Analytical mechanics

J. Avron

March 8, 2016

Plan of course

week	Topics
1	Lagrangian mechanics
2	Euler Lagrange equations
3	D'Alambert and Lorentz forces
4	Constraints
5	Noether theorem
6	Kepler problem
7	Small oscillations
8	Hamiltonian mechanics
9	Action, Hamilton Jacobi
10	Liouville theorem, Poisson brackets
11	Tops
12	Euler equations

1 Lagrangian mechanics

1.1 Reformulating Newtonian mechanics

Newton formulation of mechanics is

$$\dot{\mathbf{p}} = \mathbf{F}(\mathbf{x})$$

This is an equation of motion in vector form. If there are N interacting particles then the equation takes the form of N vectorial equation

$$\dot{\mathbf{p}}_j = \sum_k \mathbf{F}_{jk}(\mathbf{x}_j - \mathbf{x}_k)$$

where $\mathbf{F}_{jk}$ is the force the k-th particle applies on the j-th particle. Vectorial equations do not look too bad if you use Cartesian coordinates, where each particle position is given as

$$\mathbf{x}_j = (x_j, y_j, z_j)$$

and

$$\mathbf{p}_j = m_j \dot{\mathbf{x}}_j$$

Here are some reasons why one would want to reformulate Newton's laws in a more sbatraact way:

- Non-Cartesian coordinate systems: Suppose you wanted to use spherical coordinates. The position of the j-th particle is given in terms of spherical coordinates

  $$(r_j, \theta_j, \phi_j)$$

  It is not a trivial exercise to write Newton equations of motion in these coordinates. Of course, if you want to investigate planetary motions, such coordinates look like a good thing to have.

- Non-inertial coordinates: Sometimes it is convenient to use non-inertial coordinates. You may remember the pain you had with Coriolis which arises from Newton's equations in non-inertial coordinates. Wouldn't it be nice if you could write Newton equation in any coordinate system we want without this pain?

- Often just setting up Newton's equations is difficult, and even subtle. For example, in a double pendulum there are 4 forces: Gravity on the two bobs, and the forces of the rigid rod.

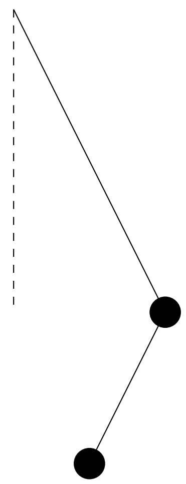


Figure 1: Double pendulum. Youtube demo of double pendulum

`http://www.youtube.com/watch?v=U39RMUzCjiU`

It would be nice to have an easier method.

Lagrangian mechanics reformulate Newtonian mechanics so you can use any coordinate you please. Even non-inertial ones. It encapsulates mechanics in a single scalar quantity from which we derives as many ordinary differential equations as we need. Let us see how this is done.

Let us start with one particle first. Consider conservative force, generated by a potential energy U

$$\mathbf{F} = -\nabla_x U$$

so the scalar potential energy U takes the place of the vectorial field $\mathbf{F}$. The position and velocity of a particle $(\mathbf{x}, \mathbf{v})$ are independent variables since I am free to choose the position and velocity at any given instant independently. Since the kinetic energy

$$T = \frac{m}{2}\mathbf{v} \cdot \mathbf{v}$$

is only aa function of $\mathbf{v}$ I can write

$$\mathbf{F} = \nabla_x L$$

where I defined a new object, the Lagrangian

$$L = T - U \tag{1.1}$$

as the difference of kinetic and potential energy.

Let us do something similar to $\mathbf{p}$. The momentum is the gradient of the kinetic energy:

$$\mathbf{p} = \nabla_v T$$

which I can write this in terms of L

$$\mathbf{p} = \nabla_v L$$

In summary, Newtotn's equations in the Lagrangian form is

$$\dot{\mathbf{p}} = \nabla_x L, \quad \mathbf{p} = \nabla_v L \tag{1.2}$$

This means that to set up the equations of motion, we only need to find the Lagrangian

$$L = L(\mathbf{v}, \mathbf{x}, t)$$

which is made from the kinetic and potential energies. All off these are are scalar quantities. We have now encapsulated the information about the dynamical system in one scalar function: The Lagrangian. L has the dimensions $[energy]$, but it is not quite the energy because it has the wrong relative sign between the kinetic and potential energy[1].

[1] If we would have picked the + sign in Eq. 1.1 we would pay with the price of a - sign in Eq. 1.2. In the next section, where we introduce the Euler-Lagrange equations, we shall see that Eq. 1.2 an be interpreted as an Euler Lagrange equation provided the sign is +.

You can now immediately see the gain by considering N particles. The Lagrangian is made from the sum of the kinetic energies[2] and potential energies

$$L = \sum_j \frac{m_j}{2} \mathbf{v}_j \cdot \mathbf{v}_j - \frac{1}{2} \sum_{jk} U_{jk}(\mathbf{x}_j - \mathbf{x}_k)$$

and the good old Newton's laws are simply

$$\boxed{\dot{\mathbf{p}}_j = \nabla_{\mathbf{x}_j} L, \quad \mathbf{p}_j = \nabla_{\mathbf{v}_j} L}$$

1.2 Generalized coordinates

One advantage of the Lagrangian formulation is that you may choose any coordinate system you want to write the kinetic and potential energy. It is ok to take Cartesian

$$(x, y, z)$$

but also just as well spherical coordinates

$$(r, \theta, \phi)$$

In Cartesian coordinates the velocities are

$$(\dot{x}, \dot{y}, \dot{z})$$

while in the spherical coordinates there are the velocities

$$(\dot{r}, \dot{\theta}, \dot{\phi})$$

Since we are not committed to any special coordinate system, we introduce a notation for generalized coordinates q_j and the generalized velocities by $\dot{q}_j$. If there are N generalized coordinates, then the Lagrangian is a function of N coordinates and N velocities:

$$L(q_1, \ldots, q_N, \dot{q}_1, \ldots, \dot{q}_N, t),$$

Definition 1.1 *The canonical momentum conjugate to the coordinate q_j is*

$$\boxed{p_j = \frac{\partial L}{\partial \dot{q}_j}}$$

If there are N generalized coordinates then Newton equations are the N ordinary differential equations:

$$\boxed{\dot{p}_j = \frac{\partial L}{\partial q_j}, \quad j = 1, \ldots, N}$$

These equations are called Euler-Lagrange equations.

[2] As viewed in an inertial frame.

Theorem 1.1 *Rescaling and shifting by constants $L \mapsto \lambda L + \mu$ does not affect the equations of motion.*

You may choose whatever units you please to measure energies in setting up the Lagrangian.

The freedom to choose whatever coordinates you want reflect the fact that L is a scalar. It does not care which coordinates you choose to represent the point and its velocity.

The Lagrangian has the advantages:

- You construct L from scalars–energies. This is easier and less confusing than vectors–forces.
- You construct the kinetic energy as if there is no potential energy and the potential energy–as if the system is frozen
- You can choose your favorite coordinates
- It is more economical because you need the minimal number of coordinates that specify the locations and velocities.

Let us look at some simpler examples:

1.3 Pendulum

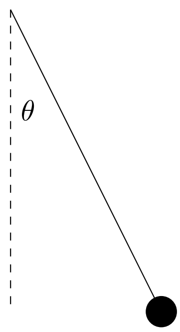


Figure 2: Pendulum

To construct the Lagrangian for a pendulum use the angle θ as general coordinate and choose the origin so that $\theta = 0$ is when the pendulum is down, and its potential energy minimal, then

$$L(\dot\theta, \theta) = \frac{m}{2}(\ell\dot\theta)^2 + mg\ell\cos\theta = m\ell^2\left(\frac{1}{2}\dot\theta^2 + \omega^2\cos\theta\right), \quad \omega^2 = \frac{g}{\ell}$$

Since an overall factor in the Lagrangian does not affect the equations of motion we may look at

$$L \mapsto \frac{1}{2}\dot\theta^2 + \omega^2\cos\theta, \quad \omega^2 = \frac{g}{\ell}$$

With this new Lagrangian

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

and Euler Lagrange is

$$\ddot\theta = \frac{\partial L}{\partial \theta} = -\omega^2 \sin\theta$$

Often the most interesting conclusions can be drawn by *not* solving the equation explicitly. Here are the some:

- The equation of motion depends on one parameter $\omega = \sqrt{\frac{g}{\ell}}$ with dimensions of frequency. This fixes the time scale in the problem.

- The velocity scale in the problem is $\omega\ell = \sqrt{g\ell}$. Our legs, and those of elephants, move like pendulums. (Actually, more like a double pendulum, but lets forget this for a moment.) An animal that is 4 times bigger moves its legs with half the frequency, but has twice the speed. Elephant and bears are faster than humans, they will outrun you.

- When θ is small

  $$\sin\theta \approx \theta$$

  and you get the Harmonic oscillator

  $$\ddot\theta = -\omega^2\,\theta$$

  with the elementary solution

  $$\theta = A\cos(\omega t + \phi)$$

  A and ϕ are constant of integration fixed by the two initial conditions.

You can appeal to Mathematica for help with solving exactly the differential equation of a pendulum:

$$\texttt{DSolve}[\theta''[t] == -\omega^2 \sin[\theta[t]], \theta[t], t]$$

It will tell you

$$\left\{\theta[t] \to \pm\texttt{JacobiAmplitude}\left[\frac{1}{2}\sqrt{(2\omega^2 + C[1])\,(t + C[2])^2}, \frac{4\omega^2}{2\omega^2 + C[1]}\right]\right\}$$

where $C[1]$ and $C[2]$ are integration constants which are determined by the two initial conditions. The trouble with this is that most people have no idea how does a JacobiAmplitude looks. Of course you can ask Mathematica to plot it for you. The moral of this little story is that insight and exact solutions are not the same thing.

1.4 Lagrangian in polar coordinates

The Euclidean plane can be equally described in terms of Euclidean or polar coordinates. The distance between two nearby points is

$$ds^2 = (dx)^2 + (dy)^2 = (dr)^2 + r^2(d\theta)^2$$

The (non-relativistic) classical Lagrangian of a free particle is

$$L(\mathbf{v}) = \frac{m}{2}\mathbf{v}^2 = \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2\right) = \underbrace{\frac{m}{2}\dot{r}^2}_{radial} + \underbrace{\frac{m}{2}r^2\dot{\theta}^2}_{azymuthal} = L'(\dot{\theta}, \theta) \tag{1.3}$$

The momenta are

$$\mathbf{p} = m\mathbf{v} \Longrightarrow \underbrace{p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r},}_{radial\ momentum} \quad \underbrace{p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta}}_{angular\ momentum}$$

The Euler-Lagrange equations for a free particle say that momentum is conserved. In the Euclidean coordinates this interpretation is what you are used to. In the polar coordinates the $\hat{\theta}$ term is what you would naturally call conservation of angular momentum.

Problem 1.1 *Write the Euler Lagrange equations for a free particle in a central potential $U(r)$ in polar coordinates.*

$$\dot{p}_r = \frac{\partial L}{\partial r} = mr\dot{\theta}^2 - \frac{\partial U}{\partial r}, \quad \dot{p}_\theta = \frac{\partial L}{\partial \theta} = 0$$

The first equation is a mess. The second says the angular momentum in a radial potential is conserved.

2 Euler-Lagrange equations

We shall now show that Lagrange equations admit an elegant interpretation as the equations that guarantee the optimization of a mechanical quantity.

2.1 Optimizing a function: critical points

Stationary (aka critical) points of a function $f(x_1, \ldots, x_N)$, are (local) optimizers, given by

$$\delta f = \sum_1^N \frac{\partial f}{\partial x_j}\delta x_j = 0$$

If the δx_j can be varied independently this implies that

$$\frac{\partial f}{\partial x_j} = 0, \quad j = 1, \ldots, N$$

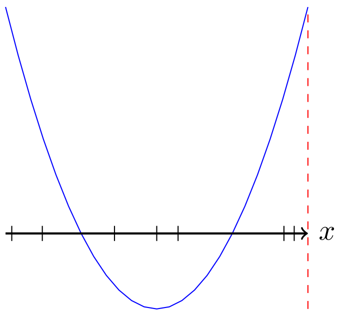


Figure 3: To identify the critical point you do not need to know how points on the x-axis are labeled–where you put the ticks. This reflects the fact that the characterization of a critical point does not care about the choice of coordinates. Note that if the independent variable x is constrained, the optimum may lie on the boundary, where the variation need not vanish.

At a critical point the variation of δx_j does not change f to first order.

The characterization of a point as critical is geometric and does not care what coordinate system you use: If you replace the coordinates x_j by y_j:

$$y_j = y_j(x_1, \dots, x_n)$$

the function f will have a different shape, but the critical *points* (not their coordinates) remain the same. This follows from the chain rule

$$\delta f = \sum_k \delta x_k \underbrace{\frac{\partial f}{\partial x_k}}_{=0} = \sum_k \delta x_k \sum_j \left(\frac{\partial y_j}{\partial x_k}\right) \frac{\partial f}{\partial y_j} = \sum_j \delta y_j \underbrace{\frac{\partial f}{\partial y_j}}_{=0} \tag{2.1}$$

2.2 Physical laws and optimizers

Certain physical laws can be formula at the solution to a variational problem. Perhaps the most famous one is Snell law. Snell law says that when light crosses the boundary between two media with different index of reflections, $n_{1,2}$, it changes its direction according to the rule

$$n_1 \sin\phi_1 = n_2 \sin\phi_2$$

This formulation is much less elegant and general than Fermat formulation, which is also easier to remember. To formulate Fermat principle recall that the meaning of n, is that it determines the velocity of light to be c/n.

Theorem 2.1 (Fermat principle) *The path of light between two points a, b minimizes the time it takes to get from a to b.*

Exercise 2.1 *Derive Snell's law from Fermat principle.*

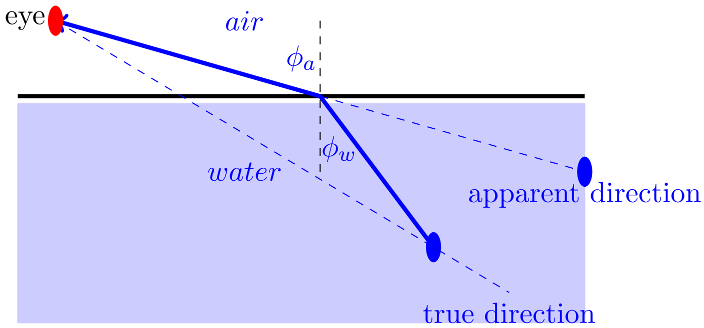


Figure 4: Fermat principle gives an interpretation of Snell law as an optimizer: Light follows the same rules as Waze: Minimize ETA. The path is then not necessarily the shortest in length, but the shortest in time. This is why a spoon in a glass filled with water looks broken.

Actually Fermat principle is stronger than Snell law because it also implies:

Exercise 2.2 *Show that Fermat principle implies that the reflection angle θ_r equals the incidence angle θ_i.*

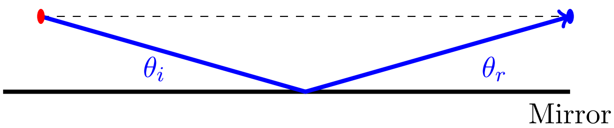


Figure 5: Fermat principle gives the rule that the reflection angle equals the angle of incidence. This is an example where the path is constrained to lie in a half space. There are then two optimizers: The dashed direct path is the global minimizer of the time of flight between the two points. The reflected pat is an optimizer that lits the boundary: The length of the path is a local minimum for moving the hitting point on the boundary. The length decreases if the hitting point moves away from the boundary.

We want to formulate a similar principle in mechanics. For this let me consider first an example that illustrates the relation between ODE and optimization problems.

The straight line between two points in Euclidean space can be characterized in two different ways:

- As a differential equation:
$$\frac{d\hat{\mathbf{v}}}{dt} = 0$$
expressing the fact that the direction $\hat{\mathbf{v}}$ is fixed.
- As the shortest path connecting the two points

We shall now build a machinery that will allow us to express the solutions to certain optimization problems as ODE. These ODE's are known as Euler-Lagrange equations.

2.3 Path space

Think of a curve in the plane connecting a and b. Such a curve γ, can be represented parametrically by a pair of functions

$$\{x(t), y(t)\}, t \in [0,1]$$

so that $\{x(0), y(0)\} = a, \quad \{x(1), y(1)\} = b$ The length of the curve, by Pythagoras, is

$$\ell(\gamma) = \int_0^1 \underbrace{\sqrt{\dot{x}^2(t) + \dot{y}^2(t)}}_{|\mathbf{v}|}\, dt$$

We are interested in that curve γ_0 that minimizes the length. This is a variational problem whose solution is a function–a straight line. Our aim will be to derive the ODE that characterizes the optimizer.

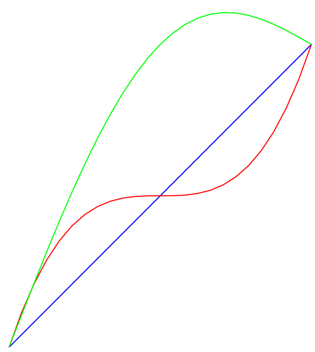


Figure 6: Three different paths between two points

Before we find the ODE let me stress the rules:

- Fix initial and final time
- Fix initial and final position
- Vary the path continuously but otherwise arbitrarily.

2.4 Euler Lagrange equations

Let $L(\mathbf{v}, \mathbf{x}, t)$ be a function of velocity, position and time. For a path $\gamma = \{\mathbf{x}(t) | t \in [0,T]\}$ connecting two points a and b in time T we associate the action

$$\underbrace{S(\gamma)}_{action} = \int_0^T dt\, L(\mathbf{v}, \mathbf{x}, t)$$

The action is a function of orbit: It is a function whose argument is a function. Such functions are sometimes called *functionals*. The dimension of the action is

$$[S] = [Energy][sec]$$

Let us find the condition that γ optimizes the action. Let $\mathbf{x}(t)$ be a path and let $\delta\mathbf{x}$ be a small variation of the path that keeps the initial and final points fixed:

$$\delta\mathbf{x}(0) = \delta\mathbf{x}(T) = 0 \tag{2.2}$$

Then upon change of the path the integrand changes by

$$\begin{aligned}
\delta L &= \delta\mathbf{v} \cdot \frac{\partial L}{\partial \mathbf{v}} + \delta\mathbf{x} \cdot \frac{\partial L}{\partial \mathbf{x}} \\
&= \delta\mathbf{v} \cdot \mathbf{p} + \delta\mathbf{x} \cdot \frac{\partial L}{\partial \mathbf{x}} \\
&= \frac{d(\delta\mathbf{x})}{dt} \cdot \mathbf{p} + \delta\mathbf{x} \cdot \frac{\partial L}{\partial \mathbf{x}} \\
&= \frac{d(\delta\mathbf{x} \cdot \mathbf{p})}{dt} - \delta\mathbf{x} \cdot \left(\frac{d\mathbf{p}}{dt} - \frac{\partial L}{\partial \mathbf{x}}\right)
\end{aligned} \tag{2.3}$$

Hence

$$\delta S = (\delta\mathbf{x}) \cdot \mathbf{p}\big|_0^T - \int_0^T dt\, \delta\mathbf{x} \cdot \left(\frac{d\mathbf{p}}{dt} - \frac{\partial L}{\partial \mathbf{x}}\right)$$

The first term is a boundary term. It vanishes by 2.2. If the variations δS $\delta\mathbf{x}$ are linearly independent then $\delta S = 0$ implies the vanishing of the brackets in the integrand:

$$\boxed{\dot{\mathbf{p}} = \frac{\partial L}{\partial \mathbf{x}}, \quad \mathbf{p} = \frac{\partial L}{\partial \mathbf{v}}} \tag{2.4}$$

These are the Euler-Lagrange equations.

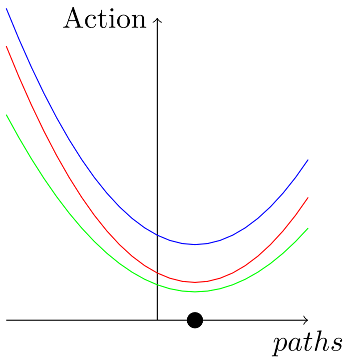


Figure 7: The minimizer of all three curves is the black dot representing a path. The blue curve is a shift of the red, $S \mapsto S + \lambda$, and the green curve is scaling of the red, $S \mapsto \lambda S$. Scaling and shifts do not affect the minimizing point. (But may affect the value of the action at the minimizer.)

Example 2.1 (Particle in an scalar potential) *Let*

$$L = \frac{m}{2}\mathbf{v}^2 - U(\mathbf{x})$$

then $\mathbf{p} = m\mathbf{v}$ *and Euler-Larange equations give Newton law*

$$\dot{\mathbf{p}} = -\nabla U(\mathbf{x})$$

Exercise 2.3 *Write the Euler-Lagrange equations as an equation for the acceleration*

$$\left(\dot{\mathbf{v}} \cdot \frac{\partial}{\partial \mathbf{v}}\right)\frac{\partial L}{\partial \mathbf{v}} + \left(\mathbf{v} \cdot \frac{\partial}{\partial x}\right)\frac{\partial L}{\partial \mathbf{v}} + \frac{\partial^2 L}{\partial t \partial \mathbf{v}} = \frac{\partial L}{\partial \mathbf{x}}$$

2.5 Relativistic Lagrangian

A free relativistic particle has the Lagrangian

$$L(\mathbf{v}) = -mc^2\sqrt{1 - (\mathbf{v}/c)^2}$$

The allowed paths have $|\mathbf{v}| < c$. Paths with $\mathbf{v} \geq c$ are unphysical (L takes imaginary values). Taylor expansion for small velocities gives

$$L = -mc^2\sqrt{1 - (\mathbf{v}/c)^2} \approx \underbrace{-mc^2}_{constant} + \frac{m}{2}\mathbf{v}^2 + \dots$$

The first term, even though large, is a constant and so does not affect the equations of motion. The second is the non-relativistic kinetic energy.

