# 1 Principle of stationary action

To specify a motion uniquely in classical mechanics, it suffices to give, at some time $t_0$, the initial positions and velocities $\mathbf{r}_i(t_0)$ and $\dot{\mathbf{r}}_i(t_0)$ for all point masses forming the system. Another formulation for the problem could be to only give the positions of all point masses in the system, at the initial time $t_1$ and final time $t_2$ of the motion under consideration.

To see why this is mathematically equivalent, consider first $t_2$ such that $\Delta t = t_2 - t_1 \to 0$. Specifying $\mathbf{r}_i(t_1)$ and $\mathbf{r}_i(t_2)$ is then tantamount to specifying $\mathbf{r}_i(t_1)$ and $\dot{\mathbf{r}}_i(t_1)$. By continuity, for $\Delta t$ sufficiently small, the motion remains unequivocally determined if one specifies $\mathbf{r}_i(t_1)$ and $\mathbf{r}_i(t_2)$. As $\Delta t$ grows arbitrarily, however, one can find several possible motions with the same end points. The *principle of stationary action*, sometimes also called – less accurately – the principle of least action, says that among all possible paths from $\mathbf{r}_i(t_1)$ and $\mathbf{r}_i(t_2)$, the physically realizable paths are the paths that extremize a functional called the *action* $S$ defined as follows

$$S[q_1, \ldots, q_N] = \int_{t_1}^{t_2} L(q_1(t), \ldots, q_N(t), \frac{dq_1}{dt}, \ldots, \frac{dq_N}{dt}, t) \tag{1}$$

where $L$ is called the *Lagrangian* of the system. Note that the functions $q_i(t)$ are *generalized coordinates*; they do not necessarily have to be Cartesian coordinates or other standard coordinates. $q_i$ does not have to have units of meter, and $dq_i/dt$ does not have to have the units of velocity.

*What the principle of stationary action is*

The principle of stationary action can be seen as a reformulation of the question of finding the proper evolution of a system in configuration space from a differential formulation to a variational/integral formulation. This has several advantages, which we briefly mentioned in the last lecture, and which we will highlight as we start to use the variational formulation in practical examples.

*What the principle of stationary action is not*

The principle of stationary action does not contain any new physics. All the physics is in finding appropriate Lagrangians do describe a system of interest. We will soon learn a simple method to construct Lagrangians for a large class of physical systems. However, there is no general method to construct a Lagrangian for every system. In fact, in new fields of physics, finding an appropriate Lagrangian to describe the dynamics of a system can be a major accomplishment in its own right.

# 2 Examples

Before we learn how to construct Lagrangians for a certain class of physical systems, let us first see explicit variational formulations for the first two simple examples we looked at in the previous lecture

## 2.1 Newton's apple

We saw that Newton's equations for Newton's apple are

$$\ddot{z} = -g \tag{2}$$

The Newtonian dynamics for the apple can be written in a variational form as follows. A Lagrangian for the free-fall of an object suject to gravity is

$$L(z, \frac{dz}{dt}, t) = \frac{1}{2}\left(\frac{dz}{dt}\right)^2 - gz \tag{3}$$

where the mass of the object has been set to $m = 1$ without any loss of generality. The action for the motion between times $t_1$ and $t_2$ thus is

$$S[z] = \int_{t_1}^{t^2} L(z, \frac{dz}{dt}, t)dt = \int_{t_1}^{t^2} \left[\frac{1}{2}\left(\frac{dz}{dt}\right)^2 - gz\right] dt$$

As we have seen in our review of the calculus of variation, extremizing the action with fixed endpoints $z(t_1)$ and $z(t_2)$ leads to the following Euler-Lagrange equation for $z(t)$:

$$\frac{d}{dt}\left(\partial_2 L(z, \frac{dz}{dt}, t)\right) = \partial_1 L(z, \frac{dz}{dt}, t) \tag{4}$$

We have

$$\partial_2 L(z, \frac{dz}{dt}, t) = \frac{dz}{dt} \quad , \quad \partial_1 L(z, \frac{dz}{dt}, t) = -g$$

So Equation (4) can be rewritten as

$$\frac{d^2 z}{dt^2} = -g$$

as expected

## 2.2 The simple pendulum

A Lagrangian for the simple pendulum is

$$L(\theta, d\theta/dt, t) = \frac{1}{2} l^2 \left(\frac{d\theta}{dt}\right)^2 + gl\cos\theta \tag{5}$$

