\mu(\partial_\mu\bar{X}^+ X^+ - \\
\partial_\mu\bar{X}^- X^-) - \tfrac{1}{2}gM\left(\bar{X}^+X^+H + \bar{X}^-X^-H + \tfrac{1}{c_w^2}\bar{X}^0X^0H\right) + \tfrac{1-2c_w^2}{2c_w}igM\left(\bar{X}^+X^0\phi^+ - \bar{X}^-X^0\phi^-\right) + \\
\tfrac{1}{2c_w}igM\left(\bar{X}^0X^-\phi^+ - \bar{X}^0X^+\phi^-\right) + igMs_w\left(\bar{X}^0X^-\phi^+ - \bar{X}^0X^+\phi^-\right) + \\
\tfrac{1}{2}igM\left(\bar{X}^+X^+\phi^0 - \bar{X}^-X^-\phi^0\right) .
\end{gathered}
$$

Figure 6: Lagrangian of the standard model of particle physics, describing three of the four known fundamental forces in the universe (electromagnetic, weak and strong interactions, but not gravity) and classifying all known elementary particles.

The second advantage of the Lagrangian formulation of mechanics is that, once we know the Lagrangian in one (typically inertial) frame of reference, we can easily write the Lagrangian in any other frame of reference, including non-inertial frames of reference, where Newton's laws become complicated because of "fictitious forces."[13] Let's see two examples of this.

[13] "Fictitious forces" is a terrible name—they do not feel fictitious at all when you are standing on the bus and the driver slams on the brakes.

Equations of motion for a moving observer. Consider a particle in one dimension that, from the point of view of an observer A at rest in an inertial reference frame, moves according to Newton's second law. A describes the position of the particle using a coordinate $x(t)$.

Now consider a second observer B that moves with respect to A—the position of B with respect to A is given by an arbitrary function

$f(t)$. This function can be anything, so B may well be accelerated at some point of its trajectory. The coordinates $X(t)$ that B assigns to the particle are[14] $X(t) = x(t) - f(t)$ but, what are the equations of motion that describe the motion of the particle from the point of view of B?

[14] Ignoring relativistic effects, of course!

Since A lives in an inertial reference frame, we can easily write down the Lagrangian for the particle in A's frame

$$\mathcal{L} = T - V = \frac{1}{2}mx'^2 - V(x) \ .$$

From here, it is also straightforward to write the Lagrangian in terms of B's coordinates. Indeed, since $x' = X' + f'$ and $V(x) = V(X)$[15], the Lagrangian in B's reference frame is[16]

[15] The potential is a scalar, so at a given location its value does not depend on the reference frame: same point, same value, different way to label the location.

[16] Note that the new kinetic energy has terms that are linear on X', in addition to the ordinary quadratic terms.

$$\mathcal{L} = \frac{1}{2}m(X' + f')^2 - V(X) \ .$$

Now, B can get the Euler-Lagrange equation to get the equations of motion[17]

[17] EXERCISE: Do it!

$$mX'' = -\frac{dV}{dX} - mf'' \ .$$

So, indeed, B sees an extra "fictitious" force $-mf''$ acting on the particle. The interesting thing is that instead of figuring out how to transform the equations of motion, we simply had to transform the Lagrangian.

EQUATIONS OF MOTION FOR A ROTATING OBSERVER. Let's now analyze a more complicated case in two dimensions. Now, B's origin is not moving with respect to A's as in the previous example, but B's reference frame is rotating at constant angular velocity ω with respect to A's inertial frame. Then the change of coordinates is given by

$$\begin{aligned} x &= +X\cos\omega t + Y\sin\omega t \\ y &= -X\sin\omega t + Y\cos\omega t \end{aligned}$$