Problem 2.1 *Show that the relativistic momentum is*

$$\mathbf{p} = m\gamma\mathbf{v}, \quad \gamma = \frac{1}{\sqrt{1 - (\mathbf{v}/c)^2}}$$

Note that while the velocity is b bounded by the speed of light, $|\mathbf{v}| < c$*, the momentum* $\mathbf{p}$ *is unbounded.*

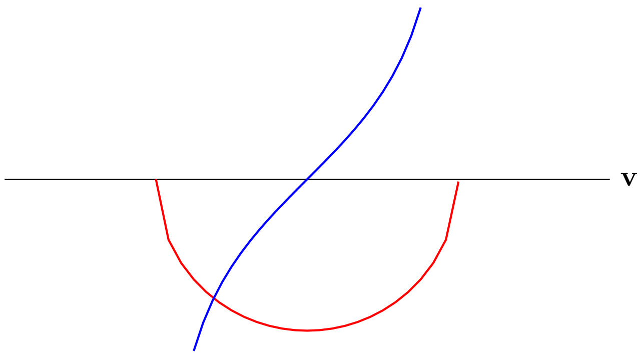


Figure 8: The red curve shows $L(\mathbf{v})$ for a relativistic particle. The graph is a semi-circle. The blue curve is the momentum. Both curves are defined only for $\mathbf{v}^2 \leq c^2$.

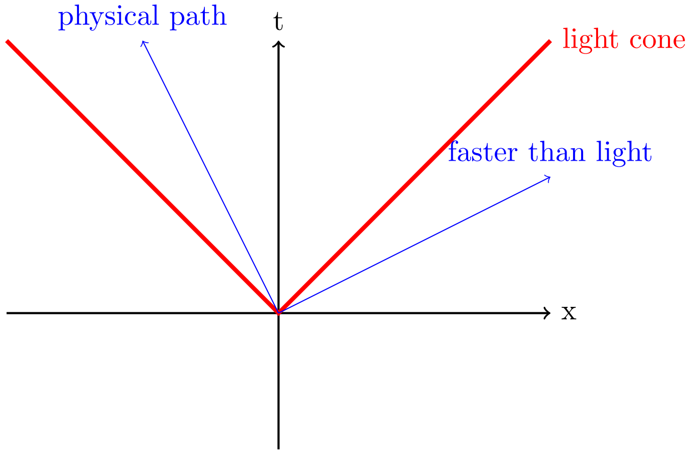


Figure 9: In a relativistic Lagrangian not you can connect the initial point $(\mathbf{x}_0, t_0)$ with the final point $(\mathbf{x}_f, t_f)$ with a physical path provided the end point lies in the forward light cone of the initial point.

2.6 Gauge freedom

Different Lagrangians can give the same Euler-Lagrange equations. Lets consider this ambiguity in more detail.

Theorem 2.2 *ITwo lagrangians that differ by a complete derivative of $\Lambda(q_j, t)$ give identical Euler-Lagrange equations:*

Proof: Since

$$L'(\dot{q}_j, q_j, t) - L(\dot{q}_j, q_j, t) = \frac{d\Lambda(q_j, t)}{dt}$$

we have that

$$S' - S = \underbrace{\Lambda(q_j(T), T) - \Lambda(q_j(0), 0)}_{\textit{boundary term}}$$

Since the paths are all anchored at the boundary the rhs is independent of the variation of the path. The two Lagrangians have the same stationary points and the same equations of motion.

Adding to the Lagrangian a complete derivative of a function $\Lambda(\mathbf{x}, t)$ is called a gauge transformation

$$\boxed{L \to L + \frac{d\Lambda(\mathbf{x}, t)}{dt}}$$

Gauge transformations affect the momentum

$$\boxed{\mathbf{p} \to \mathbf{p} + \nabla_x \Lambda}$$

This follows from:

$$\frac{d\Lambda(\mathbf{x}, t)}{dt} = (\mathbf{v} \cdot \nabla_x)\Lambda + \partial_t \Lambda$$

You see that, in a general gauge, the momentum need not point in the direction of $\mathbf{v}$.

2.7 Generalized coordinates

Suppose we choose generalized coordinate q_j, not necessarily Cartesian, not even coordinate in an inertial system. But coordinates nonetheless. This means that

$$\mathbf{x} = \mathbf{x}(q_j, t), \quad \mathbf{v} = \sum \dot{q}_j \frac{\partial \mathbf{x}}{\partial q_j} + \frac{\partial \mathbf{x}}{\partial t}$$

The Lagrangians, being scalars, take the same value at the same poitns, i.e. are related by

$$L'(q_j, \dot{q}_j, t) = L(\mathbf{v}(q_j), \mathbf{x}(q_j), t) = L(\mathbf{v}, \mathbf{x}, t)$$

(The prime indicates that the functional form is different.) The conjugate momenta to q_j are

$$\pi_j = \frac{\partial L'}{\partial \dot{q}_j} = \frac{\partial L}{\partial \mathbf{v}} \cdot \frac{\partial \mathbf{v}}{\partial \dot{q}_j} = \mathbf{p} \cdot \frac{\partial \mathbf{x}}{\partial q_j}$$

The action S does not care which coordinate you use

$$S = \int_0^T L(\mathbf{v}, \mathbf{x}, t) dt = \int_0^T L'(q_j, \dot{q}_j, t) dt$$

It is evident that the optimal value of S does not care which coordinates you use. In particular, the equation

$$\delta S = 0$$

holds in either coordinate system as long (as δq_j are linearly independent). This shows that Euler Lagrange equations have the same form in any coordinate sysstem

$$\frac{d\pi_j}{dt} - \frac{\partial L'}{\partial q_j} = 0, \quad \pi_j = \frac{\partial L'}{\partial \dot{q}_j}$$

In this sense, Euler-Lagrange equations are better than Newton's equations which change their form in different coordinate: In Cartesian coordinates the equation Newton equation is $m\ddot{x} = F_x$ but in polar coordinates $m\ddot{r} \neq F_r$.

2.8 D'Alambert principle

Here is a nice application of the freedom to choose non-inertial coordinates and gauge freedom: D'Alambert principle.

Consider two frames S', inertial, and S non inertial moving with arbitrary relative velocity $\mathbf{v}_0(t)$. The (non-relativistic) velocities in the two frames are related by

$$\mathbf{v}' = \mathbf{v} - \mathbf{v}_0(t)$$

The kinetic energy in an inertial frame expressed in the primed coordinates is

$$\begin{aligned}
L' = \frac{m}{2}(\mathbf{v}')^2 &= \frac{m}{2}(\mathbf{v} - \mathbf{v}_0)^2 \\
&= \frac{m}{2}\left(\mathbf{v}^2 - 2\mathbf{v}\cdot\mathbf{v}_0 + \mathbf{v}_0^2\right) \\
&= \frac{m}{2}\mathbf{v}^2 + m(\mathbf{x}\cdot\dot{\mathbf{v}}_0) - m\frac{d(\mathbf{x}\cdot\mathbf{v}_0)}{dt} + \frac{m}{2}\mathbf{v}_0^2 \\
&= \frac{m}{2}\mathbf{v}^2 + m(\mathbf{x}\cdot\dot{\mathbf{v}}_0)\ \underbrace{-m\frac{d(\mathbf{x}\cdot\mathbf{v}_0)}{dt} + \frac{m}{2}\mathbf{v}_0^2(t)}_{gauge\ transformation}
\end{aligned}$$

Exercise 2.4 *Explain why the last term is a gauge transformation*

In the non-inertial frame the Lagrangina is

$$\boxed{L = \frac{m}{2}\mathbf{v}^2 - \mathbf{F}(t)\cdot\mathbf{x}, \quad \mathbf{F}(t) = -m\dot{\mathbf{v}}_0(t)} \tag{2.5}$$

The extra term besides the kinetic energy gives D'Alambert fictitious force $\mathbf{F}$.

It follows that you can get rid of a constant gravitational force by moving into an accelerating frame. This observation turns out to be the key to incorporating gravity with relativity. Einstein called this "his happiest thought".

2.9 Coriolis

The earth is not an inertial frame, it is rotating about the sun and about its axis. It is more convenient to use coordinates fixed on earth and the price is Coriolis force, which is the terror of Physics 1M. Lagrangian mechanics makes Coriolis painless. (youtube demo 1, 2).

Let S' be the inertial frame and S one that rotates by $\boldsymbol{\omega}$ (possibly a function of time). If an object is at rest in the S at position $\mathbf{x}$, its velocity in S' frame is $\boldsymbol{\omega}\times\mathbf{x}$. By the rule of addition of (non-relativistic) velocities, the relation between the velocities is

$$\mathbf{v}' = \mathbf{v} + \boldsymbol{\omega}\times\mathbf{x}$$

The kinetic energy in S' is

$$L' = \frac{m}{2}(\mathbf{v}')^2 = \frac{m}{2}(\mathbf{v} + \boldsymbol{\omega} \times \mathbf{x})^2$$
$$= \underbrace{\frac{m}{2}\mathbf{v}^2}_{T} + m\mathbf{v} \cdot \boldsymbol{\omega} \times \mathbf{x} + \underbrace{\frac{m}{2}(\boldsymbol{\omega} \times \mathbf{x})^2}_{-U}$$

In the rotating frame the Lagrangian has the standard kinetic energy T, but also a potential energy

$$U = -\frac{m}{2}(\omega \times \mathbf{x})^2$$

which we identify with the centrifugal potential. But there is also an additional term which is linear $\mathbf{v}$.

The velocities in the two frames are, of course, different, but the canonical momenta are the same:

$$\mathbf{p}' = \frac{\partial L'}{\partial \mathbf{v}'} = m\mathbf{v}' = \mathbf{p} = \frac{\partial L'}{\partial \mathbf{v}} = m\left(\mathbf{v} + \boldsymbol{\omega} \times \mathbf{x}\right)$$

I used the rule

$$\nabla_{\mathbf{v}}(\mathbf{a} \cdot \mathbf{v}) = \mathbf{a}$$

In the inertial frame $\mathbf{p}' \parallel \mathbf{v}'$, but not true in the rotating frame. The canonical momenta and velocities need not be parallel.

We allow $\boldsymbol{\omega}$ to be time dependent. Euler-Lagrange equations are then

$$\dot{\mathbf{p}} = m\dot{\mathbf{v}} + m\boldsymbol{\omega} \times \mathbf{v} + m\dot{\boldsymbol{\omega}} \times \mathbf{x} = \frac{\partial L}{\partial \mathbf{x}} = \nabla_x L$$

We need to take the gradient of

$$\left(\mathbf{v} + \boldsymbol{\omega} \times \mathbf{x}\right)^2 = \left(\boldsymbol{\omega} \times \mathbf{x}\right)^2 + \underbrace{2\left(\boldsymbol{\omega} \times \mathbf{x}\right) \cdot \mathbf{v}}_{-2\mathbf{x}\cdot(\boldsymbol{\omega}\times\mathbf{v})} + \mathbf{v}^2 \tag{2.6}$$

(I used the fact that the dot and cross in the triple product can be interchanged.) For the force terms we get

$$\nabla_x L = \underbrace{\frac{m}{2}\nabla(\boldsymbol{\omega} \times \mathbf{x})^2}_{centrifugal} - \underbrace{m\boldsymbol{\omega} \times \mathbf{v}}_{half\ Coriolis}$$

The Euler-lagrange in the rotating frame is then

$$m\dot{\mathbf{v}} + m\boldsymbol{\omega} \times \mathbf{v} + m\dot{\boldsymbol{\omega}} \times \mathbf{x} = \frac{m}{2}\nabla(\boldsymbol{\omega} \times \mathbf{x})^2 + m\boldsymbol{\omega} \times \mathbf{v}$$

Assembling the fictitious forces on the right we find

$$m\ddot{\mathbf{x}} = \underbrace{\frac{m}{2}\nabla(\boldsymbol{\omega} \times \mathbf{x})^2}_{centrifugal\ force} - \underbrace{2m\boldsymbol{\omega} \times \mathbf{v}}_{Coriolis} - \underbrace{m\dot{\boldsymbol{\omega}} \times \mathbf{x}}_{Euler} \tag{2.7}$$

The last term on the tight is called after Euler, it is a force due to angular acceleration

Example 2.2 (Foucault pendulum) *Foucault*

2.10 The Levi-Civita tensor

A scalar s is an object with a single component s. A vector $\mathbf{v}$ in 3 dimensions has 3 components (v_1, v_2, v_3). A matrix M in 3 dimensions has 9 components M_{jk}. We can put all of these objects, and more in a single framework: Tensors. A tensors of rank 0 has a single component; A tensor of rank 1 in d dimensions has d components; A tensor of rank 2 has d^2 components, etc. So a general tensor has the form

$$T_{j\ldots,k}$$

Definition 2.1 *Einstein convention: A pair of identical indices are summed.*

For example, a vector can be written in shorthand form as

$$\mathbf{v} = \sum v^j \mathbf{e}_j = v^j \mathbf{e}_j \tag{2.8}$$

v^j are the components and $\mathbf{e}_j$ are the basis vectors.

We can use the basis vectors (of Cartesian coordinates in Euclidean space) to construct a very simple tensor:

$$\mathbf{e}_j \cdot \mathbf{e}_k = \delta_{jk}$$

This is the identity tensor. Clearly

$$Tr\mathbb{1} = \delta_{jj} = d$$

In d dimensions, we can use the basis vector to construct an interesting rank d tensor whose components are the volumes spanned by d basis vectors. This is the Levi-Civita tensor. In particular, in 3 dimensions:

Definition 2.2 *The Levi-Civita tensor in 3 dimensions is a 3-rd rank tensor whose components in a Euclidean space are given by*

$$\varepsilon_{ijk} = \mathbf{e}_i \times \mathbf{e}_j \cdot \mathbf{e}_k$$

where $\mathbf{e}_j$ are the unit basis vectors of any Cartesian coordinate system.

Explicitly

$$\varepsilon_{ijk} = \begin{cases} 1 & \text{if ijk=123, 231, 312} \\ -1 & \text{if ijk=321, 132, 213} \\ 0 & \text{otherwise} \end{cases} \tag{2.9}$$

The Levi-Civita tensor is useful for making computations with cross products. For example

$$(\mathbf{a} \times \mathbf{b})_i = \sum_{jk} \epsilon_{ijk} a_j b_k, \quad (\nabla \times A)_i = \epsilon_{ijk} \partial_j A_k,$$

Theorem 2.3
$$\epsilon_{ijk}\epsilon_{i\ell m} = \delta_{j\ell}\delta_{km} - \delta_{k\ell}\delta_{jm} \tag{2.10}$$

Problem 2.2 *Show*
$$v \times (\nabla \times \mathbf{A}) = \nabla(v \cdot \mathbf{A}) - (v \cdot \nabla)\mathbf{A}$$

Exercise 2.5 *Show that*
$$\nabla(\omega \times \mathbf{x})^2 = -2\omega \times (\omega \times \mathbf{x})$$

2.11 Galilei invariance

The Lagrangian of free (non-relativistic) particle is quadratic in the velocities
$$L(\mathbf{v}) = \frac{m}{2}\mathbf{v} \cdot \mathbf{v}$$
Quadratic Lagrangians are special: In a primed inertial frame moving at $\mathbf{v}_0$ the velocity $\mathbf{v}'$ is
$$\mathbf{v} = \mathbf{v}' + \mathbf{v}_0$$
The Lagrangians in the two frames, being a scalar, are related by:
$$L(\mathbf{v}) = L'(\mathbf{v}') = L(\mathbf{v}' + \mathbf{v}_0)$$
The two Lagrangians take the same values but have different functional forms:
$$L'(\mathbf{v}') = L(\mathbf{v}' + \mathbf{v}_0) = \frac{m}{2}(\mathbf{v}' + \mathbf{v}_0)^2 \tag{2.11}$$
The value of the momentum in the two frames are then also the same, but have different functional form:
$$\mathbf{p}' = m(\mathbf{v}' + \mathbf{v}_0) = \mathbf{p}$$
Quadratic Lagrangian are special since:
$$L'(\mathbf{v}') = L(\mathbf{v}') + \frac{d\Lambda}{dt}, \quad \Lambda = \frac{m}{2}(2\mathbf{x}' \cdot \mathbf{v}_0 + \mathbf{v_0}^2 t) \tag{2.12}$$
This allows us to pick $L = \frac{m}{2}\mathbf{v}^2$ in any inertial frame, even though the kinetic energies in different frames are different. This miracle is Galilei invariance.

Theorem 2.4 *The solution of Euler-Lagrange equations for a free particle gives the the absolute minimum of the action.*

Proof: The solution of Euler-Lagrange equations for a free particle is
$$\mathbf{x}(t) = \mathbf{x}_0 + \mathbf{v}_0 t$$
Choose a frame where $\mathbf{v}_0 = 0$ so the particle is at rest. By Galilei invarince we may choose the Lagrangian in the frame where the particle is at rest to be $L = \frac{m}{2}\mathbf{v}^2 \geq 0$. The action for any path is then non-negative and for the solution it is 0 which is clearly the absolute minimizer.

Exercise 2.6 *Show that the Lagrangian with pair potential*

$$L = \frac{m_1}{2}\mathbf{v}_1^2 + \frac{m_2}{2}\mathbf{v}_2^2 - U(\mathbf{x}_1 - \mathbf{x}_2)$$

is Galilei invariant.

2.12 Lorentz and Coulomb forces

Newton equation for a charges particle in and electric and magnetic field is

$$\boxed{m\dot{\mathbf{v}} = e\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right)} \tag{2.13}$$

We want to construct the Lagrangian that gives the equation as its Euler-Lagrange equation.

In the absence of magnetic fields our prescription was

$$L(\mathbf{v}, \mathbf{x}) = \frac{m}{2}\mathbf{v}^2 - e\Phi(\mathbf{x}, t)\big), \quad \mathbf{B} = 0$$

where $e\Phi(\mathbf{x}, t)$ is the electric potential. How shall we accommodate magnetic fields?

The clue come from gauge transformations. Recall that fields are related to the potentials by:

$$\mathbf{B} = \nabla \times \mathbf{A}, \quad -\mathbf{E} = \partial_t \mathbf{A} + \nabla\Phi \tag{2.14}$$

Gauge transformations

$$\mathbf{A} \mapsto \mathbf{A} + \nabla\Lambda, \quad \Phi \mapsto \Phi - \partial_t \Lambda \tag{2.15}$$

leave $\mathbf{E}$ and $\mathbf{B}$ invariant. Now note that

$$\frac{d\Lambda}{dt} = \mathbf{v} \cdot \nabla\Lambda + \frac{\partial\Lambda}{\partial t}$$

It follows that

$$\mathbf{v} \cdot \mathbf{A}(\mathbf{x}, t) - \Phi(\mathbf{x}, t)$$

changes by a complete derivative under gauge transformation. This means that the Lagrangian is gague invariant if

$$\boxed{L(\mathbf{v}, \mathbf{x}) = \frac{m}{2}\mathbf{v}^2 + e\big(\mathbf{v} \cdot \mathbf{A}(\mathbf{x}, t) - \Phi(\mathbf{x}, t)\big)}$$

This Lagrangian gives Lorentz and Coulomb forces correctly. The Lagrangian is not simply the difference between kinetic and potential energies.

Theorem 2.5 *The canonical momentum* $\mathbf{p}$ *in external electromagnetic fields is*

$$\boxed{\mathbf{p} = m\mathbf{v} + e\mathbf{A}(\mathbf{x}, t)}$$

Note that the momentum and the velocity are, in general, not parallel.