Minimizing the action

$$S[\theta] = \int_{t_1}^{t_2} \left[\frac{1}{2} l^2 \left(\frac{d\theta}{dt}\right)^2 + gl\cos\theta\right] dt$$

leads to the E-L equations

$$\frac{d}{dt}\left(\partial_2 L(\theta, d\theta/dt, t)\right) = \partial_1 L(\theta, d\theta/dt, t) \tag{6}$$

We have

$$\partial_2 L(\theta, d\theta/dt, t) = l^2 \frac{d\theta}{dt} \quad , \quad \partial_1 L(\theta, d\theta/dt, t) = -gl\sin\theta$$

so Equation (6) takes the form

$$l^2 \ddot{\theta} = -gl\sin\theta \quad \Leftrightarrow \quad \ddot{\theta} + \frac{g}{l}\sin\theta = 0$$

which agrees with the second-order ODE we had previously obtained for that system.

# 3 Arbitrariness in the definition of the Lagrangian

Suppose a Lagrangian $L$ describes the dynamics of a system of point masses. Then, given the form of the E-L equation

$$\frac{d}{dt}\left(\partial_2 L\right) = \partial_1 L$$

it is clear that the Lagrangian $L' = AL + B$, where $A$ and $B$ are two constants, leads to the same equations of motion for the system. More interestingly, the equations of motion are also recovered if one replaces $L$ with the Lagrangian $L'$ defined by

$$L' = L + \frac{d}{dt} f(t, q(t)) \tag{7}$$

Indeed, if $S'$ is the action associated with the Lagrangian $L'$ and $S$ is the action associated with the Lagrangian $L$, then Equation (7) implies

$$S'[q] = S[q] + \int_{t_1}^{t_2} \frac{d}{dt} f(t, q(t)) dt = S[q] + f(t_2, q(t_2)) - f(t_1, q(t_1)) \tag{8}$$

The principle of stationary action states that the physically relevant trajectory in configuration space is obtained by extremals of the action holding the initial and final times and positions fixed. Since $t_1$, $t_2$, $q(t_1)$ and $q(t_2)$ are fixed, the extremals of $S'$ are the same as those of $S$, so both $L'$ and $L$ can be used to describe the motion of the system.

# 4 Hamilton's principle

## 4.1 Statement of the principle

An important class of problems in classical mechanics is the class of problems that involves *conservative forces* only. Conservative forces are forces that can be expressed as the gradient of a scalar function one often calls a potential, that depends only on time and the positions of the particles: $\mathbf{F} = -\nabla V(t, \mathbf{r}(t))$. For these situations, the construction of a Lagrangian is straightforward, and given by *Hamilton's principle*, which we state below:

A system of point masses for which the forces are derived from a potential energy that is independent of velocity evolves along a path $q$ for which the action

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

is stationary with respect to variations of the path $q$ that leave the endpoints fixed, **where $L = T - V$ is the difference between the kinetic energy $T$ and the potential energy $V$.**

## 4.2 Illustration

The examples provided in section 2 are illustrations of Hamilton's principle. Let us study one more important example.

Newton's law of gravitation tells us that two objects with masses $m$ and $M$ separated by a distance $R$ exert a force with magnitude

$$F_{\mathcal{G}} = G\frac{mM}{R^2}$$

on each other, and the force is an attractive force, directed along the line joining the centers of gravity of the two objects. $G \approx 6.67 \times 10^{-11} N.m^2.kg^{-2}$ is called the gravitational constant. We will consider the particular case in which $M \gg m$, so the object of mass $M$ is essentially at rest, and the lighter object, with mass $m$, orbits around the larger object. This is a very good approximation for the motion of the Earth around the sun, or satellites, natural (the Moon) or not, orbiting around Earth.

It easy to show that the motion of the object of mass $m$ is contained in a plane. If we take the center of gravity of the massive object as the origin of our coordinate system, we can write that the force acting on the light object is

$$\mathbf{F}_{\mathcal{G}} = -GmM\frac{\mathbf{r}}{|\mathbf{r}|^3}$$

Let us compute the quantity

$$\frac{d}{dt}\left(\mathbf{r} \times \dot{\mathbf{r}}\right) = \mathbf{r} \times \ddot{\mathbf{r}} = -\mathbf{r} \times GM\frac{\mathbf{r}}{|\mathbf{r}|^3} = \mathbf{0}$$

where the second equality was obtained by applying Newton's law. We conclude that

$$\mathbf{L}(t) = \mathbf{r} \times \dot{\mathbf{r}} = \overrightarrow{Cst} = \mathbf{L}(0) = \mathbf{r}(0) \times \dot{\mathbf{r}}(0) \tag{9}$$

This means that the motion of the object of mass $m$ is at all times in the plane perpendicular to $\mathbf{L}(0)$. The analysis of the motion of the object can thus be reduced to studying a system with two degrees of freedom.