For simplicity, we now assume that there is no potential, so A writes the Lagrangian as

$$\mathcal{L} = \frac{m}{2}\left(x'^2 + y'^2\right) \ .$$

To rewrite the Lagrangian in B's frame, we need to obtain x' and y' in B's coordinates, namely

$$\begin{aligned} x' &= +X'\cos\omega t - \omega X\sin\omega t + Y'\sin\omega t + \omega Y\cos\omega t \\ y' &= -X'\sin\omega t - \omega X\cos\omega t + Y'\cos\omega t - \omega Y\sin\omega t \ . \end{aligned}$$

After some algebra,[18] B's Lagrangian reads

[18] EXERCISE: Do it!

$$\mathcal{L} = \frac{m}{2}\left(X'^2 + Y'^2\right) + \frac{m\omega^2}{2}\left(X^2 + Y^2\right) + m\omega\left(X'Y - Y'X\right)$$

and, after using the Euler-Lagrange equations, one gets the following equations of motion[19]

[19] Exercise: Do it!

$$\begin{aligned} mX'' &= m\omega^2 X - 2m\omega Y' \\ mY'' &= m\omega^2 Y + 2m\omega X' \end{aligned}$$

which are exactly Newton's laws with centrifugal and Coriolis forces!

Constraints and forces of constraint

An additional advantage of the Lagrangian formalism is that many constraints[20] can be incorporated into the Lagrangian without even thinking much about them. These same constraints would often be difficult to translate into forces in the Newtonian framework. Let's illustrate this with an example.

[20] Constraints are often divided into *holonomic* and *non-holonomic*. Holonomic constraints are geometric and can be expressed in terms of the coordinates $\mathbf{q}$ only, whereas non-holonomic constraints depend also on the velocities $\mathbf{q}'$. In general, incorporating non-holonomic constraints into the Lagrangian formalism is more involved (unless they can be integrated), and we will not consider this here.

Consider a mass m sliding down the surface of a frictionless hemisphere of radius R (Fig. 7). By using polar coordinates $x = R\sin\theta$ and $y = R\cos\theta$, the Lagrangian of the system can be written as[21]

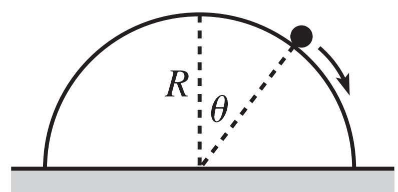

Figure 7: A mass m slides on the surface of a hemisphere of radius R.

[21] Exercise: Do it!

$$\mathcal{L} = \frac{m}{2}R^2\theta'^2 - mgR\cos\theta\,,$$

and by applying the Euler-Lagrange equation to θ we get the equation of motion

$$\theta'' = (g/R)\sin\theta\,.$$

Importantly, we solved the problem without paying any attention to the normal force that keeps the mass on top of the hemisphere. Rather, by a judicious choice of coordinates, we transformed the problem into a 1-dimensional problem and solved it easily. In other words, we dealt with the constraints of the problems simply by choosing the most suitable generalized coordinate.

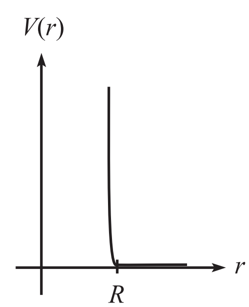

Figure 8: Potential simulating a hard surface. The potential is $V(r > R) = 0$ and becomes extremely large for $r < R$.

But what if we wanted to know the value of the CONSTRAINING NORMAL FORCE? To do this, let's solve the problem in a different way and write things in terms of both r and θ, as if r weren't exactly constrained to be R. Indeed, we assume that the particle pushes and sinks inward a tiny distance until the hemisphere gets squashed enough to push back with the appropriate force to keep the particle from sinking in any more.[22] The particle is therefore subject to a (very steep) potential $V(r)$ arising from the hemisphere's force (Fig. 8). Then, the Lagrangian for the system is[23]

[22] In fact, this is a much more accurate picture of what happens in the real world!