Example 2.3 (Charged particle in constant electric and magnetic fields)

2.13 Geodesics

The shortest path between two points is called geodesic. In Euclidean space geodesics are straight lines. When you study geodesics you want to minimize the length of the path. The geodesic equation is the Euler-Lagrange equations for the Lagrangian

$$d\ell = |\mathbf{v}|dt, \quad \underbrace{\ell}_{length} = \int_0^T d\ell$$

Since

$$\nabla_{\mathbf{v}}(\mathbf{v}\cdot\mathbf{v}) = 2\mathbf{v}$$

Exercise 2.7 *Use* $\frac{\partial v_j}{\partial v_k} = \delta_{jk}$ *and* $(\mathbf{v}\cdot\mathbf{v}) = \sum v_j v_j$ *to show this.*

It follows that

$$L = |\mathbf{v}| = (\mathbf{v}\cdot\mathbf{v})^{1/2} \Longrightarrow \nabla_{\mathbf{v}} L = \frac{1}{2}\frac{\nabla_{\mathbf{v}}(\mathbf{v}\cdot\mathbf{v})}{|\mathbf{v}|} = \frac{\mathbf{v}}{|\mathbf{v}|} = \hat{\mathbf{v}}$$

The momentum is the direction of the velocity:

$$\mathbf{p} = \frac{\mathbf{v}}{|\mathbf{v}|} = \hat{\mathbf{v}} \Longrightarrow \|\mathbf{p}\| = 1$$

By Euler-Lagrange $\dot{\mathbf{p}} = 0$ the direction of the path is conserved: This is what we mean by the path being a straight line. Note that Euler-Lagrange equations do not determine the magnitude of the velocity, only its direction. This is a somewhat pathological aspect of geodesics. It is a consequence of the parametrization invariance of time

$$d\ell = |\mathbf{v}|dt = d|\mathbf{x}|$$

2.14 Interacting systems: Adding actions

The action for many interacting degrees of freedom is constructed by adding the actions. Let us look at an example with two degrees of freedom:

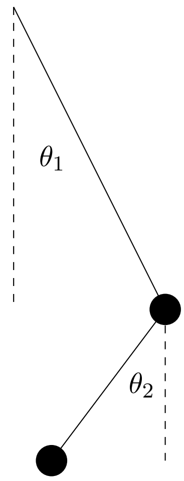
Chaotic systems: Double pendulum

Let us write down the equations of motion for a double pendulum assuming the two masses are identical and so are the lengths. We may choose the energy scale by setting $m\ell^2 = 1$.

$$L = L_1 + L_2, \quad L_1 = \frac{1}{2}\dot{\theta}_1^2 + \omega^2 \cos\theta_1$$

L_2 is more complicate. The potential energy is not so bad

$$U_2 = -mg\ell(\cos\theta_2 + \cos\theta_1) = -\underbrace{(m\ell^2)}_{=1}\omega^2(\cos\theta_2 + \cos\theta_1)$$



We proceed slowly with the kinetic energy

$$x_2 = \ell(\sin\theta_1 + \sin\theta_2), \quad y_2 = -\ell(\cos\theta_1 + \cos\theta_2)$$

It follows that

$$\dot{x}_2 = \ell(\dot\theta_1 \cos\theta_1 + \dot\theta_2 \cos\theta_2), \quad \dot{y}_2 = \ell(\dot\theta_1 \sin\theta_1 + \dot\theta_2 \sin\theta_2)$$

Therefore

$$T_2 = \frac{m}{2}(\dot{x}_2^2 + \dot{y}_2^2) = \frac{m\ell^2}{2}(\dot\theta_1^2 + 2\underbrace{\cos(\theta_1-\theta_2)}_{\cos\theta_1\cos\theta_2+\sin\theta_1\sin\theta_2}\dot\theta_1\dot\theta_2 + \dot\theta_2^2)$$

Hence

$$\boxed{L = \frac{1}{2}\Big(2\dot\theta_1^2 + 2\cos(\theta_1-\theta_2)\dot\theta_1\dot\theta_2 + \dot\theta_2^2\Big) + \omega^2(2\cos\theta_1 + \cos\theta_2)}$$

The generalized momenta are

$$p_1 = 2\dot\theta_1 + \cos(\theta_1-\theta_2)\dot\theta_2, \quad p_2 = \dot\theta_2 + \cos(\theta_1-\theta_2)\dot\theta_1$$

There are two Euler-Lagrange equations, one for each degree of freedom. For example

$$\dot{p}_1 = \frac{\partial L}{\partial\theta_1} = -\sin(\theta_1-\theta_2)\dot\theta_1\dot\theta_2 - 2\omega^2\sin\theta_1$$

where $\dot{p}_1$ is rather complicated, it is not just $\ddot\theta_1$, there is more to it

$$\dot{p}_1 = 2\ddot\theta_1 + \cos(\theta_1-\theta_2)\ddot\theta_2 - \sin(\theta_1-\theta_2)\dot\theta_2(\dot\theta_1 - \dot\theta_2)$$

And a similar equation for p_2.

We make the following observations

- There is just one frequency scale, ω, but this does not mean that the motion is simple.

- The double pendulum is an example of a chaotic system. See Youtube demo of double pendulum. The solution of the equations of motion can not be written down in terms of standard known functions, e.g. ones known to Mathematica.
- The kinetic energy is is quadratic in the velocities. but also depends on the coordinates, $T(\dot{\theta}_1, \dot{\theta}_2, \theta_1 - \theta_2)$.
- The momenta are not easy to figure out from what you were told in in high school. In particular, p_1 is not proportional to $\dot{\theta}_1$ as you may naively expect.

2.15 Driven system: Kapitza pendulum

Lagrangians mechanics allows to write the equations of motion of pretty complicated problems which would be forbiddingly difficult if you only knew Newton's laws.

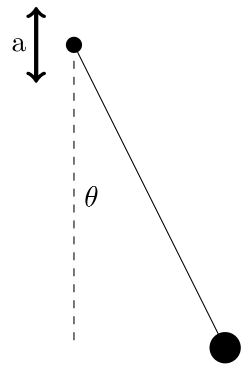


Figure 10: The pivot of the Kapitza pendulum oscillates vertically with amplitude a

Problem 2.3 *Write the Lagrangian for a planar pendulum whose pivot is oscillating at frequency ω and amplitude a. The pendulum has length ℓ and mass m.*

The x and y coordinates are

$$x = \ell \sin\theta, \quad y = -a\cos\omega t - \ell\cos\theta$$

The kinetic energy is

$$T = \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2\right) = \frac{m}{2}\left(\ell^2\dot{\theta}^2 + 2a\ell\omega\dot{\theta}\sin\omega t\sin\theta + (a\omega)^2\sin^2\omega t\right)$$

Integrating by parts, this can be written as

$$T = \frac{m}{2}\left(\ell^2\dot{\theta}^2 + 2a\ell\omega^2\cos\omega t\cos\theta\right) + \underbrace{\frac{m}{2}\left(-2a\ell\omega\frac{d(\sin\omega t\cos\theta)}{dt} + (a\omega)^2\sin^2\omega t\right)}_{\frac{dF}{dt}}$$

Terms of form $\frac{dF}{dt}(\theta, t)$ can be dropped from the Lagrangean. The potential energy is

$$U = mgy = -mg\left(\ell\cos\theta + a\cos\omega t\right)$$

Again, the term on the right is of the form $\frac{dF}{dt}(\theta, t)$. The Lagrangian is $L = T - U$. dropping terms of the form $\frac{dF}{dt}(\theta, t)$ we finally get

$$L = \frac{m}{2}\left(\ell^2\dot\theta^2 + 2a\ell\omega^2\cos\omega t\cos\theta\right) + mg\ell\cos\theta$$

This equation has many parameters, (m, ℓ, ω, g) that make it difficult to see what it says. To simplify matters, let me scale time so we measure time in multiples of the driving period $2\pi/\omega$. Namely, let me denote $s = \omega t$. In this dimensionless time s the Lagrangian takes the form

$$\boxed{\left(\frac{1}{m\omega^2\ell^2}\right) L(\dot\theta, \theta, s) = \frac{1}{2}\left(\frac{d\theta}{ds}\right)^2 + \frac{a}{\ell}\cos s\cos\theta + \frac{\omega_g^2}{\omega^2}\cos\theta, \quad \omega_g^2 = \frac{g}{\ell}} \tag{2.16}$$

The overall factor in front of of the Lagrangian doe not affect the Euler Lagrange equation. The E-L equation then depend on two dimensionless parameters: a/ℓ, the ratio between the driving amplitude and the length of the pendulum and $(\omega_g/\omega)^2$, the ratio of the natural frequency to the driving frequency. In actual Kapitza pendulums these to parameters are small numbers. You can say quite a lot about the solution of the equation, without solving it explicitly. If

$$\frac{a}{\ell} \ll \left(\frac{\omega_g}{\omega}\right)^2$$

we can forget about the driving, and the behavior is close to an ordinary pendulum. In the opposite limit

$$\frac{a}{\ell} \gg \left(\frac{\omega_g}{\omega}\right)^2$$

we can forget about gravity.

Problem 2.4 *Derive the Euler-Lagrange equations for the Lagrangian in 2.16*

A Youtube demo can be found here.

2.16 On the existence and uniqueness of the minimizer

For a free particle the solution of Euler Lagrange equations give a minimizer of the action. This is often, but not always the case.

Consider the Lagrangian of a Harmonic oscillator

$$L = \frac{m}{2}v^2 - \frac{k}{2}x^2 = \frac{m}{2}\left(v^2 - \omega^2 x^2\right), \quad \omega^2 = \frac{k}{m}$$

Euler-Lagrange equations are:

$$\ddot x = -\omega x \tag{2.17}$$

Theorem 2.6 *The general solution of the ODE*

$$\ddot{x} = -\omega x$$

is

$$x(t) = Ae^{i\omega t} + Be^{-i\omega t}$$

where $A, B \in \mathbb{C}$ are free (complex) parameters. In particular, the most general real solution (this is what is relevant to mechanics) is

$$x(t) = A\sin\omega t + B\cos\omega t$$

where $A, B \in \mathbb{R}$ are free (real) parameters.

It follows from this that:

Fact 2.1 *The action for a classical path that starts at $x = 0$ at time $t = 0$ and satisfies Euler-Lagrange equations is*

$$S = \frac{m}{2}\int_0^T (v^2 - \omega^2 x^2)dt = \frac{A^2\omega^2 m}{2}\int_0^T (\cos^2\omega t - \sin^2\omega t)dt = \frac{A^2\omega m}{4}\sin 2\omega T$$

It follows that the minimizer of the action need not be unique:

Fact 2.2 *If T is a period, or half a period, then the action $S = 0$ independent of A. The Action clearly does not have a unique stationary point.*

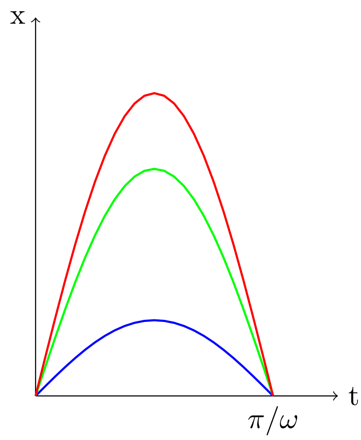


Figure 11: All these paths are solutions of Newton's equation, and start and end at the same points. All have action 0.

The fact that the solutions of Newton's equations are periodic leads to an amusing puzzle: What is the optimal path that connect $(x = 0, t = 0)$ with $(x_f, t = \pi/\omega)$ with $x_f \neq 0$?

To see what happens let us consider the family that connects the initial and final points and has a free parameter A:

$$x(t) = \frac{x_f}{\pi}\omega t + A \sin \omega t \tag{2.18}$$

The action can be computed explicitly:

$$S = -m x_f^2 \omega \left(\frac{A}{x_f} + \frac{\pi^2 - 3}{6\pi} \right)$$

If $x_f = 0$ the action vanishes independent of A, as we have already seen. But if $x_f \neq 0$ then you can make S as large (and negative) as you want by letting $A/x_f \to \infty$. The action has *no* minimizer since it is unbounded below. The "optimal path" connecting the two points does this by first going to infinity.

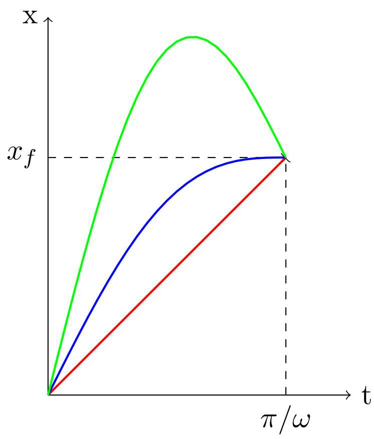


Figure 12: The figure shows three paths for with A of Eq. (2.18) taking values $A = 0, 1, 3$. The larger A the smaller the action. The minimizer is a path that connects the point 0 to the point 1 by first going through infinity.

3 Constraints

3.1 Holonomic and an-holonomic constraints

So far we considered generalized coordinates q_j which could be varied independently. Now let us consider situations where this is not the case. There are basically 3 scenarios

- The coordinates q_j are related by one (or more) function

  $$g(q_1, \dots, q_n) = 0$$

  Such a constraint is called Holonomic.

- The coordinates q_j are constrained by an inequality

$$g(q_1, \dots, q_n) \geq 0$$

(or several inequalities).

- The coordinates q_j and velocities $\dot{q}_j$ are constrained by one or more function

$$g(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n) = 0$$

which does not reduce to a function of the coordinates only. Such constraints are called an-holonomic.

Example 3.1 *Consider a free particle of mass m constrained to move on a sphere of radius R. In spherical coordinates the constraint is*

$$g(r) = r - R = 0$$

and the free Lagrangian is

$$L = \frac{m}{2}\left(\dot{r}^2 + r^2(\dot{\theta}^2 + \dot{\varphi}^2 \sin^2\theta)\right) \tag{3.1}$$

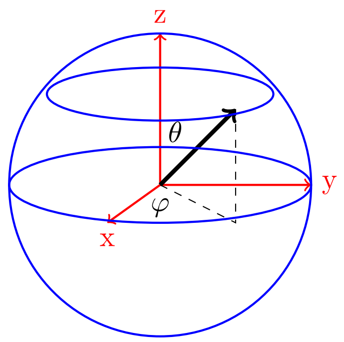


Figure 13: Spherical coordinates

It is tempting, and even correct, to substitute the constraint in the Lagrangian to get a new Lagrangian for the remaining generalized coordinates (θ, ϕ). Since the constraint forces the kinetic energy in the radial direction to vanish we have

$$L = \frac{mR^2}{2}\left(\dot{\theta}^2 + \dot{\varphi}^2 \sin^2\theta\right)$$

But, why is this correct? At the end of the day a Lagrangian is correct if it reproduces Newton's equation. The constraint represents the action of forces. Where are they?

Example 3.2 *A single ball on a billiard table is can be modeled by a free Lagrangian with an inequality constraint, e.g. $g(x, y) = y \geq 0$. This is a case where you know that if you pick two points on the table, there are usually several classical paths the connect them.*

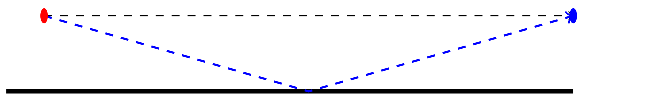


Figure 14: We need to consider both paths that avoid the boundary and paths that do not. The dashed black line is the absolute minimizer. The dashed blue line is a path that hits the boundary. For variations of the hitting point along the boundary the local minimizer has the angle of incidence equal the angle of reflection. It is clearly a good solution of Newton's equations.

Critical points with constraints–Lagrange multipliers

Suppose we want to find a stationary point of the function $f(x)$ subject to the constraint $g(x) \geq 0$.

Example 3.3 (Constrained critical points) *Find the critical points of* $f(x, y, z) = z$ *subject to the constraint* $g(x, y, z) = x^2 + y^2 + z^2 - 1 = 0$.

- *First method: A simple way to do that: Switch coordinates to spherical* (r, θ, ϕ). *Then the problem reduces to maximizing* $f = r \cos\theta$ *subject to the constraint* $g(r) = r - 1 = 0$. *The critical points are* $\partial_\theta \cos\theta = \sin\theta = 0$ *i.e.* $\theta = 0, \pi$ *or* $z = \pm 1$.

- *Second method: Suppose you did not hit the idea of using spherical coordinates. You could then try to solve the constraint*

$$z = \pm\sqrt{1 - x^2 - y^2}$$

The critical points lie at $x = y = 0$ *and hence* $z = \pm 1$

If you can find a point (x_0, y_0, z_0) which solves $g(x, y, z) = 0$, and if $\partial_z g \neq 0$, then, in principle, you can solve for $z = g^{-1}(x, y)$ in the neighborhood of this point. You can then substitute z in $f = f(x, y, g^{-1}(x, y))$ and get an unconstrained optimization problem. However, this will, in general not work globally. In the example 3.3 there are actually two functions: One for $z > 0$ and another for $z < 0$.

Suppose we want to find a stationary point of the function $f(x)$ in the region defined by $g(x) \geq 0$. Here g is the constraint. The variation of f is

$$\delta f = \sum (\partial_j f) \delta x_j$$

In the interior points, where $g(x) > 0$, you can vary δx_j independently. In this case the condition that f is stationary is

$$\partial_j f = 0, \quad \forall j$$

However, on the boundary you can not move normal to the boundary.

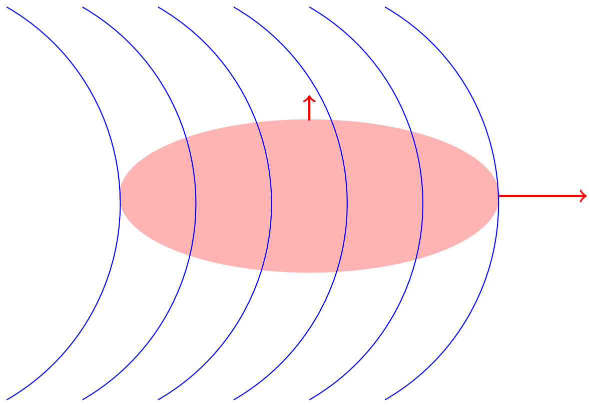


Figure 15: Lagrange methods: The level sets of f are the blue curves and the region $g \geq 0$ is the red ellipse. A stationary points of f that satisfies the constraint is characterized by the fact that ∇f is parallel to ∇g. If the ∇f and ∇g are not parallel, you can go either up or down in f staying on $g = 0$.

This means that you want to satisfy two variational principles: You want the variation of f to vanish because it is optimized and you want the variation of g to vanish because you move on a surface of fixed level:

$$\underbrace{\sum_i \delta x_i \partial_i f = 0,}_{optimizer} \quad \underbrace{\sum_i \delta x_i \, \partial_i g = 0}_{constraint}$$

Lagrange realized that the solution has a geometric meaning that leads to a nice equation: Since the gradient of f is perpendicular to the level surface it must be that ∇f and ∇g must be parallel

$$\nabla f \| \nabla g \Longrightarrow \nabla f + \lambda \nabla g = 0 \tag{3.2}$$

with λ a scalar[3]. This is illustrated in the picture.

Eq. 3.2 can be written as a variational problem for $f + \lambda g$ where we treat λ as an additional coordinate: $(x_1, \ldots x_n, \lambda)$:

$$\sum_i \delta x_i \underbrace{(\partial_i f + \lambda \partial_i g)}_{0 \ by \ 3.2} + (\delta \lambda) \underbrace{g}_{0 \ by \ constraint} = 0$$

δx_i and $\delta \lambda$ are treated as independent and one gets the n equations 3.2 plus the constraint equation for the $n + 1$ unknowns.

In summary:

Theorem 3.1 *The stationary points of the function $f(q_1, \ldots, q_n)$ restricted to the surfaces $g_j(q_1, \ldots, q_n) = 0$, with $j = 1, \ldots, r < n$ are given by the solutions of the $n + r$*

[3] A number in this case

equations for the $n + r$ *unknowns* $q_1, \ldots, q_n$ *and* $\lambda_1, \ldots, \lambda_r$

$$\partial_j f + \sum \lambda_k \partial_j g_k = 0, \quad j = 1, \ldots, n$$

and

$$g_k = 0, \quad k = 1, \ldots, r$$

assuming $\partial_j g_k \neq 0$.

Example 3.4 *We can now solve example 3.3 in a third way:*

$$\nabla f = (0, 0, 1), \quad \nabla g = 2(x, y, z)$$

The gradients are parallel at $x = y = 0$ *and* $z = \pm 1$.

3.2 Holonomic constraints

Consider a motion constrained by a single constraint on the generalized coordinates and time

$$G(q_1, \ldots, q_n, t) = 0 \tag{3.3}$$

with the Lagrangian $L(\dot{q}_1, q_1, \ldots, \dot{q}_n, q_n)$. The variation of the Lagrangian gives

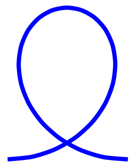


Figure 16: A curve $G(q_1, q_2) = 0$ that constrain the two coordinates in the plane to a motion in one dimension.

$$\delta L = \sum A_j \delta q_j = \mathbf{A} \cdot \delta \mathbf{q}, \quad A_j = \dot{p}_j - \partial_j L \tag{3.4}$$

Now, however, the $\delta L = 0$ does not imply $\mathbf{A} = 0$ because δq_j can not be varied independently. The constraint forces them to satisfy

$$\sum_j (\partial_j G) \delta q_j = \nabla G \cdot \delta \mathbf{q} = 0 \tag{3.5}$$

This means that we need to solve simultaneously

$$\mathbf{A} \cdot \delta \mathbf{q} = 0, \quad \nabla G \cdot \delta \mathbf{q} = 0$$

This is precisely the problem solved by Lagrange. We must have at every point along the orbit,

$$\mathbf{A} = \lambda(t) \nabla G$$

In summary:

Theorem 3.2 *Euler-Lagrange equations with a holonomic, possibly time dependent, constraint $G(q_1, \ldots, q_n, t) = 0$ are*

$$\frac{dp_j}{dt} - \frac{\partial L}{\partial q_j} = \lambda(t) \frac{\partial G}{\partial q_j}, \tag{3.6}$$

The rhs is naturally interpreted as the force of the constraint.

The n differential equations in the theorem are the Euler Lagrange equations for the augmented Lagrangian

$$L(\dot{q}_1, \ldots, q_n, t) - \lambda G(q_1, \ldots, q_n, t)$$

where we append to the generalized coordinate q_j the Lagrange multiplier as an additional coordinate. Since the augmented Lagrangian does not depend on $\dot{\lambda}$ the associated momentum vanishes $p_\lambda = 0$ and the corresponding Euler-Lagrange equation is the constraint equation $G(q_j) = 0$.