Let us consider the polar coordinate system $(r, \theta)$ in the plane perpendicular to $\mathbf{L}(0)$, with the origin fixed at the center of gravity of the massive object. In this coordinate system, the force $\mathbf{F}_{\mathcal{G}}$ can be written as

$$\mathbf{F}_{\mathcal{G}} = -\nabla V$$

where

$$V = -\frac{GmM}{r} \tag{10}$$

is the gravitational potential. We can therefore use Hamilton's principle to write a Lagrangian for the object of mass $m$. The kinetic energy of the object is

$$T = \frac{m}{2}\left[\frac{dr}{dt}^2 + \left(r\frac{d\theta}{dt}\right)^2\right]$$

so a Lagrangian for the object is

$$L(r,\theta,\frac{dr}{dt},\frac{d\theta}{dt},t) = T - V = \frac{m}{2}\left[\frac{dr}{dt}^2 + \left(r\frac{d\theta}{dt}\right)^2\right] + \frac{GmM}{r} \tag{11}$$

The equations for the motion of the object are given by the E-L equations. For the $r$ coordinate, the E-L equation is

$$\frac{d}{dt}\partial_3 L(r,\theta,\frac{dr}{dt},\frac{d\theta}{dt},t) = \partial_1 L(r,\theta,\frac{dr}{dt},\frac{d\theta}{dt},t) \qquad \Leftrightarrow \frac{d^2r}{dt^2} = r\left(\frac{d\theta}{dt}\right)^2 - \frac{GM}{r^2} \tag{12}$$

The E-L equation for the $\theta$ coordinate is

$$\frac{d}{dt}\partial_4 L(r,\theta,\frac{dr}{dt},\frac{d\theta}{dt},t) = \partial_2 L(r,\theta,\frac{dr}{dt},\frac{d\theta}{dt},t) \qquad \Leftrightarrow \frac{d}{dt}\left(r^2\frac{d\theta}{dt}\right) = 0 \tag{13}$$

Equation (13) tells us that the quantity $r^2\frac{d\theta}{dt}$ is a conserved quantity of the motion. This quantity is called the angular momentum. It is the magnitude of the vector $\mathbf{L}$ introduced in Equation (9). We will soon see that this quantity is conserved for all potentials $V$ such that $V = V(|\mathbf{r}|)$, sometimes called *central potentials* (the associated forces are called *central forces*). Let us write

$$r^2\frac{d\theta}{dt} = \Gamma = Cst$$

Plugging $\Gamma$ into Equation (12), we have

$$\frac{d^2r}{dt^2} = \frac{\Gamma^2}{r^3} - \frac{GM}{r^2}$$

Multiplying this equation by $dr/dt$, one finds

$$\frac{d}{dt}\left[\frac{1}{2}\left(\frac{dr}{dt}\right)^2 + \frac{\Gamma^2}{2r^2} - \frac{GM}{r}\right] = 0$$

We conclude that

$$\frac{1}{2}\left(\frac{dr}{dt}\right)^2 + \frac{\Gamma^2}{2r^2} - \frac{GM}{r} = Cst = E \tag{14}$$

The total energy of the system is conserved: the first two terms correspond to the kinetic energy, the last term is the potential energy.

Equation (14) can be seen as the energy equation for a system with one degree of freedom, with an effective potential energy

$$V_{eff}(r) = \frac{\Gamma^2}{2r^2} - \frac{GM}{r}$$

This potential is dominated by the attractive gravitational part at large distances, but dominated by the repulsive angular momentum part at short distances. $V_{eff}(r)$ is shown in Figure 1 for the particular case $\Gamma = GM = 1$.

Since (14) can be rewritten as

$$\frac{1}{2}\left(\frac{dr}{dt}\right)^2 = E - V_{eff}(r)$$

point masses with $E < 0$ have bounded orbits: there exists $r_1$ in the attractive part of the potential such that $E = V_{eff}(r_1)$, and this $r_1$ is a turning point of the trajectory. On the other hand, point masses with $E > 0$ have unbounded orbits.