[23] Exercise: Do it!

$$\mathcal{L} = \frac{m}{2}\left(r'^2 + r^2\theta'^2\right) - mgr\cos\theta - V(r)$$

and the equations of motion (now for θ *and* r) become[24]

[24] Exercise: Do it!

$$mr^2\theta'' + 2mrr'\theta' = mgr\sin\theta$$
$$mr'' = mr\theta'^2 - mg\cos\theta - \frac{dV(r)}{dr}\,.$$

Now we can impose the constraint $r = R$, which implies $r' = r'' = 0$. The equation for θ reduces, as it should, to what we had obtained before, and the equation for r gives

$$-\left.\frac{dV}{dr}\right|_{r=R} = mg\cos\theta - mR\theta'^2$$

which is the normal force that we aimed to obtain.[25]

[25] When dealing with a system in which a nonconservative force such as friction is present, the Lagrangian method loses much of its appeal. The reason for this is that nonconservative forces don't have a potential energy associated with them, so there isn't a specific $V(x)$ that you can write down in the Lagrangian. Although friction forces can in fact be incorporated in the Lagrangian method, you have to include them in the Euler-Lagrange equations essentially by hand.

Method of Lagrange multipliers. Although illustrative, this method can become cumbersome in problems where the constraining force is not neatly aligned with one of the natural generalized coordinates. To finish this part, let's see how we can generalize the result above to a method that can be applied more broadly.

Let's continue with the same example, but use Cartesian coordinates. Start by defining a function $\eta(x, y) = \sqrt{x^2 + y^2} - R$ that measures how far we are from satisfying the constraint $\eta = 0$. Then, the Lagrangian of the system, including the constraining potential, is

$$\mathcal{L} = \frac{m}{2}(x'^2 + y'^2) - mgy - V(\eta(x, y))\,.$$

Now, when we use the Euler-Lagrange equations we get an extra term in the equations of motion (with respect to the case without constraint), equivalent to a "force"

$$-\frac{\partial V}{\partial \eta}\frac{\partial \eta}{\partial q_i} \equiv -\lambda\frac{\partial \eta}{\partial q_i}$$

with $q_i \in \{x, y\}$.

Now consider the following modified Lagrangian

$$\tilde{\mathcal{L}}(\mathbf{q}, \mathbf{q}', \lambda, t) = \mathcal{L}(\mathbf{q}, \mathbf{q}', t) + \lambda\eta(\mathbf{q}, t)$$

which is a Lagrangian which depends on the coordinates $\mathbf{q}$ *and* a new "coordinate" $\lambda(t)$ (but not on the corresponding "velocity" λ'). It is straightforward to show that the Euler-Lagrange equations for the $\mathbf{q}$ lead to[26]

[26] Exercise: Do it!

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial q_i'}\right) - \frac{\partial \mathcal{L}}{\partial q_i} = \lambda\frac{\partial \eta}{\partial q_i}\,,$$

which is exactly the Euler-Lagrange equation for the original Lagrangian with the extra force of constraint. Additionally, the Euler-Lagrange equation for the new "coordinate" $\lambda(t)$ leads to[27]

[27] Exercise: Do it!

$$\frac{d}{dt}\left(\frac{\partial \tilde{\mathcal{L}}}{\partial \lambda'}\right) - \frac{\partial \tilde{\mathcal{L}}}{\partial \lambda} = 0 \implies \eta(\mathbf{q}) = 0\,,$$

that is, to the satisfaction of the constraint. All in all, in general, when we have m (holonomic) constrains, we can impose all of them by working with the modified Lagrangian

$$\boxed{\tilde{\mathcal{L}}(\mathbf{q}, \mathbf{q}', \lambda, t) = \mathcal{L}(\mathbf{q}, \mathbf{q}', t) + \sum_{a=1}^{m} \lambda_a \eta_a(\mathbf{q}, t)}$$

or, equivalently, work directly with the constrained Euler-Lagrange equations

$$\boxed{\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial q_i'}\right) - \frac{\partial \mathcal{L}}{\partial q_i} = \sum_{a=1}^{m} \lambda_a \frac{\partial \eta_a}{\partial q_i}\,.}$$