Example 3.5 (Free particle on a sphere) *Consider again the free particle on a sphere of radius R.*

$$L = \frac{m}{2}\mathbf{v}^2 + \frac{\lambda}{2}(\mathbf{x}^2 - R^2)$$

Euler-Lagrange gives

$$\mathbf{p} = m\mathbf{v}, \quad \dot{\mathbf{p}} = \lambda \, \mathbf{x}$$

The force of the constraint is therefore in the radial direction. This is nice, because it implies that the force of the constraint does no work.

- *Constrained force and power: From $\mathbf{x}^2 = R^2$ we get*

$$\mathbf{x} \cdot \mathbf{v} = 0 \tag{3.7}$$

  It follows that the constraint does not do work–the power of the constraint force vanishes:

$$\dot{\mathbf{p}} \cdot \mathbf{v} = \lambda \mathbf{x} \cdot \mathbf{v} = 0$$

- *Magnitude of the constraint force: To find the magnitude of the constraint force we need to find λ. This follows from*

$$\dot{\mathbf{p}} = \lambda \, \mathbf{x} \Longrightarrow \mathbf{x} \cdot \dot{\mathbf{p}} = \lambda \mathbf{x}^2 = \lambda R^2$$

  Now note that by differentiating Eq. 3.7

$$\mathbf{v}^2 + \mathbf{x} \cdot \dot{\mathbf{v}} = 0 \tag{3.8}$$

  We substitute this in

$$\lambda = \frac{\mathbf{x} \cdot \dot{\mathbf{p}}}{R^2} = -\frac{m\mathbf{v}^2}{R^2}$$

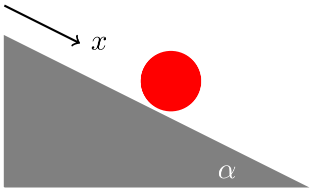


Figure 17: Cylinder rolling without slipping on an inclined plane.

We conclude that the constraint force is simply the centripetal force that keeps the bead on the sphere:

$$\lambda \mathbf{x} = \underbrace{-\frac{m\mathbf{v}^2}{R}\hat{\mathbf{x}}}_{centripetal\ force}$$

Example 3.6 (Holonomic velocity constraints) *Consider a cylinder, of mass M, radius R and moment of inertia I, rolling without slipping, along an inclined plane. Pick the two generalized coordinates as (θ, x). The Lagrangian in a gravitational field is*

$$L = \frac{m}{2}\dot{x}^2 + \frac{I}{2}\dot{\theta}^2 + mgx\sin\alpha$$

The constraint is the relation between the velocities

$$\dot{x} = R\dot{\theta}$$

This first looks like an an-holonomic constraint since it is not a relation between the coordinates. However, the constraint can be integrated to such a relation, namely

$$x = R\theta$$

The constraint is therefore holonomic.

Let us now look at the constraint forces using the method of Lagrange

$$L \mapsto L + \lambda(x - R\theta)$$

There are now two constraint forces:

$$F_x = \lambda, \quad F_\theta = -\lambda R$$

By dimension analysis F_x has dimension of force while F_θ has dimension of torque. Let us compute the total power associated with the constraint

$$\dot{x}F_x + \dot{\theta}F_\theta = \dot{x}\lambda - \dot{\theta}R\lambda = 0$$

The forces of friction transfer energy from the linear velocity to the angular velocity, but do not dissipate or generate energy.

Example 3.7 (Great circles are geodesics) *The Lagrangian describing geodesics on a sphere of radius R is*

$$L = |\dot{\mathbf{x}}| - \frac{\lambda}{2}(\mathbf{x}^2 - R^2)$$

The canonical momentum is the direction of the velocity

$$\mathbf{p} = \frac{\dot{\mathbf{x}}}{|\dot{\mathbf{x}}|}$$

and the equation of motion is

$$\dot{\mathbf{p}} = -\lambda \mathbf{x}, \quad \dot{\mathbf{p}} = \frac{\ddot{\mathbf{x}}|\dot{\mathbf{x}}|^2 - \dot{\mathbf{x}}(\dot{\mathbf{x}} \cdot \ddot{\mathbf{x}})}{|\dot{\mathbf{x}}|^3}$$

Great circles on the sphere are the intersection of a plane with the sphere

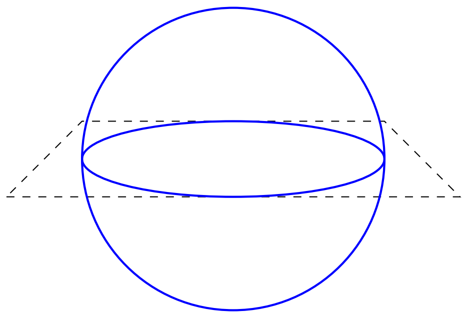


Figure 18: The intersection of the sphere with a plane through the origin is a great circle

$$\mathbf{x} \cdot \hat{\mathbf{a}} = 0, \quad \mathbf{x}^2 = R^2 \tag{3.9}$$

where $\hat{\mathbf{a}}$ is a constant (in time) unit vector in the direction perpendicular or the plane. We want to show that this is a property of the solutions of the equations of motion. To do this we need to guess $\hat{\mathbf{a}}$. Here is a natural guess: Clearly $\dot{\mathbf{x}}$ must be a vector in this plane. But if $\dot{\mathbf{x}}$ is to stay in the plane then $\ddot{\mathbf{x}}$ must be a second vector in the plane. But, for a particle moving on a sphere $\ddot{\mathbf{x}}$ is parallel to $\mathbf{x}$. Since $\mathbf{p}$ is parallel to $\dot{\mathbf{x}}$ we get for the vector perpendicular to the plane

$$\mathbf{a} = \mathbf{x} \times \mathbf{p} \tag{3.10}$$

which clearly satisfies Eq. 3.9. What we need to check is that $\mathbf{a}$ is a constant vector, that does not change in time: A constant of motion.

Indeed, since $\mathbf{p}$ is parallel to $\dot{\mathbf{x}}$

$$\dot{\mathbf{a}} = \dot{\mathbf{x}} \times \mathbf{p} + \mathbf{x} \times \dot{\mathbf{p}} = 0 - \lambda \mathbf{x} \times \mathbf{x} = 0$$

and the proof is complete.

Exercise 3.1 1. *Find the magnitude of the acceleration $a(\theta)$ for a roller coaster that starts at rest at the top on circular tracks in a gravitational field.*

2. *Show that a is independent of the radius of the hoop R but depends on the angle θ*

3. *Plot the function $a(\theta)$. If you are too busy to do this, ask Wolfram alpha.*

4. *The function has a minimum where $a < g$. This is called by pilots negative g. This is the sick feeling you get when you fall.*

5. *Why isn't the minimum at $\theta = \pi/2$?*

6. *Find the sector $[\theta_0, \pi]$ where $a \geq 3g$*

7. *Show that it takes about*
$$2\sqrt{R/g}$$
seconds to traverse the region $a \geq 3g$.

8. *What can you conclude about the size of loops in amusement parks assuming that some people will pass out if they feel $a \geq 3g$ for more than one second.*

Example 3.8 (Motion on a sphere with Magnetic monopole)

3.3 Non-holonomic constraints

Definition 3.1 *A non-holonomic constraint is a constraint on the generalized coordinates and velocities, $F(\dot{q}_j, q_t, t) = 0$, which is not equivalent to a constraint on the generalized coordinates alone $G(q_t, t) = 0$*

Non-holonomic constraints are a mess. This is one place where you will find one textbook claiming that another famous textbook has mistakes. Some papers in Journals devoted to physics education that attempt to address old mistakes make new mistakes. This does not mean nobody understands an-holonomic constraints. It only means that the subject is tricky and subtle. Since the literature is full with wrong formulas. Beware (this includes my notes). The safer places to learn this tend to be in math literature and engineering devoted to control theory.

In order not to make a fool of myself, I will only consider one special class of anholonomic constraints, namely, constraint that are linear and homogeneous in the velocities. This appears to be the simplest case.

$$\sum \dot{q}_j f_j(q_1, \ldots, q_n, t) = 0 \tag{3.11}$$

Because the constraint is linear and homogeneous in the velocities we can write it as a relation between differentials

$$\sum \delta q_j f_j(q_1, \ldots, q_n, t) = 0$$

This is the setting in Eq. (3.5) with $\partial_j G \mapsto f_j$. We therefor have the analog of Theorem 3.2

Theorem 3.3 *Euler-Lagrange equations with an anholonomic constraint linear in the velocities, $\sum \dot{q}_j f_j(q_1, \dots, q_n, t) = 0$ are*

$$\frac{dp_j}{dt} - \frac{\partial L}{\partial q_j} = \lambda(t) f_j, \tag{3.12}$$

The rhs is naturally interpreted as the force of the constraint.

Note that the constraint equation 3.11 can be interpreted as the statement that the constraint does no work.

Example 3.9 *Consider a wheel rolling without slipping on the plane. There are 4 generalized coordinates $(x, y, \theta, \phi, \psi)$. The Lagrangian is*

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) + \frac{I_3}{2}\dot{\theta}^2 + \frac{I_1}{2}(\dot{\psi}^2 + \dot{\phi}^2)$$

The constraints of rolling without slipping give

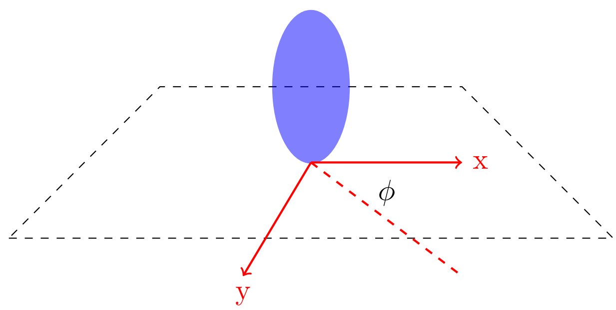


Figure 19: θ is the rotation angle of the disc about its axis. ϕ is the rolling direction and ψ the inclination angle of the disc with the table.

$$R(\cos\phi\, d\theta + \sin\phi\, d\psi) = dx, \quad R(\sin\phi\, d\theta - \cos\phi\, d\psi) = dy$$

You may think of these as (linear) relations between the velocities.

The relations do not translate to two relations among the coordinates alone: If I give you the three angles θ, ϕ ,and ψ for the front wheel of your bike you still do not know where your bike is.

Example 3.10 (Euler's disc) *This youtube shows an amusing illusion of acceleration when actually the motion slows down. What happens when a coin falls with friction. Euler's disc.*

Example 3.11 *A nice and surprising movie showing what an-holonoomic constraints can do to a ball rolling inside a vertical cylinder is here.*

Example 3.12 (Demo of nontrivial motions) *Museum demos.*

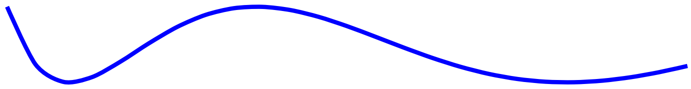


Figure 20: Slalom skier controls the direction but not the speed.

Example 3.13 (Skiing) *When you ski you control the direction you go, but not the speed. If the slope has angle θ and x is the coordinate along the slope the Lagrangian is:*

Let us see how to handle the variational principle in the case that you skate on an inclined plane with coordinates $\mathbf{x} = x\mathbf{e}_x + y\mathbf{e}_y$. The Lagrangian is

$$L/m = \frac{\dot{x}^2 + \dot{y}^2}{2} - \gamma x, \quad \gamma = g \sin\theta \tag{3.13}$$

Write

$$\mathbf{v} = \hat{\mathbf{v}}|\mathbf{v}|, \quad |\mathbf{v}| = \frac{d\ell}{dt}$$

$\dot{\ell}$ is free to vary but the direction is not. Hence, the allowed variation is $d\ell$ and the direction of the velocity $\hat{\mathbf{v}}$ is contained

$$\delta L/m = |\mathbf{v}|\delta|\mathbf{v}| - \gamma\delta x = \frac{d(|\mathbf{v}|\delta\ell)}{dt} - \frac{d|\mathbf{v}|}{dt}\delta\ell - \gamma\delta x$$

$\delta\ell$ and δx are not independent:

$$\hat{\mathbf{v}} \cdot \hat{\mathbf{e}}_x \, \delta\ell = \delta x$$

The variation over $\delta\ell$ gives one equation

$$\frac{d|\mathbf{v}|}{dt} + \gamma\hat{\mathbf{v}} \cdot \hat{\mathbf{e}}_x = 0 \tag{3.14}$$

This says that the speed increases when you point the skis downhill. The speed is then determined by Euler-Lagrange and its direction by the control. Since you control $\hat{\mathbf{v}}$ this is a known function of time. We reduced the problem to solving an ODE with a known function of time (a.k.a. source term), the term on the rrright.

3.4 Global constraints

This subsection is an aside. It is not about constraints that naturally arise in mechanics, but about constraints that arise in the study of variational problems that teaching assistants like to give.

Consider constraint that has the same structure as the action S, i.e.

$$\int dt \, G(q_j, \dot{q}_j, t) = 0$$

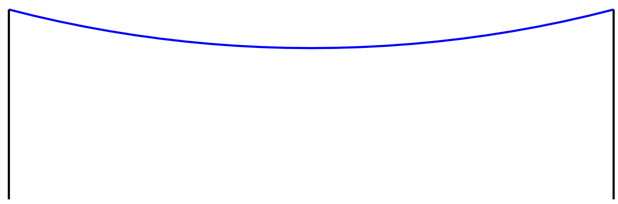


Figure 21: The shape of an electric wire of length ℓ between two poles is determined by solving variational problem with a constraint.

Example 3.14 *The wire is described by function $y(x)$ with $x \in [a, b,]$. a and b are the locations of the poles. The length of the wire is the constraint*

$$\ell = \int_a^b \sqrt{1 + \dot{y}^2} dx$$

The wire wants to minimize its potential energy. So, the analog of the action is

$$S = \int_a^b y\sqrt{1 + \dot{y}^2}\, dx$$

Varying a path gives a variation in G

$$\delta G = \underbrace{\sum_j \frac{d}{dt}\left(\frac{\partial G}{\partial \dot{q}_j}\delta q_j\right)}_{\textit{boundary term}} + \sum_j \left(-\frac{d}{dt}\frac{\partial G}{\partial \dot{q}_j} + \frac{\partial G}{\partial q_j}\right)\delta q_j$$

The boundary term does not matter since $\delta q_j = 0$ at the end points. To translate this problem to the problem of Lagrange, where we optimize a function subject to constraints, you have to think of this as a function of infinitely many variables:

$$\{q_1(t_1), q_2(t_1), \ldots, q_1(t_2), q_2(t_2), \ldots q_n(t_n)\}$$

with one Lagrange multiplier. The dictionary is

$$\begin{aligned} f &\Longleftrightarrow S \\ \sum_j &\Longleftrightarrow \int dt \\ \partial_j f &\Longleftrightarrow \left(-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} + \frac{\partial L}{\partial q_j}\right) \\ \delta x_j &\Longleftrightarrow \delta q_j(t_k) \end{aligned}$$

We appeal again to Lagrange and get the Euler Lagrange equations

$$\sum_j \left(-\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} + \frac{\partial L}{\partial q_j}\right) + \lambda\left(-\frac{d}{dt}\frac{\partial G}{\partial \dot{q}_j} + \frac{\partial G}{\partial q_j}\right) = 0$$

In this case λ is a constant, independent of time, reflecting the fact the the constraint is global.

4 Conservation laws: Nöther theorem

Emmy Nöther was one of the great mathematicians of the early 20 century. She was German, Jewish and a woman. Bad luck. When D. Hilbert brought up her appointment as a professor of Math in Göttingen, some of his colleagues at the senate said it was unthinkable that veterans of the (first) world war would have to sit at the feet of a woman. Hilbert responded: Meine Herren, I do not see that the sex of the candidate is an argument against her admission as a Privatdozent. After all, the Senate is not a bathhouse.

Emmy Nöther was removed from office by the Nazis since "she did not demonstrate her loyalty to the fatherland". You can read more about her in Wikipedia.

Nöther theorem reinterprets conservation laws, such as conservation of momentum, angular momentum and energy, as consequences of symmetries. This insight had profound effect on the development of modern physics.

4.1 Physical spaces

The only space known to the Greek was Euclidean space, and of course, everybody thought, or believed that the physical space we live in is Euclidean. The first person to realize that it is possible to contemplate other types of spaces was Gauss. Once he realized the other spaces were, in principle, possible, he also wondered if our physical world is indeed Euclidean. To answer this question he made an experiment.

In a Euclidean space the angle of triangles add up to π. So Gauss, who was also the father of surveying, measured the angles of a triangle, whose vertexes were three mountaintops, assuming that light moves on geodesics. The deviation from π, if any, was much smaller than the experimental errors he found.

In either case, Gauss, and probably everyone else, thought that the space we live in, be it Euclidean or not, is God given. A constant of nature, like the velocity of light. This changed with General Relativity. In general relativity the geometry of space time is dynamics and evolves, just as electromagnetic fields evolve.

However, in practice, unless you live too close to a black hole, taking space to be Euclidean is an excellent approximation.

It might be instructive to have in mind an example of a space that is not Euclidean: A sphere. A case you can visualize in a 2-dimensional sphere embedded in 3 Dimensional Euclidean space

$$x^2 + y^2 + z^2 = R^2$$

Similarly, you may consider a 3-dimensional sphere embedded in a 4-dimensional space

$$w^2 + x^2 + y^2 + z^2 = R^2$$

If R is humongous, and we happened to live near $w = R$, and never get far away from home, then our everyday experience is consistent with thinking that (x, y, z) are Euclidean coordinates. To see this is more detail, consider a neighboring points on the

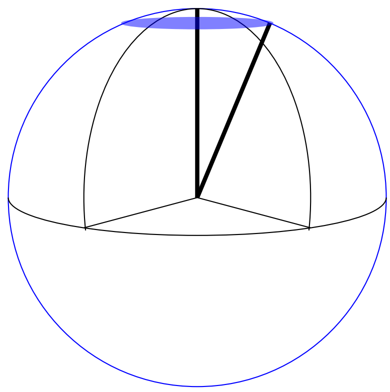


Figure 22: If you lived near the north-pole of a sphere or large radius R, and you never traveled far you might be tempted to think that you lived on a Euclidean disc. This is what many people in the ancient world thought. You could still find out that you live on a sphere if you could measure angles accurately: The angles in a spherical triangle sum to more than π. Eratosthenes accurately computed the radius of earth in the 2-nd century b.c.

sphere. The Euclidean distance between two neighboring points is

$$\begin{aligned}(d\ell)^2 &= (dx)^2 + (dy)^2 + (dz)^2 + (dw)^2 \\ &= (dx)^2 + (dy)^2 + (dz)^2 + \left(\frac{xdx + ydy + zdz}{w}\right)^2\end{aligned}$$

where I used the constraint

$$wdw + xdx + ydy + zdz = 0$$

Hence, if $w \approx R$ and x, y, z never get to be comparable to R the distance between two points on the 3-sphere is close to the Euclidean distance.

Homogeneity of the physical space means that you do not need to change your physics books when you immigrate to Berlin. In some sense this is the essence of the Copernican revolution: Contrary to what most religious leaders tell you, neither Mecca, nor Jerusalem are distinguished as far as the laws of physics are concerned.

4.2 Euclidean spaces

Euclidean space is the first space we learn about in school and for most people, this is also the last one. Euclid (implicitly) defined the space through 5 axioms, stated loosely:

- Every two points define a straight line
- Lines can be extended to infinity

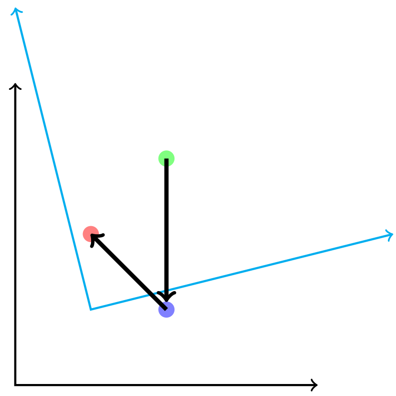


Figure 23: Euclidean space has no distinguished origin and no distinguished orientation. Here we have chosen arbitrarily two frames. Any two points in the space define a vector. The scalar product between vector allows to associate lengths with vectors and angles between them. These properties do not care what frame you picked.

- All right angles are equal
- One can draw circles with arbitrary radius and center
- There is a unique parallel the rough a point not on the line

Below I shall give a more modern, Cartesian, definition of Euclidean space. But let me first say something about its symmetries.

A Euclidean spaces is homogeneous and isotropic:

Definition 4.1 (Homogeneous) *A space is called homogeneous if it admits an action, called translation, that allows to take any point to any other point, so that the distances between any two points remains the same.*

This means that if you have a rigid body, you can still place it anywhere you want. You may even throw it. If space was not homogeneous, translating an object will result in buildup of stresses in the new location. Here is a material analogy: A sea of pure water looks homogeneous–you are surrounded by the same molecule everywhere–but diving results in deformation due to hydrostatic pressure.

Definition 4.2 (Isotropy) *A space is called isotropic if it admits an action, called rotation, which moves all points not on the axis, so that the distances between any two points remains the same.*

This says that the world has no distinguished direction. Isotropy allows you to rotate and spin rigid bodies.