We will soon see that there are very efficient ways of finding out from the outset, from the functional dependence of the Lagrangian, that energy and angular momentum are conserved in the system studied here. We could have saved a bit of algebra. Still, it does not hurt too much to get practice with the dumb ways at first.

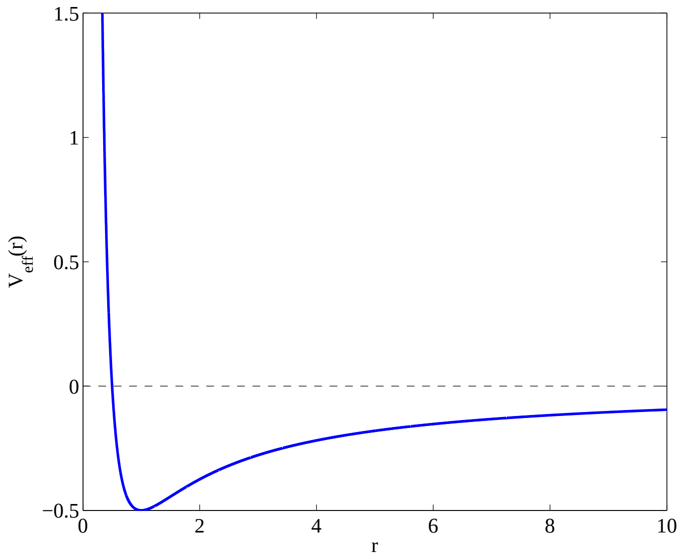


Figure 1: $V_{eff}(r)$ as a function of $r$ for $\Gamma = GM = 1$

***A little digression regarding the classical 2-body problem***

It turns out that any system consisting of two bodies with masses $m_1$ and $m_2$, at positions $\mathbf{r}_1(t)$ and $\mathbf{r}_2(t)$ and each exerting the force

$$\mathbf{F} = \pm G \frac{m_1 m_2}{|\mathbf{r}_1 - \mathbf{r}_2|^3} (\mathbf{r}_1 - \mathbf{r}_2)$$

on the other body can be studied within the framework we presented above, even if the masses $m_1$ and $m_2$ are comparable and none of the bodies can be considered static. Here is the reason why. Newton's equations for the bodies are

$$m_1 \ddot{\mathbf{r_1}} = G \frac{m_1 m_2}{|\mathbf{r}_2 - \mathbf{r}_1|^3} (\mathbf{r}_2 - \mathbf{r}_1) \tag{15}$$

$$m_2 \ddot{\mathbf{r_2}} = G \frac{m_1 m_2}{|\mathbf{r}_2 - \mathbf{r}_1|^3} (\mathbf{r}_1 - \mathbf{r}_2) \tag{16}$$

Let us define the position vectors

$$\mathbf{R} = \frac{m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2}{m_1 + m_2}$$

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$$

$\mathbf{R}$ is the position of the center of mass of the system. Adding Equations (15) and (16), one finds the following equation for $\mathbf{R}$:

$$\ddot{\mathbf{R}} = \mathbf{0} \tag{17}$$

This is expected. The system of the two bodies is isolated, so in the absence of external forces, the center of mass of the whole system has a constant velocity.

Subtracting $(16)/m_2$ from $(15)/m_1$, one finds a second order ODE for $\mathbf{r}$:

$$\mu \ddot{\mathbf{r}} = -G \frac{m_1 m_2}{|\mathbf{r}|^3} \mathbf{r} \tag{18}$$

with $\mu = m_1 m_2/(m_1 + m_2)$, which is often called the reduced mass. Equation (18) describes the motion of a particle of mass $\mu$ and position vector $\mathbf{r}$ with respect to a fixed center of force. This is precisely the situation we have studied with the Lagrangian formulation above.

The separate motion of the two bodies is obtained by solving for $\mathbf{r}$ (using either Newton's approach or a variational approach), and reconstructing $\mathbf{r}_1(t)$ and $\mathbf{r}_2(t)$ with the relations

$$\mathbf{r}_1 = \mathbf{R} + \frac{m_2}{m_1 + m_2}\mathbf{r}$$
$$\mathbf{r}_2 = \mathbf{R} - \frac{m_1}{m_1 + m_2}\mathbf{r}$$