Definition 4.3 *A Euclidean space is a space that admits a Cartesian coordinate system. This means:*

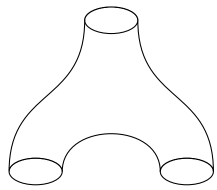


Figure 24: The pants in the picture is an example of a two dimensional space that is neither homogeneous nor isotropic. If you cut a piece somewhere it will not fit in most places.

- *Any point in the space can be described (uniquely) in terms of (Cartesian) coordinates as a (column) vector*
$$\mathbf{x} = (x_1, \ldots, x_n)^t$$

- *Translation by* $\mathbf{y}$ *acts on* $\mathbf{x}$ *by*
$$\mathbf{T_y x} = (x_1 + y_1, \ldots, x_n + y_n)^t$$

- *Rotation about the origin is given by an* $n \times n$ *real orthogonal matrix* R *acting on* $\mathbf{x}$ *by*
$$\mathbf{x} \mapsto R\mathbf{x}, \quad \underbrace{R^t = R^{-1}}_{orthogonal}$$

- *In two dimensions rotations move all points except the origin. In 3 dimensions rotations move all points not on the axis which is given by the eigenvector of* R *with eigenvalue* 1.

- *The scalar product between two vectors is defined by*
$$\mathbf{x} \cdot \mathbf{y} = \sum x_1 y_1, + \ldots, x_n y_n$$

- *The distance between two point* $\mathbf{x}$ *and* $\mathbf{y}$ *is determined by the scalar product*
$$dist(\mathbf{x}, \mathbf{y}) = \sqrt{(\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y})}$$
(and satisfies the requirements from a distance, e.g. the triangle inequality).

- *The distance is translation invariant*
$$dist(\mathbf{x}, \mathbf{y}) = dist(\mathbf{x} - \mathbf{a}, \mathbf{y} - \mathbf{a})$$

- *The angles between two vectors is defined*
$$\cos \alpha = \hat{\mathbf{x}} \cdot \hat{\mathbf{y}}$$
and is invariant under rotation.

Remark: Some of the assertions in the "definition" are actually "theorems". For example, the last assertion follows from:

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a} \cdot (\mathbb{1}\mathbf{b}) = \mathbf{a} \cdot (R^{-1}R\mathbf{b}) = \mathbf{a} \cdot (R^t R\mathbf{b}) = (R\mathbf{a}) \cdot (R\mathbf{b})$$

Exercise 4.1 *Show that translations and rotations preserve distances and angles.*

Example 4.1 *Rotation about the z-axis by angle θ is given by*

$$R = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & \\ 0 & 0 & 1 \end{pmatrix}$$

Example 4.2 *You can use Einstein summation convention to show that length of a vector $\mathbf{x} \cdot \mathbf{x}$ is invariant under rotations*

$$\begin{aligned} (x')_k(x')_k &= R_{kj}x_jR_{km}x_m = (R_{kj}R_{km})x_jx_m \\ &= (R^t_{jk}R_{km})x_jx_m = (R^tR)_{jm}x_jx_m \\ &= \delta_{jm}x_jx_m = x_jx_j \end{aligned}$$

4.3 Vectors and tensors in Euclidean space

You are told in high school that a vector is an object that has a magnitude and a direction. Here is a formal definition:

Definition 4.4 *A vector $\mathbf{v}$ is an object that transforms as the Euclidean coordinates do under rotation, i.e*

$$\mathbf{v} \mapsto R\mathbf{v}$$

In coordinates:

$$v_j \mapsto \sum_k R_{jk}v_k = \underbrace{R_{jk}v_k}_{\textit{implied summation}} \tag{4.1}$$

Theorem 4.1 *The only vector that is left invariant under arbitrary rotation is the zero vector.*

Consider a real symmetric matrix M. In general, such a matrix will have distinct eigenvalues and then its eigenvectors are mutually orthogonal:

$$M\mathbf{e}_j = \lambda_j\mathbf{e_j}, \quad \mathbf{e}_j \cdot \mathbf{e}_k = \delta_{jk}$$

The matrix M gives an orthogonal frame. This gives much more information than the information in a single vector.

Under a rotation R the frame rotates

$$\mathbf{e}_j \mapsto R\mathbf{e}_j$$

Since

$$RM\mathbf{e}_j = (RMR^t)(R\mathbf{e}_j) = \lambda_j R\mathbf{e_j}$$

The law of transformation for M is fixed by

$$M \mapsto RMR^t \tag{4.2}$$

In components

$$M_{jk} \mapsto R_{jn} M_{nm} R^t_{mk} = R_{jn} R_{km} M_{nm} \tag{4.3}$$

(implied summation).

Exercise 4.2 *Show that Eq. 4.2 takes a symmetric matrix to a symmetric matrix*

This motivates:

Definition 4.5 *A second rank tensor is an object that transforms under rotations according to Eq. 4.3.*

Theorem 4.2 *The only second rank tensor that is left invariant under arbitrary rotation is a multiple of the identity matrix.*

In 3 dimensions the Levi-Civita tensor allows us to translate vectors to anti-symmetric second rank tensors

Definition 4.6 *In 3 dimensions define*

$$(V^*)_{jk} = \varepsilon_{ijk} V_i$$

In matrix form

$$V^* = \begin{pmatrix} 0 & v_3 & -v_2 \\ -v_3 & 0 & v_1 \\ v_2 & -v_1 & 0 \end{pmatrix}$$

Exercise 4.3 *Show that V is in the kernel of the matrix V^*.*

Exercise 4.4 *Show that if V^* transforms like a second rank tensor then V transforms like a vector.*

Theorem 4.3 *The Levi-Civita tensor is invariant under rotation .*

4.4 Translation invariance and Conservation of momentum

Theorem 4.4 *Suppose that the Lagrangian is invariant under common translation of the coordinates $q_1, \dots, q_m$, i.e.*

$$\frac{\partial}{\partial \varepsilon} L(\dot{q}_1, \dots, \dot{q}_n, q_1 + \varepsilon, \dots, q_m + \varepsilon, q_{m+1}, \dots, q_n, t) = 0$$

Then

$$P = \sum_{j=1}^{m} p_j$$

is a constant of motion.

Proof:

$$0 = \frac{\partial L}{\partial \varepsilon} = \sum_{j=1}^{m} \frac{\partial L}{\partial q_j} = \sum_{j=1}^{m} \dot{p}_j$$

and the last equality used Euler-Lagrange equations.

Euclidean space is homogeneous: If we translate a pair of points the vector connecting them does not change. It follows from this that the velocity in a Euclidean space does not care about shifts:

$$\mathbf{v} dt = \mathbf{x}(t + dt) - \mathbf{x}(t) = \big(\mathbf{x}(t + dt) + \mathbf{a}\big) - \big(\mathbf{x}(t) + \mathbf{a}\big)$$

The kinetic energy, $T(\mathbf{v})$ being a function of $\mathbf{v}$ only (in Cartesian coordinates) is translation invariant. This is true for a single particle. Since the kinetic energy is additive, it is also true for many particles.

Remark 4.1 *It is instructive to contrast this with the kinetic energy of a free particle in spherical coordinates:*

$$T(\dot{r}, \dot{\theta}, \dot{\phi}, r, \theta) = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \cos\theta^2 \dot{\phi}^2) \tag{4.4}$$

T now depends both on the velocities and the generalized coordinates r and θ but is independent of ϕ.

In the absence of external fields the potential energy of a system of particles depends on their relative position. In Euclidean space this means

$$U(\mathbf{x_1} + \mathbf{a}, \dots, \mathbf{x_n} + \mathbf{a}) = U(\mathbf{x_1}, \dots, \mathbf{x_n})$$

It follows that the Lagrangian (in Euclidean space)

$$L(\mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{x}_1 \dots, \mathbf{x}_n) = L(\mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{x}_1 + \mathbf{a_1} \dots, \mathbf{x}_n + \mathbf{a}_n)$$

does not depend on the shift $\mathbf{a}$ hence

$$0 = \nabla_{\mathbf{a}} L = \sum_j \nabla_{\mathbf{x}_j} L$$

Combined with Nöther gives conservation of the total momentum

$$0 = \sum_j \nabla_{\mathbf{x}_j} L = \sum_j \dot{\mathbf{p}}_j = \frac{d}{dt}\left(\sum \mathbf{p}_j\right)$$

Remark 4.2 *Gauge freedom allows to replace L by*

$$L \mapsto L + \frac{d\Lambda}{dt} = L + \left(\sum \dot{q}_j A_j\right) + \partial_t \Lambda, \quad A_j = \frac{\partial \Lambda}{\partial q_j}$$

The new Lagrangian remains translation invariant if Λ is. Gauge transformations change the momenta to

$$p_j \mapsto p_j + A_j$$

In the new gauge it is the new total momentum which is conserved.

- This justifies the definition of momentum as $\frac{\partial L}{\partial \dot{q}}$
- Suppose there is translation invariance in some direction, but not in others. In this case, the conserved quantity is the appropriate component of the momentum.

A lovely youtube demo of conservation of momentum on the International Space Station

4.5 Center of mass

Definition 4.7 *In a Euclidean space, where the coordinates are vectors, the center of mass is defined by*

$$\mathbf{X} = \sum \frac{m_j}{M} \mathbf{x}_j, \quad M = \sum m_j$$

For the quadratic kinetic energy the total momentum is

$$\mathbf{P} = \sum m_j \dot{\mathbf{x}}_j$$

It follows

Theorem 4.5 *Conservation of momentum in Euclidean space implies that the center of mass moves at uniform speed*

$$\mathbf{X}(t) = \frac{\mathbf{P}}{M} t + \mathbf{X}_0$$

In particular, if is at rest at $t = 0$ it will stay at rest forever.

This is why Baron von Munchhaussen could not lift himself by pulling on his hair.

The center of mass is a more subtle notion than you may first think.

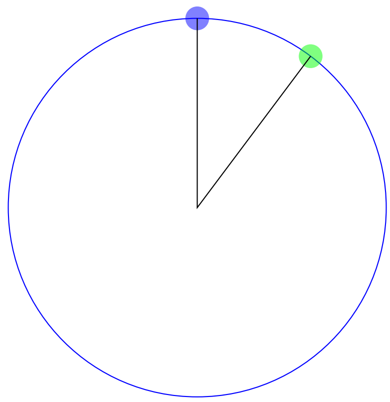


Figure 25: The center of mass on a circle is ill defined

Example 4.3 *Consider for two masses moving on a circle of radius R and connected by a spring. The Lagrangian is*

$$L = \frac{R^2}{2}\left(\sum m_j \dot{\varphi}_j^2\right) - 2kR^2 \sin^2\left(\frac{\varphi_1 - \varphi_2}{2}\right)$$

By Nöther, the total conserved momentum is

$$R^2(m_1\dot{\varphi}_1 + m_2\dot{\varphi}_2)$$

However, since φ is an angle there is no good notion of center of mass. If you define the center of mass by

$$\varphi_{cm} = \frac{m_1}{M}\varphi_1 + \frac{m_2}{M}\varphi_2$$

then the natural equivalence $\varphi \cong \varphi + 2\pi$ forces bizarre equivalence for the center of mass:

$$\varphi_{cm} \cong \varphi_{cm} + 2\pi\left(\frac{m_1}{M}n_1 + \frac{m_2}{M}n_2\right)$$

for all $n_{1,2} \in \mathbb{Z}$.

The absence of a =center of mass is not a disaster. You can still simplify the problem by defining the new generalized coordinates, which do not suffer from this disease:

$$\varphi = \varphi_1, \quad \theta = \varphi_1 - \varphi_2$$

In these

$$L = \frac{R^2}{2}\left(m_1\dot{\varphi}^2 + m_2(\dot{\theta} - \dot{\varphi})^2\right) - 2kR^2 \sin^2\left(\frac{\theta}{2}\right)$$

Since φ is a cyclic coordinate you get the constant of motion

$$p_\varphi = R^2\left((m_1 + m_2)\dot{\varphi} - m_2\dot{\theta}\right)$$

The energy is another constant of motion

$$E = \frac{R^2}{2}\left(m_1\dot{\varphi}^2 + m_2(\dot{\theta} - \dot{\varphi})^2\right) + 2kR^2 \sin^2\left(\frac{\theta}{2}\right)$$

Example 4.4 *Suppose you had two equal masses on a sphere. You could then define their center of mass to be at the midpoint of a geodesic. But, then where would you put the center of mass if the two masses are antipodal?*

Exercise 4.5 *Show that the center of mass of three equal masses on a sphere that are all close by, is ambiguous: It depends on whether you first find the center of mass of 1-2 and then 3 or some other arrangement.*

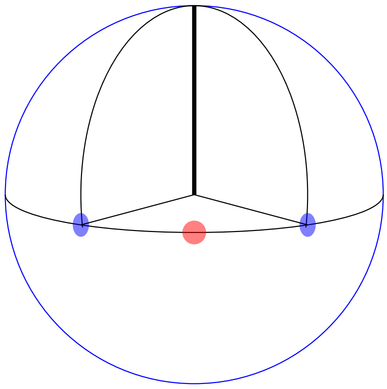


Figure 26: The center of mass in a sphere is an ambiguous notion. The red dot is a good candidate for the center of mass of the blue dots, unless the blue dots are antipodal.

Because of the ambiguity of the notion of center of mass on a sphere, Baron von Muchhausen could move himself if he lived on a sphere .

The notion of center of mass depends on the fact that the kinetic energy is quadratic in velocities. This is illustrated in the following example

Remark 4.3 (Defining the center of mass for relativistic particles) . *The total (conserved) momentum is*

$$P = \sum m_j \mathbf{v}_j \gamma_j = \underbrace{\frac{d}{dt}\left(\sum m_j \gamma_j \mathbf{x}_j\right)}_{center\ mass} - \underbrace{\sum \frac{1}{c^2} \mathbf{x}_j \dot{E}_j}_{extra}, \quad E_j = \underbrace{m_j c^2 \gamma_j}_{kinetic\ energy}$$

The middle term is what you would want to identify with center of mass, but what about the term on the right?. If the particles do not interact you expect the kinetic energy to be conserved and then this term vanishes. What if they interact? You may then worry that you could trade kinetic for potential energy and the last term may be non-zero. This would be bad. Fortunately In relativistic theory potentials of the form $U(\mathbf{x}_j - \mathbf{x}_k)$ are not allowed because they imply that influence can propagate at infinite speed. A legitimate interaction is collision. For a colliding pair, at the moment of collision $\mathbf{x}_j = \mathbf{x}_k$ so

$$\dot{E}_j \mathbf{x}_j + \dot{E}_k \mathbf{x}_k = \mathbf{x}_j \frac{d}{dt}(E_j + E_k)$$

Conservation of energy in a collision implies the vanishing of the last term. We find that the notions of center of mass is well defined but depended on conservation of energy. For quadratic dispersion the notion of center of mass was well defined even without energy conservation. You see from this the the notion of center of mass for non-quadratic dispersions can be subtle.

4.6 Conservation of energy, time translation invariance and the Jacobi integral

Suppose L is time translation symmetric, i.e. does not depend explicitly on time

$$L(\dot{q}, q, \cancel{t}) = L(q, \dot{q})$$

which I can write as

$$\frac{\partial L}{\partial t} = 0$$

By Euler-Lagrange

$$\frac{dL}{dt} = \sum \frac{\partial L}{\partial q_j}\dot{q}_j + \sum \frac{\partial L}{\partial \dot{q}_j}\ddot{q}_j + \underbrace{\frac{\partial L}{\partial t}}_{=0} = \sum \dot{p}_j \dot{q}_j + p_j \ddot{q}_j = \frac{d}{dt}\left(\sum_j p_j \dot{q}_j\right)$$

Rearranging we find a constant of motion, known as the Jacobi integral

$$\frac{dE}{dt} = 0, \quad E = \sum_j p_j \dot{q}_j - L$$

In summary

Theorem 4.6 *If the Lagrangian is invariant under time translation L then the Jacobi integral*

$$E = \left(\sum_j p_j \dot{q}_j\right) - L$$

is a constant of motion. This defines the energy.

4.7 Energy and gauge transformations

Under a gauge transformation

$$L \mapsto L + \sum_j \dot{q}_j \frac{\partial \Lambda}{\partial q_j} + \frac{\partial \Lambda}{\partial t}$$

The new L is still time-translation invariant if

$$\frac{\partial^2 \Lambda}{\partial q_j \partial t} = 0, \quad \frac{\partial^2 \Lambda}{\partial t^2} = 0$$

These two conditions imply that Λ is of the form:

$$\Lambda(q,t) = E_0 t + f(q)$$

This guage transformation changes the momentum and changes the energy by a constant:

$$E \mapsto E + E_0, \quad p_j \mapsto p_j + \frac{\partial \Lambda}{\partial q_j}$$

Exercise 4.6 *Show that for a charged particle in time-independent potentials, with the Lagrangian*

$$L = \frac{m}{2}\mathbf{v}^2 + e\mathbf{v}\cdot\mathbf{A}(\mathbf{x}) - e\Phi(\mathbf{x}),$$

the Jacobi integral is

$$E = \frac{m}{2}\mathbf{v}^2 + e\Phi(\mathbf{x},t) = \frac{(\mathbf{p}-e\mathbf{A})^2}{2m} + e\Phi(\mathbf{x},t),$$

Exercise 4.7 *Show that the Jacobi integral for the relativistic Lagrangian*

$$L = -mc^2/\gamma, \quad \gamma = \frac{1}{\sqrt{1-(\mathbf{v}/c)^2}}$$

is

$$E = mc^2\gamma$$

4.8 Euler theorem

Definition 4.8 *A function f is homogenous of degree n if*

$$f(\lambda q_1, \dots, \lambda q_n) = \lambda^n f(q_1, \dots, q_n)$$

Euler theorem says:

Theorem 4.7 *If f is homogeneous of degree n then*

$$\sum_j q_j \frac{\partial f}{\partial q_j} = nf$$

The proof follows by differentiating wrt λ

$$\frac{\partial f}{\partial \lambda} = \sum_j \frac{\partial f}{\partial(\lambda q_j)}\frac{d(\lambda q_j)}{d\lambda} = \sum_j \frac{q_j}{\lambda}\frac{\partial f}{\partial q_j} = n\lambda^{n-1} f(q)$$

Substituting $\lambda = 1$ we get Euler theorem [4].

[4] $(\mathbf{x}\cdot\nabla)f = \mathbf{x}\cdot\nabla f$

4.9 Homogeneous kinetic energy and Jacobi integral

Theorem 4.8 *Suppose*

$$L = T(\dot q_1, \ldots, \dot q_n, q_1, \ldots, q_n) - U(q_1, \ldots, q_n)$$

where T is homogeneous of degree n in the velocities $\dot q_j$. Then the Jacobi integral is

$$E = (n-1)T + U$$

In particular, if $n = 2$ the Jacobi integral is the sum of the kinetic and potential energies.

Proof:

$$E = \sum \dot q_j p_j - L = \sum \dot q_j \frac{\partial T}{\partial \dot q_j} - L = nT - L = (n-1)T + U$$

Theorem 4.9 *The kinetic energy of a non-relativistic particle is homogeneous of degree 2 in $\dot q_j$ for arbitrary choice of generalized coordinates, (e.g. spherical)*

$$T = \frac{m}{2}\mathbf{v}^2 = \frac{m}{2}\left(\sum \dot x_j^2(q_1, q_2, q_3)\right) \frac{m}{2}\sum_j \left(\sum_k \dot q_k \frac{\partial x_j}{\partial q_k}\right)^2$$

4.10 Energy for time dependent Lagrangians

How shall we define the energy if L is time dependent? Often the time dependence comes from the fact that we selected part of the degrees of freedom as our system and think of the rest as external forces. For example, consider a solar system with a Sun, Jupiter and Earth. The Lagrangian for the full system is time independent

$$L_{sje} = \underbrace{T_s + T_j - U_{sj}}_{sun+jupiter} + \underbrace{T_e - U_{se} - U_{je}}_{earth}$$

and I have arranged the Lagrangian so that the first term is large and the second small. The smallness of the second term comes from the fact that both T_e and $U_{e,su}$ are proportional to the mass of the earth which is the lightest of the three.

The total energy of the system

$$E_{sje} = \underbrace{T_s + T_j + U_{sj}}_{sun+jupiter} + \underbrace{T_e + U_{se} + U_{je}}_{earth}$$

is conserved. Note that there is no natural way to partition the energies of 3 bodies to energies of 2 bodies and 1 body in general. The partition we made is motivated by the different scales off masses: The first term is much larger than the second.

The system is a three body problem and is quite complicated. Since the earth is much smaller that the Sun and Jupiter a good strategy to approximately solve it is by divide and conquer: Solve for the motion of Sun and Jupiter pretending that earth does not affect their motion. This means, first reduce the 3 body problem to the 2-body problem, with a time-independent but approximate Lagrangian

$$L_{sj} = T_s + T_j - U_{sj}$$

Since this is a 2-doby problem we can solve it explicitly and find the orbits $\mathbf{x}_s(t)$ and $\mathbf{x}_j(t)$. Now that the orbits of the sun and Jupiter are known we consider the earth effective Lagrangian

$$L_e = T_e - U_{se} - U_{je}$$

L_e is time dependent because the orbits of the Sun and Jupiter are prescribed

$$U_{se} = U_{se}\big(\mathbf{x}_e - \mathbf{x}_s(t)\big), \quad U_{je} = U_{je}\big(\mathbf{x}_e - \mathbf{x}_j(t)\big)$$

This procedure gives a reasonable partition of the energies:

$$E_{sje} = E_{sj} + (T_e + U_{se} + U_{je})$$

The second term is formally the Jacobi integral for the effective Lagrangian of earth. The first term on the right, the Jupiter-Sun energy, is determined by the global conservation law. Note that E_{sj} is not quite the energy associated to L_{sj}. The latter is approximate and time independent. The former is time dependent. This motivates

Definition 4.9 *For a Lagrangian L, irrespective if it is time dependent or not, the energy is defined by*

$$E = \sum p_j \dot{q}_j - L$$

Youtube demo of conservation of energy

Remark 4.4 (Gauge freedom) *When the Lagrangian is time dependent the definition of the energy has the unpleasant feature of being gauge dependent. There is no such problem when L is time independent. Under gauge transformations*

$$E \mapsto E - \frac{\partial \Lambda}{\partial t}$$

The energy is a gauge dependent quantity. It is gauge independent if we allow only for gauge transformations that are not explicitly time dependent:

$$\frac{\partial \Lambda}{\partial t} = 0$$

4.11 Conserved angular momentum

If L is isotropic, the corresponding conserved quantity is the angular momentum. This can be viewed as a special case of conservation of momentum, when the generalized coordinates are angles. For example consider the Lagrangian with cylindrical symmetry

$$L = \underbrace{\frac{m}{2}(\dot{\rho}^2 + (\rho\dot{\theta})^2 + \dot{z}^2)}_{T} - U(z, \rho)$$

Since L is independent of θ we get the conserved quantity

$$J_z = \frac{\partial L}{\partial \dot{\theta}} = m\rho^2\dot{\theta}$$

which you recognize as angular momentum about the z axis.

4.12 More on rotations

Rotations in 3 dimensions are represented by 3×3 orthogonal matrices

$$\mathbf{x} \mapsto R\mathbf{x}, \quad R^t = R^{-1}$$

(with real coefficients). No rotation at all is represented by the identity matrix $\mathbb{1}$. A small rotation must then be of the form

$$R = \mathbb{1} - \delta F$$

with δF a small matrix. Orthogonal implies

$$\mathbb{1} = R^t R = \mathbb{1} - (F + F^t) \Longrightarrow F + F^t = 0$$

It follows that F is anti-symmetric.

Now, the Levi-Civita tensor allows us to map the anti-symmetric 3×3 matrix δF to a vector $\delta\boldsymbol{\phi}$

$$\delta\boldsymbol{\phi}_i = \frac{1}{2}\varepsilon_{ijk}\delta F_{jk} \Longleftrightarrow \delta F_{jk} = \varepsilon_{jki}\delta\boldsymbol{\phi}_i$$

It follows that

$$\begin{aligned}
(R\mathbf{x})_i &= \mathbf{x}_i - (\delta F\mathbf{x})_i \\
&= \mathbf{x}_i - \delta F_{ij}\mathbf{x}_j \\
&= \mathbf{x}_i - \varepsilon_{ijk}\delta\boldsymbol{\phi}_k\mathbf{x}_j \\
&= \mathbf{x}_i + (\delta\boldsymbol{\phi} \times \mathbf{x})_i
\end{aligned}$$

Hence, a small rotation can be represented by an infinitesimal vector $\delta\boldsymbol{\phi}$ that acts by

$$\mathbf{x} \mapsto \mathbf{x} + \delta\boldsymbol{\phi} \times \mathbf{x} \tag{4.5}$$

(and I neglect terms quadratic in $\delta\boldsymbol{\phi}$. Since vectors $\mathbf{x} \| \delta\boldsymbol{\phi}$ do not move, we identify $\widehat{\delta\boldsymbol{\phi}}$ with the axis of rotation and the (small) rotation angle with the length.

Exercise 4.8 *Show that any rotation, not necessarily small, can be written as*

$$\mathbf{x} \mapsto \left(\mathbf{x} \cdot \widehat{\boldsymbol{\phi}}\right) \widehat{\boldsymbol{\phi}} + (\sin\phi)\, \widehat{\boldsymbol{\phi}} \times \mathbf{x} - (\cos\phi)\, \widehat{\boldsymbol{\phi}} \times (\widehat{\boldsymbol{\phi}} \times \mathbf{x})$$

By the rules of tensor calculus, this then holds for any vector, e.g. the velocity

$$\mathbf{v} \mapsto \mathbf{v} + \delta\boldsymbol{\phi} \times \mathbf{v}$$

Euclidean space is isotropic. This means that the scalar product $\mathbf{a} \cdot \mathbf{b}$ of any two vectors does not care about rotation.

4.13 Isotropy and conservation of angular momentum

Isotropy means that L is a function of scalar

$$L = L\left(\mathbf{v}_1^2, \ldots, \mathbf{v}_n^2, \mathbf{x}_1^2, \ldots, \mathbf{x}_n^2, (\mathbf{x}_j - \mathbf{x}_k)^2, \ldots, (\mathbf{x}_n - \mathbf{x}_m)^2\right) \tag{4.6}$$

This is the case for Coulomb and gravitational forces.

For simplicity, let me look at the case of a single particle. Using Einstein summation convention over repeated indexes

$$\begin{aligned}
0 = \frac{\partial L}{\partial \phi_k} &= \underbrace{\frac{\partial L}{\partial x_a}\frac{\partial x_a}{\partial \phi_k} + \frac{\partial L}{\partial \dot{x}_a}\frac{\partial \dot{x}_a}{\partial \phi_k}}_{use\ Euler-Lagrange} \\
&= \underbrace{\dot{p}_a \frac{\partial x_a}{\partial \phi_k} + p_a \frac{\partial \dot{x}_a}{\partial \phi_k}}_{use\ Eq.(4.5)} \\
&= \underbrace{\dot{p}_a \epsilon_{akm} x_m + p_a \epsilon_{akm} \dot{x}_m}_{Leibnitz\ rule} \\
&= \underbrace{\frac{d}{dt}\left(p_a \epsilon_{akm} x_m\right)}_{cyclicity\ of\ \varepsilon} \\
&= \frac{d}{dt}\left(x_m p_a \epsilon_{mak}\right) \\
&= \frac{d}{dt}\left(\mathbf{x} \times \mathbf{p}\right)_k
\end{aligned}$$

It follows that if L is a Euclidean scalar, i.e. isotropic, then the angular momentum

$$\mathbf{J} = \mathbf{x} \times \mathbf{p}$$

is conserved.

For a multi-particle system you similarly get the conservation of the total angular momentum:

$$\mathbf{J} = \sum_j \mathbf{x}_j \times \mathbf{p}_j$$

Youtube demo conservation of angular momentum in space

Remark 4.5 *Conservation laws follow from symmetry and Euler-Lagrange equations. If you put back the conservation laws into the Lagrangian and compute Euler-Lagrange you get nonsense. For example, for a free particle in the plane with polar coordinates*

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2)$$

Angular momentum is conserved

$$J = mr^2\dot{\theta}$$

It is not allowed to put that back into L. If you do you get

$$L = \frac{m}{2}\dot{r}^2 + \frac{J^2}{2mr^2}$$

which gives correctly the size of the centrifugal barrier but a wrong sign.

5 Rotations without angular momentum

We shall talk later about rigid body motions. But you probably know from previous classes that a rigid body can rotate only if it has angular momentum. This is not true for deformable bodies, as every cat knows.

Youtube movie of falling cat

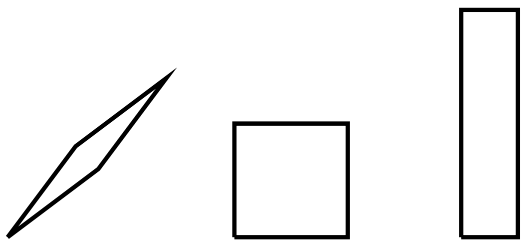


Figure 27: The figure shows two types of scaling transformation, call them A and B. For both, by symmetry, there is no angular momentum about the center of the square. This is a transformations that a deformable body. like a cat, is allowed to make with internal forces that satisfy Newton third law, alone. They are not associated with angular momentum: Scaling does not generate angular momentum. The order matters, however. So $ABA^{-1}B^{-1}$ is not the identity. If A, B are close to the identity, the remainder is a rotation.

An animation showing that non commuting deformations give a rotation

5.1 Constrained deformable triangles

Consider three point masses with $m = 1$ in the plane, located at $\mathbf{x}_j$, $j = 1, 2, 3$. The masses are connected by mass-less rods of lengths $\ell_{ij}(t)$ which are viewed as time dependent constraints. There are three holonomic constraints:

$$(\mathbf{x}_j - \mathbf{x}_k)^2 = \ell_{jk}^2(t)$$

The corresponding Lagrangian has three (time dependent) Lagrange multipliers:

$$L = \frac{1}{2}\left(\mathbf{v}_1^2 + \mathbf{v}_2^2 + \mathbf{v}_3^2\right) + \frac{\lambda_{12}}{2}(\mathbf{x}_1 - \mathbf{x}_2)^2 + \frac{\lambda_{23}}{2}(\mathbf{x}_2 - \mathbf{x}_3)^2 + \frac{\lambda_{31}}{2}(\mathbf{x}_3 - \mathbf{x}_1)^2$$

The Lagrangian is invariant under identical shifts of all coordinates, so momentum is

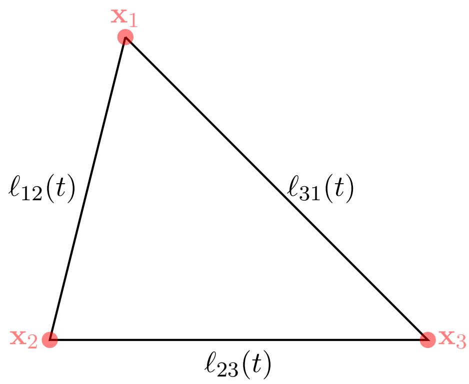


Figure 28: Triangular cat: The lengths of the edges are constrained to be given functions of time. You can make the triangle rotate with 0 angular momentum.

conserved. It follows that if the center of mass starts at rest at the origin, it will remain at rest at the origin:

$$\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3 = 0$$

The Lagrangian is made from scalars $\dot{\mathbf{x}}^2$ and $(\mathbf{x}_j - \mathbf{x}_k)^2$ and consequently it is invariant under joint rotations and angular momentum is conserved. If we assume zero angular momentum to start with then it will remain zero

$$\mathbf{x}_1 \times \mathbf{v}_1 + \mathbf{x}_2 \times \mathbf{v}_2 + \mathbf{x}_3 \times \mathbf{v}_3 = 0$$

There is no energy conservation, however. This is because the constraints are time dependent.

The Euler Lagrange equations with constraints are

$$\dot{\mathbf{v}}_1 = \lambda_{12}(t)(\mathbf{x}_1 - \mathbf{x}_2) - \lambda_{31}(t)(\mathbf{x}_3 - \mathbf{x}_1) \tag{5.1}$$

and cyclic permutations. The equations of motion depend on the λ which are not known a-priori.

This makes setting up the equations of motion for the constrained problem hard. Normally, we use the equations of motion to update $(\mathbf{x}, \mathbf{v})$

$$\mathbf{x}(t + dt) = \mathbf{x}(t) + \mathbf{v}dt, \quad \mathbf{v}(t + dt) = \mathbf{v}(t) + \mathbf{a}(t)dt$$

but here we do not know really $\mathbf{a}$ from Eq. 5.1 because we do not know λ.

We can get rid of λ's by sacrificing 3 equations of motions: If you hit Eq. **??** with cross product by $\mathbf{x} - \mathbf{x}_2$ you get

$$\mathbf{a}_1 \times (\mathbf{x}_1 - \mathbf{x}_2) = \lambda_{31}(\mathbf{x}_3 - x_1) \times (\mathbf{x}_1 - \mathbf{x}_2)$$

(and cyclic permutations.) The vector is pointing perpendicular to the plane of the triangle, which we call $\hat{\mathbf{z}}$. This allows us to expressing λ in terms of the coordinates and their acceleration:

$$\lambda_{31}(t) = \underbrace{\frac{\mathbf{a}_1 \times (\mathbf{x}_1 - \mathbf{x}_2) \cdot \hat{\mathbf{z}}}{(\mathbf{x}_3 - \mathbf{x}_1) \times (\mathbf{x}_1 - \mathbf{x}_2) \cdot \hat{\mathbf{z}}}}_{linear\ in\ acceleration} \tag{5.2}$$

and its cyclic permutations. Now that we have λ we return to the components of the equation of motion we have not used, namely

$$\mathbf{a}_1 \cdot (\mathbf{x}_1 - \mathbf{x}_2) = \underbrace{\lambda_{12}\ell_{12}^2 + \lambda_{31}(\mathbf{x}_3 - \mathbf{x}_1) \cdot (\mathbf{x}_1 - \mathbf{x}_2)}_{linear\ in\ acceleration} \tag{5.3}$$

and its cyclic permutations. If we now substitute Eq. 5.2 in Eq. 5.3 we get three homogeneous, linear equations for the (six) accelerations. We get three more linear, but non-homogeneous equations for the acceleration by differentiating twice the constraint:

$$\underbrace{(\mathbf{x}_1 - \mathbf{x}_2) \cdot (\mathbf{a}_1 - \mathbf{a}_2) + (\mathbf{v}_1 - \mathbf{v}_2)^2}_{linear\ in\ acceleration} = \ell_{12}(t)\ddot{\ell}_{12}(t) + \dot{\ell}_{12}^2(t)$$

At the end of the day we get 6 *linear* equations for the 6 accelerations with coefficients that are functions of the coordinates, the velocities and the constraints and their derivatives. So, if we assume that we know how to solve 6 linear equations (computers do that with no sweat) we can bring the equations of motion to Newton's standard form

$$\mathbf{a_j} = \mathbf{F}_j(\mathbf{x}_1, \ldots, \mathbf{v}_1, \ldots, \ell, \dot{\ell}, \ddot{\ell})$$

This allows for updating $(\mathbf{x}, \mathbf{v})$ in time.

6 Virial theorem

The (non-relativistic) kinetic energy is quadratic in velocities

$$T(\mathbf{v}) = \sum \frac{m_j}{2}\mathbf{v}_j^2$$

and so homogeneous of degree 2:

$$T(\lambda\mathbf{v}) = \lambda^2 T(\mathbf{v})$$

The gravitational energy of a system of particles

$$U(\mathbf{x}) = -\frac{G}{2}\sum \frac{m_j m_k}{|\mathbf{x}_j - \mathbf{x}_k|}$$

is homogeneous of degree -1

$$U(\lambda \mathbf{x}) = \lambda^{-1} U(\mathbf{x})$$

The same is true for the electrostatic energy of a system of particles.

The potential energy of a collections of oscillator

$$U(\mathbf{q}) = \frac{1}{2} \sum K_{jk} q_j q_k$$

is homogeneous of degree 2, i.e.

$$U(\lambda \mathbf{q}) = \lambda^2 U(\mathbf{q})$$

K is the matrix of spring constants.

This motivates the interest in Lagrangians of the form

$$L = T(\mathbf{v}) - U(\mathbf{x})$$

with T and U homogeneous functions of in general different degrees.

6.1 Time averages

Assume that $L = T - U$ with T homogeneous of degree m and U homogeneous of degree n

$$\begin{aligned}
\frac{d(\mathbf{p} \cdot \mathbf{q})}{dt} &= \mathbf{p} \cdot \dot{\mathbf{q}} + \dot{\mathbf{p}} \cdot \mathbf{q} \\
&= \frac{\partial L}{\partial \dot{\mathbf{q}}} \cdot \dot{\mathbf{q}} + \frac{\partial L}{\partial \mathbf{q}} \cdot \mathbf{q} \\
&= \dot{\mathbf{q}} \cdot \frac{\partial T}{\partial \dot{\mathbf{q}}} - \mathbf{q} \cdot \frac{\partial U}{\partial \mathbf{q}} \\
&= (\dot{\mathbf{q}} \cdot \nabla_{\dot{q}}) T - (\mathbf{q} \cdot \nabla_q) U \\
&= mT - nU
\end{aligned}$$

Definition 6.1 *Let $F(t)$ be a function of time. The time average of F is defined by*

$$\langle F \rangle = \lim_{T \to \infty} \frac{1}{T} \int_0^T F\big(t\big) dt$$

Theorem 6.1 *Let G be a bounded function then*

$$\left\langle \frac{dG}{dt} \right\rangle = 0$$

Proof:

$$\left\langle \frac{dG}{dt} \right\rangle = \lim_{T \to \infty} \frac{1}{T} \int_0^T \frac{dG}{dt} dt = \lim_{T \to \infty} \frac{G(T) - G(0}{T} = 0$$

It then follows

Theorem 6.2 (Virial theorem) *Suppose that the kinetic energy is homogeneous in the velocities of degree m and that the potential energy is homogeneous of degree n and that neither $\mathbf{p}(t)$ not $\mathbf{q}(t)$ escape to infinity even if you wait for a long time. Then $\langle \mathbf{p} \cdot \mathbf{q} \rangle = 0$ and the average kinetic and potential energies are related by*

$$m \langle T \rangle = n \langle U \rangle$$

- For Harmonic potential the average kinetic energy and potential energies are the same. This lies at the heart of the equipartition theorem in statistical mechanics.
- In a bound gravitating system

  $$2 \langle T \rangle = - \langle U \rangle$$

  and the binding energy is minus the average kinetic energy:

  $$E = \langle T \rangle + \langle U \rangle = - \langle T \rangle$$

6.2 Dark matter

Weighing galaxies: Let M be the mass of a nice and uniform galaxy with $N \gg 1$ stars. Then

$$\langle T \rangle = \frac{1}{2} \sum_j m_j \mathbf{v}_j^2 = \frac{M}{2} \left\langle \mathbf{v}^2 \right\rangle, \quad M = \sum_j m_j$$

The average velocity (here the average weighted by the mass) is something we can hope to measure.

Recall from E&M that the self energy of a uniformly charged ball of radius R

$$\frac{3}{5} \frac{Q^2}{R}$$

This means that a galaxy that looks like a uniform ball stars with radius R

$$\langle U \rangle \approx -\frac{3}{5R} G M^2$$

R is something we can measure.

Exercise 6.1 *Why the minus?*

The virial can be used to estimate the mass of a galaxy M from measurable quantities:

$$M \left\langle \mathbf{v}^2 \right\rangle \approx \frac{3}{5R} G M^2 \Longrightarrow M \approx \frac{5}{3G} \left\langle \mathbf{v}^2 \right\rangle R$$

A puzzle arises because there is also a second way to estimate M. The theory of stellar structures relates the mass of Galaxies (and stars) to the radiation it emits: The type of light spectrum a star emits is related to its mass. So by looking at the light emitted we get another way to estimate M. It turns out that the mass estimate from the virial is much larger than the mass estimate from observing the visible light. This was taken by Zwicky as evidence that there is dark matter that you do not see.

7 The Kepler problem

We turn to study planetary orbits.

7.1 Reduction of the two body problem

Sometimes a clever choice of generalized coordinates allows to write the Lagrangian as a sum of two independent Lagrangians:

$$L_1(\dot{\mathbf{q}}_1, \dot{\mathbf{q}}_2, \mathbf{q}_1, \mathbf{q}_2) = L_1(\dot{\mathbf{q}}_1, \mathbf{q}_1) + L_2(\dot{\mathbf{q}}_2, \mathbf{q}_2)$$

When this is case the , $\mathbf{q}_1$ coordinates move independently of the $\mathbf{q}_2$ coordinates: Divide and and conquer.

Here is how this works for two bodies:

The center of mass and relative distance, $(\mathbf{X}, \mathbf{r})$ are related to the coordinates of the particles $(\mathbf{x}_1, \mathbf{x}_2)$ by

$$\mathbf{X} = \frac{m_1\mathbf{x}_1 + m_2\mathbf{x}_2}{M}, \quad \mathbf{r} = \mathbf{x}_2 - \mathbf{x}_1$$

and the converse

$$\mathbf{x}_2 = \mathbf{X} + \frac{m_1}{M}\mathbf{r}, \quad \mathbf{x}_1 = \mathbf{X} - \frac{m_2}{M}\mathbf{r}$$

If the kinetic energy is quadratic in the velocities a magic occurs:

Theorem 7.1

$$\frac{m_1}{2}\dot{\mathbf{x}}_1^2 + \frac{m_2}{2}\dot{\mathbf{x}}_2^2 = \frac{M}{2}\dot{\mathbf{X}}^2 + \frac{\mu}{2}\dot{\mathbf{r}}^2,$$

where

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2}, \quad M = m_1 + m_2$$

μ is called the reduced mass.

The proof is by algebra:

$$\begin{aligned}
T &= \frac{m_1}{2}\mathbf{v}_1^2 + \frac{m_2}{2}\mathbf{v}_2^2 \\
&= \underbrace{\frac{m_1}{2}\left(\dot{\mathbf{X}} - \frac{m_2}{M}\,\dot{\mathbf{r}}\right)^2 + \frac{m_2}{2}\left(\dot{\mathbf{X}} + \frac{m_1}{M}\,\dot{\mathbf{r}}\right)^2}_{\textit{Mixed terms drop}} \\
&= \frac{M}{2}\dot{\mathbf{X}}^2 + \frac{\mu}{2}\dot{\mathbf{r}}^2,
\end{aligned}$$

Fact 7.1 *If $m_1 \gg m_2$ then $M \approx m_1$ and $\mu \approx m_2$*

It follows

Theorem 7.2 *If two particles interact with a potential that is a function of their relative distance and the Kinetic energy is quadratic the two body problem reduces to two independent one body problems:*

$$L = T - U = \frac{M}{2}\dot{\mathbf{X}}^2 + \left(\frac{\mu}{2}\dot{\mathbf{r}}^2 - U(\mathbf{r})\right)$$

The center of mass motion decouples from the relative motion. The center of mass move on a straight line at uniform velocity. The relative motion is a one body problem in an external potential independent of the center of mass.

7.2 The case of 3 particles–The Mass matrix

To appreciate the simplicity of the two body problem consider the three body problem where the kinetic energy T is

$$T = \frac{m_1}{2}\dot{\mathbf{x}}_1^2 + \frac{m_2}{2}\dot{\mathbf{x}}_2^2 + \frac{m_3}{2}\dot{\mathbf{x}}_3^2$$

The center of mass is

$$\mathbf{X} = \frac{m_1\mathbf{x}_1 + m_2\mathbf{x}_2 + m_3\mathbf{x}_3}{M},$$

There are many ways to choose the relative coordinates. For example, you may chose the relative coordinates to be

$$\mathbf{r}_{13} = \mathbf{x}_1 - \mathbf{x}_3, \quad \mathbf{r}_{23} = \mathbf{x}_2 - \mathbf{x}_3$$

but you could make other choices. The kinetic energy can be written as the sum of the kinetic energy of center of mass and the kinetic energy of the relative coordinates:

$$T = \frac{M}{2}\dot{\mathbf{X}}^2 + T_r$$

But now T_r is given by

$$\begin{aligned} T_r &= \frac{m_1}{2}\left(1 - \frac{m_1}{M}\right)\dot{\mathbf{r}}_{13}^2 - \frac{m_1 m_2}{M}\dot{\mathbf{r}}_{13}\cdot\dot{\mathbf{r}}_{23} + \frac{m_2}{2}\left(1 - \frac{m_2}{M}\right)\dot{\mathbf{r}}_{23}^2 \\ &= \frac{1}{2}(\dot{\mathbf{r}}_{13}, \dot{\mathbf{r}}_{23})\,\mathbf{M}\,(\dot{\mathbf{r}}_{13}, \dot{\mathbf{r}}_{23})^t \end{aligned} \tag{7.1}$$

where

$$\mathbf{M} = \begin{pmatrix} m_1(1 - m_1/M) & -m_2 m_2/M \\ -m_1 m_2/M & m_2(1 - m_2/M) \end{pmatrix}$$

is a mass matrix.

The off diagonal terms couples the two relative coordinates (it even comes with a minus sign!) If $m_3 \gg m_1, m_2$ the off-diagonal terms are much smaller than the terms on the diagonal as you may expect.

The point of this example is to show that the general form of the relative kinetic energy once the center of mass has been removed, is a quadratic function of the velocities with a mass matrix that is not diagonal in general:

$$T_r = \frac{1}{2}\sum \dot{q}_j M_{jk}\dot{q}_k$$

Since $T_r \geq 0$ the matrix M is a positive matrix. We shall come back to the mass matrix when we study small oscillations.

7.3 Central force: Kepler second law

Definition 7.1 *U is central if*

$$U(\mathbf{r}) = U(|\mathbf{r}|)$$

It follows from Nöther theorem that the motion in a central potential has conserved angular momentum:

Theorem 7.3 *If*

$$L = T(|\mathbf{v}|) - U(|\mathbf{r}|), \quad \mathbf{v} = \dot{\mathbf{r}}$$

then

- $\mathbf{J} = \mathbf{r} \times \mathbf{p}$ *is a constant of motion.*
- $\mathbf{p} \| \mathbf{v}$
- $\dot{\mathbf{v}}$ *is parallel to* $\mathbf{r}$ *if* T *is quadratic. For general* T *the acceleration is in the plane spanned by* $\mathbf{r}$ *and* $\mathbf{v}$*, i.e the plane associated with* $\mathbf{J}$*. In other words* $\mathbf{r} \times \dot{\mathbf{v}} \| \mathbf{J}$

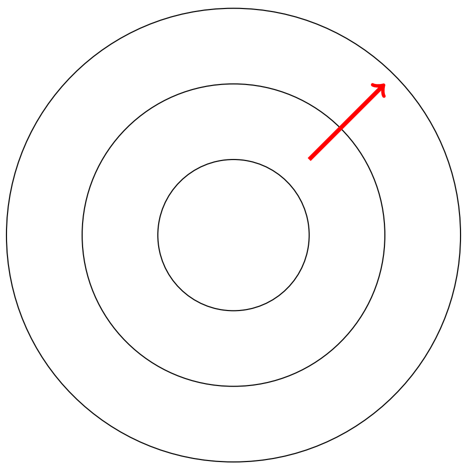


Figure 29: The gradient of T is radial

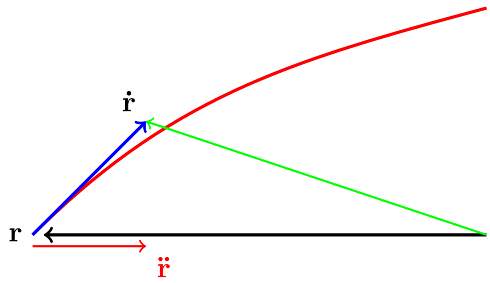


Figure 30: The red curve is part of the orbit. The black vector is $\mathbf{r}$ and the blue vector is $d\mathbf{r}$. $\mathbf{J}dt/\mu$ is twice the area of the triangle.

Proof: $\mathbf{p}$ is parallel to $\mathbf{v}$ since T is a function of $|\mathbf{v}|$. In fact

$$\mathbf{p} = \hat{\mathbf{v}}T' = m(|\mathbf{v}|)\mathbf{v}$$

That the acceleration $\dot{\mathbf{v}}$ is in the plane spanned by $\mathbf{r}$ and $\dot{\mathbf{r}}$ follows from the conservation of $\mathbf{J}$

$$0 = \dot{\mathbf{J}} = \underbrace{\dot{\mathbf{r}} \times \mathbf{p}}_{=0} + \mathbf{r} \times \dot{\mathbf{p}} = m\mathbf{r} \times \dot{\mathbf{v}} + \underbrace{\dot{m}\mathbf{r} \times \mathbf{v}}_{\|\mathbf{J}}$$

In the case that T is quadratic, $\dot{m} = 0$ and the acceleration is parallel to $\mathbf{r}$. It follows

Theorem 7.4 *The three vectors $\mathbf{r}, \dot{\mathbf{r}}, \ddot{\mathbf{r}}$ are co-planar and the motion takes place in one plane. The vector $\mathbf{J}$ is orthogonal to this plane. Finally the radius vector covers equal areas at equal times*

$$\mathbf{J}\,dt = \mathbf{r} \times \mathbf{p}\,dt = \mu \underbrace{\mathbf{r} \times \mathbf{dr}}_{area}$$

Proof: Since

$$\mathbf{r} \mapsto \mathbf{r} + \dot{\mathbf{r}}dt, \quad \dot{\mathbf{r}} \mapsto \dot{\mathbf{r}} + \ddot{\mathbf{r}}dt,$$

and all these vectors are co-planar, the motion stays in the same plane. An alternate way to say this is that since $\mathbf{J}$ is conserved the plane has a fixed orientation.

We have thus proved:

Law 7.1 (Kepler second law) *The orbit of planets lie in a plane that includes the sun. The radius vector to the sun covers equal areas in equal times.*

Exercise 7.1 (Impossible orbits) *Conservation of angular momentum can be used to rule out impossible orbits. In the figure, the central force is centered at the black dot. But you do not know if it is attracting or repelling. Use conservation of angular momentum to decide.*

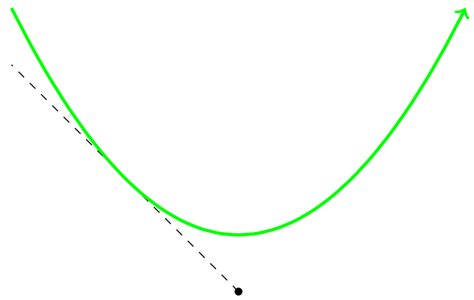


Figure 31: The central force is centered at the black dots. The green orbit does not conserve angular momentum since it changes sign as it touches the dashed tangent.

7.4 Effective potential

The Lagrangian for a planar motion in a central potential in polar coordinates[5]

$$L = \frac{\mu}{2}(\dot{r}^2 + (r\dot{\theta})^2) - U(r) \tag{7.2}$$

θ is cyclic and the associated constant of motion is the angular momentum

$$J = \mu r^2 \dot{\theta} \tag{7.3}$$

Since the Lagrangian is time independent energy is conserved. Combined with the conservation of J

$$\begin{aligned} E &= \frac{\mu}{2}(\dot{r}^2 + (r\dot{\theta})^2) + U(r) \\ &= \frac{\mu}{2}\left(\dot{r}^2 + \left(\frac{J}{\mu r}\right)^2\right) + U(r) \\ &= \frac{\mu}{2}\dot{r}^2 + U(r) + \underbrace{\frac{J^2}{2\mu r^2}}_{centrifugal} \\ &= \frac{\mu}{2}\dot{r}^2 + \underbrace{U_J(r)}_{effective\ potential} \end{aligned} \tag{7.4}$$

Non-zero J raises barrier that prevents falling to the origin even if $U(r)$ is an attractive. The centrifugal barrier beats the Coulomb attraction at small distances.

Exercise 7.2 *Use the centrifugal barrier to argue why many Galaxies and solar systems are disk shaped.*

Example 7.1 *For a relativistic particle*

$$L = -mc^2\sqrt{1-\beta^2} - U(r), \quad \beta^2 = (\dot{r}^2 + (r\dot{\theta})^2)/c^2 \tag{7.5}$$

[5]This is actually subtle since we substituted a conservation law into the equation. Try to justify this.

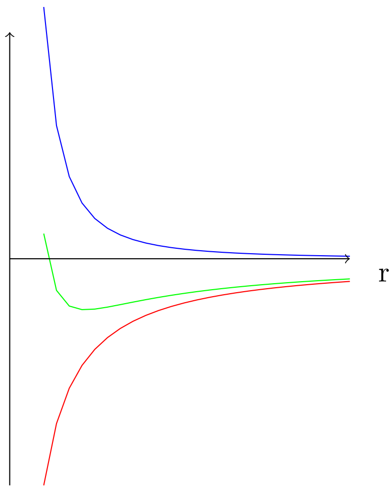


Figure 32: The red curve is the gravitational potential. The blue curve is the centrifugal barrier. The green is the effective potential.

Conservation of angular momentum is

$$J = \partial_{\dot\theta} L = (\partial_{\beta^2} L)(\partial_{\dot\theta}\beta^2) = m\gamma r^2 \dot\theta$$

The energy is

$$E = mc^2\gamma + U(r) \tag{7.6}$$

The relation between γ *and* J *can be found from*

$$c^2\beta^2 = \dot r^2 + (r\dot\theta)^2 = \dot r^2 + \frac{J^2}{m^2\gamma^2 r^2} = c^2\left(1 - \frac{1}{\gamma^2}\right) \tag{7.7}$$

or

$$\dot r^2 + \frac{1}{\gamma^2}\left(\frac{J^2}{m^2 r^2} + c^2\right) = c^2 \tag{7.8}$$

This gives a quadratic equation for γ. *In analogy with the non relativistic case let me define the centrifugal potential as the total kinetic energy when the radial velocity vanishes. This gives the relativistic analog of the centrifugal potential*

$$mc^2\gamma\Big|_{\dot r = 0} = mc^2\sqrt{1 + \frac{J^2}{m^2c^2r^2}} \quad \lim_{r\to 0} \quad \frac{|J|c}{r} \tag{7.9}$$

The centrifugal barrier is weaker at short distances for relativistic particle. It may or may not balance gravity depending on the sign on $|J|c - GMm$ *is. This is why massive stars can collapse into black holes.*

7.5 The period in a radial potential

Eq. 7.4 can be written as a non-linear, first order ODE:

$$dt = \frac{\sqrt{\mu}\, dr}{\sqrt{2(E - U_J(r))}}$$

It gives $t(r)$. In particular, for a periodic orbit the period (in r) can be computed by integrating

$$\frac{T}{2} = \int_{r_{min}}^{r_{max}} \frac{\sqrt{\mu}\, dr}{\sqrt{2(E - U_J(r))}} \tag{7.10}$$

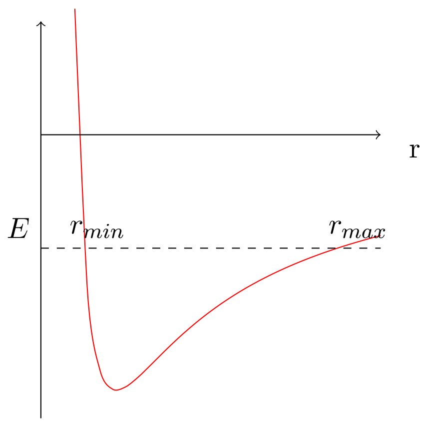


Figure 33: The effective potential, r_{min} and r_{max}.

Remark 7.1 *The period in r does not mean that the motion in the plane is periodic*

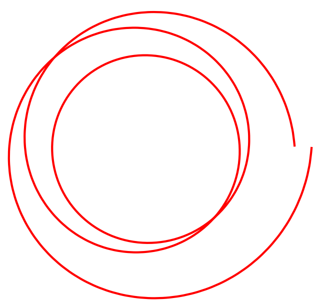


Figure 34: An orbit in the plane that does not close. The period of the radial motion between r_{min} and r_{max} is not a period of the full motion.

7.6 Ellipses

We shall now derive Kepler first law from Newton's equations:

Law 7.2 (Kepler first law) *Planets move on ellipses with the sun at the focal point.*

Newton showed this using ingenious geometric constructions. The derivation I shall present is modern, simpler, but relies on having the path paved by Newton: We know what to look for.

Definition 7.2
- *An ellipse (hyperbola) is the locus of points in the plane so that the sum (difference) of distances from the two focii is $2a > 0$ fixed. The two focii are at separated by $2f \geq 0$.*
- *From the triangle inequality inn an ellipse $2a \geq 2f > 0$ while in a hyperbola $2f \geq 2a > 0$.*
- *The the eccentricity of an ellipse is defined by*

$$0 \leq e = \frac{f}{a} < 1$$

- *When $f = 0$ the ellipse degenerates to a circle of radius a. When $f = a$ it degenerates to the straight line connecting the focii.*

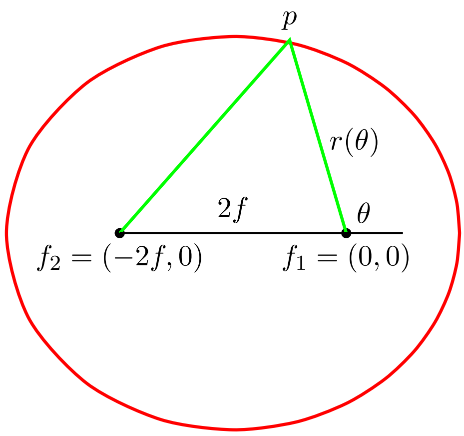


Figure 35: The red curve is an ellipse. The two black dots are the focal points one at the origin $(0,0)$ and one at $(-2f, 0)$.

We can now use elementary analytic geometry to turn the geometric definition in to a formula.

Theorem 7.5 *The equation of an ellipse with a focal point at the origin in polar coordinates is*

$$r(\theta) = \frac{a^2 - f^2}{a + f\cos(\theta)} > 0, \quad 0 \leq f \leq a \tag{7.11}$$

Proof: By the law of cosines

$$\underbrace{\sqrt{(2f)^2 + 4fr(\theta)\cos\theta + r^2(\theta)}}_{dist(f_2,p)} + \underbrace{r(\theta)}_{dist(f_1,p)} = 2a \tag{7.12}$$

Isolating and squaring the root in Eq. 7.12 (and canceling common factor 4) gives

$$f^2 + fr(\theta)\cos\theta = ar(\theta) + a^2$$

which gives a linear equation for $r(\theta)$:

$$r(\theta) = \frac{a^2 - f^2}{a + f\cos(\theta)} \tag{7.13}$$

Remark 7.2 *For an hyperbola $f > a$ and $r > 0$ in the two sectors where $a + f\cos\theta < 0$.*

Exercise 7.3 *Show that the equation of the ellipse with the two focal points at $(\pm f, 0)$ in Cartesian coordinates is*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad b^2 = a^2(1 - e^2)$$

Definition 7.3 *The closest point of an orbit to the sun is called perihelion, and the furthest from the sun aphelion.*

By inspection the distance of the perihelion and aphelion from the sun is

$$r_{p/a} = \frac{a^2 - f^2}{a \pm f} = a \mp f \tag{7.14}$$

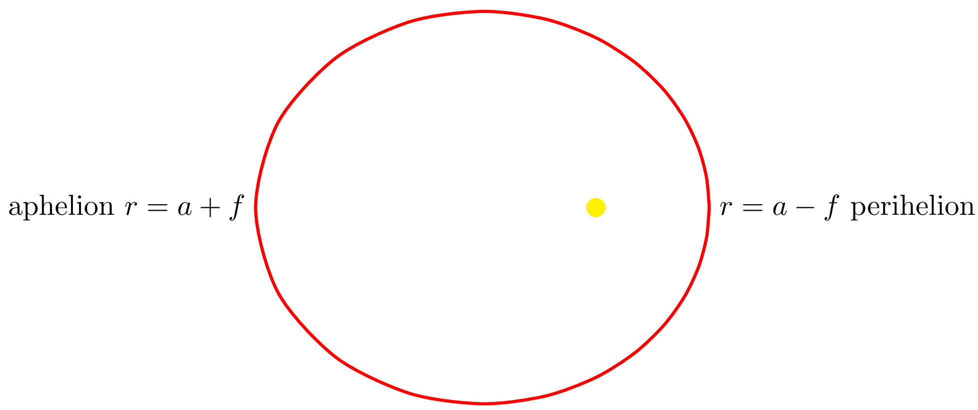


Figure 36: The perihelion and aphelion.

Example 7.2 *In the case of Mars*

$$r_p = 2 \times 10^8 [km], \quad r_a = 2.5 \times 10^8, \quad e \approx 0.1$$

The ellipticity was large enough for Kepler to conclude that the Copernicus belief that planets moved around the sun in circles was not right. An ugly fact killed a beautiful theory.

Fact 7.2 *Eq. (7.11) for $f > a$ (and $a < 0$) is the equation of hyperbolas.*

Exercise 7.4 *Show that if $e = \sqrt{2}$ the hyperbola lies in a quadrant.*

It turns out to be convenient to define $u = 1/r$ so that

$$u(\theta) = \frac{a}{a^2 - f^2} + \frac{f}{a^2 - f^2} \cos(\theta) \tag{7.15}$$

What is the differential equation that u satisfies? Differentiating Eq. 7.15 gives

$$u'' = -\frac{f}{a^2 - f^2} \cos(\theta) = -u + \frac{a}{a^2 - f^2} \tag{7.16}$$

(which is the equation of motion of a harmonic oscillator with a constant force.)

7.7 The orbits in a central potential

Euler Lagrange equations in a central potential are

$$\mu \ddot{r} = F(r) + \mu r \dot{\theta}^2 = f(r) + \frac{J^2}{\mu r^3}, \quad F(r) = -\frac{dU}{dr} \tag{7.17}$$

and we used conservation of angular momentum *in the equations of motion.*

We now want to convert the equation for the orbit as a function of time to an equation for the geometric shape of the orbit. This means, trade t for θ and turn the differential equation for the motion $r(t)$ to a differential equation for the orbit $R(\theta)$:

$$r(t) = R(\theta) = R\big(\theta(t)\big) \tag{7.18}$$

We shall now prove:

Theorem 7.6 *The orbit $R(\theta)$ satisfies the differential equation*

$$RR''(\theta) - 2(R'(\theta))^2 = \frac{\mu R^5}{J^2} F(R) + R^2 \tag{7.19}$$

Proof: Let dot and prime denote derivations wrt to t and θ. Then

$$\dot{r} = R' \dot{\theta},$$

And then also

$$\ddot{r} = R''(\dot{\theta})^2 + R'\ddot{\theta}$$

We now eliminate $\ddot{\theta}$ and $\dot{\theta}$. First

$$\dot{\theta} = \frac{J}{\mu r^2} = \frac{J}{\mu R^2},$$

Next

$$\ddot{\theta} = -2\frac{J}{\mu r^3}\dot{r} = -\frac{2J}{\mu r^3}R'\,\dot{\theta} = -\frac{2J^2}{\mu^2 R^5}\,R'$$

This then allows us to write $\ddot{r}$ in terms of R' and R'', namely

$$\ddot{r} = \dot{\theta}^2 R'' + \ddot{\theta}R' = \frac{J^2}{\mu^2 R^4}R'' - \frac{2J^2}{\mu^2 R^5}(R')^2$$

This allows us to turn Newton's equation of motion to a differential equation for the orbit $r(\theta)$

$$\frac{J^2}{\mu R^4}R''(\theta) - \frac{2J^2}{\mu R^5}(R'(\theta))^2 = F(R) + \frac{J^2}{\mu R^3}$$

Multiplying by $\mu R^5/J^2$ finishes the proof.

The trouble with this result is that we got a non-linear ugly differential equation with for $R(\theta)$. Is there a clever transformation of R that would simplify the equation? We take as a clue the fact that the differential equation for the ellipse is second order and linear if we look at $u = 1/R$. This motivates the substitution:

Fact 7.3 *For*

$$R(\theta) = \frac{1}{u(\theta)}$$

one has by direct computation

$$R' = -\frac{u'}{u^2} \to R'' = -\frac{u''}{u^2} + 2\frac{(u')^2}{u^3}$$

Now a miracle happens when we substitute this in Eq. 7.19, the non-linear terms cancel.

Theorem 7.7 *The orbit $u = 1/R$ satisfies the second order differential equation*

$$u'' = \frac{\mu}{J^2 u^2}F(1/u) - u$$

Proof:

$$RR'' - 2(R')^2 = -\frac{u''(\theta)}{u^3(\theta)} + \underbrace{2\frac{(u'(\theta))^2}{u^4(\theta)} - 2\left(\frac{u'(\theta)}{u^2(\theta)}\right)^2}_{cancel}$$

Multiplying by u^3 finishes the proof.

The differential equation for the orbit is second order, but for general central force f, non-linear. In the Kepler problem a miracle happens.

7.8 The first law: Keplerian orbits are ellipses

In the Kepler problem

$$F(R) = -\frac{k}{R^2}, \quad k = GM_\odot m > 0$$

and the differential equation for the orbit becomes linear:

Theorem 7.8 *The orbits in the Kepler problem satisfy the linear ODE*

$$u'' = K - u \tag{7.20}$$

which is the differential equation of either ellipses or hyperbolas, Eq. 7.16. The parameters of the ellipse are related to K by

$$0 < K = \frac{k\mu}{J^2} = \frac{a}{a^2 - f^2} = \frac{1}{a(1-e^2)} \tag{7.21}$$

We thus proved Kepler first law from Newton's equations.

7.9 The parameters of the ellipse and conservation laws

The orbit is an ellipse if the potential is attractive, $k > 0$, and the planet is bound to the sun, $E < 0$.

Theorem 7.9 *The ellipse parameters (a, f) are related to the constants of motions (E, J) by:*

$$\frac{1}{2a} = -\frac{E}{k}, \quad e = \frac{f}{a} = \sqrt{1 - \frac{2|E|J^2}{\mu k^2}} \tag{7.22}$$

Proof: At the perihelion, r_p is smallest and u_p is largest. There is no kinetic energy in the radial direction so, using 7.14

$$E = U_J(r_m) = -\frac{k}{r_p} + \frac{J^2}{2\mu r_p^2} = ku_p\left(-1 + \frac{J^2}{2\mu k}u_p\right)$$

This relates the energy to the major axis a

$$\frac{E}{k} = \frac{1}{a-f}\left(-1 + \frac{a^2 - f^2}{2a}\frac{1}{a-f}\right) = \frac{1}{a-f}\left(-1 + \frac{a+f}{2a}\right) = -\frac{1}{2a} \tag{7.23}$$

From Eq. 7.21, the eccentricity of the ellipse is

$$1 \geq e = \frac{f}{a} = \sqrt{1 - \frac{2|E|J^2}{\mu k^2}} = \sqrt{1 + \frac{2EJ^2}{\mu k^2}} \geq 0$$

The eccentricity is an *increasing* function of E. In particular, for fixed J the energy takes it minimal value for circular orbit

$$E \geq -\frac{\mu k^2}{2J^2}$$

7.10 Circularization of planetary orbits

The eccentricities of the orbits in the solar system are small. This could have been a coincidence. But a better explanation is: Suppose we could show that

$$\frac{d(e^2)}{dt} = \frac{2}{\mu k^2}\frac{d(EJ^2)}{dt} < 0$$

then orbits would become more circular in time and the smallness of the eccentricity would then be a consequence of the old age of the solar system.

Exercise 7.5 *Show that a friction force $-\gamma(E_k)\mathbf{p}$ dissipates energy at the rate*

$$\dot{E} = -2\gamma(E_k)E_k$$

and angular momentum at the rate

$$\dot{\mathbf{J}} = -\gamma(E_k)\mathbf{J}$$

irrespective of the nature of the central force

It follows that

$$\left(e\,\mu k^2\right)\frac{de}{dt} = -2\gamma(E_k)J^2\left(E_k + E\right)$$

If friction is small, the orbits are still to a good approximation Keplerian. Let us compute the average over one Kepler orbit, using the conservation of E and J in Kepler orbits and using the virial theorem

$$\begin{aligned}\left(e\,\mu k^2\right)\left\langle\frac{de}{dt}\right\rangle &= -2\left\langle J^2\gamma(E_k)\left(E_k + E\right)\right\rangle \\ &= -2J^2\big(\left\langle\gamma(E_k)E_k\right\rangle - \left\langle E_k\right\rangle\left\langle\gamma(E_k)\right\rangle\big)\end{aligned}$$

I need the following fact about averages[6]

Theorem 7.10 *Suppose $f(x)$ is an increasing function of x, then*

$$\left\langle xf(x)\right\rangle \geq \left\langle x\right\rangle\left\langle f(x)\right\rangle \tag{7.24}$$

Prove it. For f an increasing function $\left\langle (x - \langle x\rangle)(f(x) - f(\langle x\rangle)\right\rangle \geq 0$

It follows that for γ a *an increasing* function of its argument,

$$\left(\frac{e\,\mu k^2}{2J^2}\right)\left\langle\frac{de}{dt}\right\rangle = -\big(\left\langle\gamma(E_k)E_k\right\rangle - \left\langle E_k\right\rangle\left\langle\gamma(E_k)\right\rangle\big) \leq 0$$

Exercise 7.6 *How does the right hand side depend on e for an orbit with small eccentricity? What can you conclude from this? Try to estimate the change in e for the earth in one year.*

[6] Thanks to Oded Kenneth.

7.11 Kepler third law

Fact 7.4 (Kepler third law) *For all planets in the solar system the period and semi-major axis*

$$T^2 = const \times a^3$$

Now that we know that the orbits are ellipses, we can compute the period from the period of the radial motion (Eq. 7.10):

$$\frac{T}{2} = \int_{r_{min}}^{r_{max}} \frac{dr}{\dot{r}}$$

where

$$\frac{\mu \dot{r}^2}{2} = E + \frac{k}{r} - \frac{J^2}{2\mu r^2}$$

Theorem 7.11

$$E + \frac{k}{r} - \frac{J^2}{2\mu r^2} = -\frac{J^2}{2\mu}\left(\frac{1}{r} - \frac{1}{r_p}\right)\left(\frac{1}{r} - \frac{1}{r_a}\right)$$

Proof: Both terms are quadratic in $1/r$ and vanish at the same points $r_{p/a}$ where the radial kinetic energy vanishes. It follows that they must be proportional. The coefficients of proportionality $J^2/2m$ is determined by comparing the coefficient of $1/r^2$.

Theorem 7.12 *For all planets in the Solar system $m \approx \mu$ and Kepler third law holds with the constant given by*

$$T^2 \approx \left(\frac{4\pi^2}{GM_\odot}\right) a^3$$

Proof: For T we get the integral

$$\frac{T}{2} = \frac{\mu}{J} \int_{r_{min}}^{r_{max}} \frac{dr}{\sqrt{-\left(\frac{1}{r} - \frac{1}{r_p}\right)\left(\frac{1}{r} - \frac{1}{r_a}\right)}}$$

You can ask Wolfram alpha to compute the integral for you

$$\int_{r_p}^{r_a} \frac{dr}{\sqrt{-\left(\frac{1}{r} - \frac{1}{r_p}\right)\left(\frac{1}{r} - \frac{1}{r_a}\right)}} = \frac{\pi}{2}(r_p + r_a)\sqrt{r_p r_a} = \pi a \sqrt{a^2 - f^2}$$

and using 7.21

$$\frac{T^2}{4} = \pi^2 \frac{\mu^2}{J^2} a^2 (a^2 - f^2) = \pi^2 \frac{\mu}{k} a^3$$

In the Kepler problem $k = GM_\odot m$ with $M_\odot$ the mass of the sun and m that of the planet. Since $\mu \approx m$ in the solar system we get

$$T^2 = \underbrace{\left(\frac{4\pi^2}{GM}\right)}_{constant\ solar\ system} a^3$$

independent of the mass of the planet. So, if we can measure a we weigh the sun. This is Kepler third law.

7.12 Laplace Runge Lenz vector

The Kepler orbits are closed. This looks like a coincidence, or a miracle. It reflects a hidden symmetry and hidden constants of motion. The symmetry is hard to describe, but the constant of motion is the Laplace-Runge-Lenz vector

Theorem 7.13 *The Laplace-Runge-Lenz vector*

$$\mathbf{e} = \left(\frac{1}{\mu k}\right)\mathbf{p}\times\mathbf{J} - \frac{\mathbf{x}}{|\mathbf{x}|}$$

is a constant of motion for the Kepler problem:

$$\dot{\mathbf{e}} = 0$$

The magnitude of $\mathbf{e}$ *is the eccentricity and it points from the empty focal point to the focal point where the sun is located.*

Both are constants of motion in Kepler.

Proof: Since $\mathbf{e}\cdot\mathbf{J} = 0$ the vector $\mathbf{e}$ lies in the plane of the ellipse. It is a dimensionless quantity by inspection. Expand

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = \mathbf{b}(\mathbf{a}\cdot\mathbf{c}) - \mathbf{c}(\mathbf{a}\cdot\mathbf{b})$$

whose mnemonics is *bac-cab* to get

$$\mathbf{e} = \frac{1}{\mu k}\mathbf{p}\times(\mathbf{x}\times\mathbf{p}) - \frac{\mathbf{x}}{|\mathbf{x}|} = \frac{\mathbf{x}(\mathbf{p}\cdot\mathbf{p}) - \mathbf{p}(\mathbf{x}\cdot\mathbf{p})}{\mu k} - \frac{\mathbf{x}}{|\mathbf{x}|}$$

Suppose first that $\mathbf{e}$ is indeed a constant of motion. We can then compute it anywhere on the orbit. At the perihelion $\mathbf{p}\cdot\mathbf{x} = 0$ and hence:

$$\begin{aligned}
\mathbf{e} &= \mathbf{x}\left(\frac{\mathbf{p}\cdot\mathbf{p}}{\mu k} - \frac{1}{|\mathbf{x}|}\right) \qquad (7.25)\\
&= \mathbf{x}\left(2\left(\frac{E}{k} + \frac{1}{|\mathbf{x}|}\right) - \frac{1}{|\mathbf{x}|}\right)\\
&= \mathbf{x}\left(2\frac{E}{k} + \frac{1}{|\mathbf{x}|}\right) = \mathbf{x}\left(-\frac{1}{a} + \frac{1}{a-f}\right) = \mathbf{x}\frac{f}{a(a-f)}\\
&= \underbrace{\frac{\mathbf{x}}{a-f}}_{unit\ vector}\ \frac{f}{a} = \hat{\mathbf{x}}\,e
\end{aligned}$$

To show that $\mathbf{e}$ is indeed a constant of motion, use $\dot{\mathbf{J}} = 0$

$$\begin{aligned}\dot{\mathbf{e}} &= \frac{1}{mk}\dot{\mathbf{p}} \times \mathbf{J} - \frac{\dot{\mathbf{x}}}{|\mathbf{x}|} + \mathbf{x}\frac{\mathbf{x}\cdot\dot{\mathbf{x}}}{|\mathbf{x}|^3} \\ &= \underbrace{-\frac{\mathbf{x}}{m|\mathbf{x}|^3}}_{Newton} \times \mathbf{J} - \frac{\mathbf{v}}{|\mathbf{x}|} + \mathbf{x}\frac{\mathbf{x}\cdot\mathbf{v}}{|\mathbf{x}|^3} \\ &= \frac{-\mathbf{x}\times(\mathbf{x}\times\mathbf{v}) - \mathbf{v}(\mathbf{x}\cdot\mathbf{x}) + \mathbf{x}(\mathbf{x}\cdot\mathbf{v})}{|\mathbf{x}|^3} \\ &= 0\end{aligned} \tag{7.26}$$

The numerator vanishes by *bac-cab*.

Now that you know that $\mathbf{e}$ is conserved painful results follow easily. For example, that the orbit is an ellipse follows from:

$$\underbrace{\mathbf{e}\cdot\mathbf{x}}_{re\cos\theta} = \frac{1}{mk}\mathbf{p}\times\mathbf{J}\cdot\mathbf{x} - |\mathbf{x}| = \frac{1}{mk}\mathbf{x}\times\mathbf{p}\cdot\mathbf{J} - |\mathbf{x}| = \frac{\mathbf{J}^2}{mk} - \underbrace{|\mathbf{x}|}_{-r}$$

from which we recover the equation of the ellipse

$$\left(\frac{J^2}{mk}\right) = r(1 + e\cos\theta)$$

8 Small Oscillations

In this section we shall consider Lagrangians that are quadratic in q and $\dot{q}$. There are two reasons to study them. The physical reason is that these are the universal description of a motion near equilibrium and so ubiquitous. The second is that they lead to Euler-Lagrange equations that are linear ODE. There is an arsenal to study such equations.

8.1 Simple Harmonic oscillator

The simplest quadratic Lagrangian is the Harmonic oscillator:

$$L = \frac{m}{2}\dot{q}^2 - \frac{k}{2}q^2 = \frac{m}{2}(\dot{q}^2 - \omega^2 q^2), \quad \omega^2 = \frac{k}{m} > 0$$

By dimensional analysis $[\omega] = [1/sec]$. Since a multiple of the Lagrangian does not affect the Euler Lagrange equations, only ω shows up in the equations. Indeed

$$\ddot{q} = -\omega^2 q, \quad \omega^2 > 0$$

This is a homogeneous, linear, second order differential equation, with (real) constant coefficients.

- Since the equation is homogeneous $q = 0$ is a solution. This is the equilibrium point.

- Since it is both homogeneous and linear if $q(t)$ is a solution so is $\lambda q(t)$. This gives a remarkable feature of harmonic motion, namely that the period is independent of the amplitude.

- Since the equation is real if $q(t)$ is a solution so is $q^*(t)$ equivalently, if $q(t)$ is a complex solution so are $Re\big(q(t)\big)$ and $Im\big(q(t)\big)$.

- The period is
$$T = \frac{2\pi}{\omega}$$

The standard method to solve ODE with constant coefficients is to make the ansatz

$$q = e^{i\omega t}$$

The general (real) solution can be written in many equivalent forms

$$q(t) = Re\big(Ae^{i\omega t}\big) = |A|Re\big(e^{i(\omega t+\alpha)}\big) = |A|\cos(\omega t + \alpha), \quad A = |A|e^{i\alpha} \in \mathbb{C}$$

Exercise 8.1 *Assuming that the leg is a harmonic pendulum, estimate the walking speed of a man.*

Exercise 8.2 *Suppose*

$$L = \frac{m}{2}\dot{q}^2 - \frac{k}{2}q^2 + g\, q\dot{q}$$

Find the equations of motions and determine the allowed values for g so that the motion is stable.

8.2 The general form of the Lagrangian of small oscillations

In this section I want to derive the most general form of the Lagrangian to second order obtained as Taylor expansion of a general Lagrangian.

Assumption 8.1 *We assume that the Lagrangian is at least twice differentiable at $q = \dot{q} = 0$ and the motion that stay near $q \approx 0$ and $\dot{q} \approx 0$.*

Expand the Lagrangian about the origin to second order:

$$L(q,\dot{q},t) = \underbrace{L(0,0,t)}_{0\ order} + \underbrace{\frac{\partial L}{\partial q}q + \left(\frac{\partial L}{\partial \dot{q}}\right)\dot{q}}_{1-st\ order} + \underbrace{\tfrac{1}{2}\left(\frac{\partial^2 L}{\partial q^2}q^2 + 2\frac{\partial^2 L}{\partial q \partial \dot{q}}q\dot{q} + \frac{\partial^2 L}{\partial \dot{q}^2}(\dot{q})^2\right)}_{2-nd\ order} + \dots \quad (8.1)$$

where all the derivatives are computed at $q = \dot{q} = 0$ and are therefore, either constants, or functions of time. For example

$$\frac{\partial L}{\partial q} = \frac{\partial L}{\partial q}(0,0,t), \quad \frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial \dot{q}}(0,0,t),$$

etc. Clearly, the 0 order term, being a complete derivative

$$L(0,0,t) = \frac{d}{dt}\left(\int^t L(0,0,s)ds\right)$$

does not affect Euler-Lagrange equations. We can set it to 0 without loss.

The first order term can be organized as

$$\frac{\partial L}{\partial q}q + \left(\frac{\partial L}{\partial \dot{q}}\right)\dot{q} = \underbrace{\frac{d}{dt}\left(\left(\frac{\partial L}{\partial \dot{q}}\right)q\right)}_{complete\ derivative} + \underbrace{\left(-\frac{\partial^2 L}{\partial \dot{q}\partial t} + \frac{\partial L}{\partial q}\right)}_{F(t)} q \tag{8.2}$$

The second order term can be organized as follows

$$\frac{m_{\alpha\beta}}{2}\dot{q}_\alpha\dot{q}_\beta - \frac{k_{\alpha\beta}}{2}q_\alpha q_\beta + \omega_{\alpha\beta}q_\alpha\dot{q}_\beta \tag{8.3}$$

summation over repeated indexes implied. It is clear that we can assume, without loss

$$m_{\alpha\beta} = m_{\beta\alpha}, \quad k_{\alpha\beta} = k_{\beta\alpha} \tag{8.4}$$

for the antisymmetric part drops upon summation. For the thirds term note that

$$\begin{aligned}\frac{d}{dt}\left(\omega_{\alpha\beta}q_\alpha q_\beta\right) &= \omega_{\alpha\beta}\dot{q}_\alpha q_\beta + \omega_{\alpha\beta}q_\alpha\dot{q}_\beta + \dot{\omega}_{\alpha\beta}q_\alpha q_\beta \\ &= \left(\omega_{\alpha\beta} + \omega_{\beta\alpha}\right)\dot{q}_\alpha q_\beta + \dot{\omega}_{\alpha\beta}q_\alpha q_\beta\end{aligned} \tag{8.5}$$

Let me write

$$\omega_{\alpha\beta} = \omega^s_{\alpha\beta} + \omega^a_{\alpha\beta} = \frac{1}{2}(\omega_{\alpha\beta} + \omega_{\beta\alpha}) + \frac{1}{2}(\omega_{\alpha\beta} - \omega_{\beta\alpha})$$

Then, from Eq. 8.5 we have that

$$\frac{d}{dt}\left(\omega^s_{\alpha\beta}q_\alpha q_\beta\right) = 2\omega^s_{\alpha\beta}\dot{q}_\alpha q_\beta + \dot{\omega}^s_{\alpha\beta}q_\alpha q_\beta \tag{8.6}$$

We can then get rid of the symmetric part of $\omega_{\alpha\beta}$ from the second order term and are then left with

$$\frac{m_{\alpha\beta}}{2}\dot{q}_\alpha\dot{q}_\beta - \frac{k_{\alpha\beta} + \dot{\omega}^s_{\alpha\beta}}{2}q_\alpha q_\beta + \omega^a_{\alpha\beta}q_\alpha\dot{q}_\beta \tag{8.7}$$

We have therefore shown that to second order, the Lagrangian is of the form

Theorem 8.1 *The general form of the Lagrangian to second oder is:*

$$L = \frac{1}{2}\dot{\mathbf{q}}^t\mathbf{M}\dot{\mathbf{q}} - \frac{1}{2}\mathbf{q}^t\mathbf{K}\mathbf{q} + \mathbf{q}^t\boldsymbol{\Omega}\dot{\mathbf{q}} + \mathbf{F}\cdot\mathbf{q} \tag{8.8}$$

where $\mathbf{q} = (q_1,\ldots,q_N)^t$, $\mathbf{M}$ *and* $\mathbf{K}$ *are symmetric matrices and* $\boldsymbol{\Omega}$ *anti-symmetric and* $\mathbf{F} = (F_1,\ldots,F_N)$ *(all possibly time dependent). If* L *is time reversal symmetric then* $\boldsymbol{\Omega} = 0$.

This Lagrangian leads to linear Euler-Lagrange equations.

Theorem 8.2 *The canonical momenta are*

$$\mathbf{p} = \mathbf{M}\dot{\mathbf{q}} - \mathbf{\Omega}\mathbf{q} \tag{8.9}$$

The Euler-Lagrange equations are

$$\dot{\mathbf{p}} = \frac{\partial L}{\partial \mathbf{q}} = -\mathbf{K}\mathbf{q} + \mathbf{\Omega}\dot{\mathbf{q}} + \mathbf{F} \tag{8.10}$$

They may be written explicitly as

$$\mathbf{M}\ddot{\mathbf{q}} = -(\mathbf{K} - \dot{\mathbf{\Omega}})\mathbf{q} + (2\mathbf{\Omega} - \dot{\mathbf{M}})\dot{\mathbf{q}} + \mathbf{F} \tag{8.11}$$

Finally, the energy is

$$E = \frac{1}{2}\dot{\mathbf{q}}^t\mathbf{M}\dot{\mathbf{q}} + \frac{1}{2}\mathbf{q}^t\mathbf{K}\mathbf{q} - \mathbf{F}\cdot\mathbf{q} \tag{8.12}$$

Since the Euler-Lagrange equations are a linear system of ODE we have that

Theorem 8.3 *The general solution can be written as a sum of any special solution and the general solution of the homogeneous system, i.e. the system with* $\mathbf{F} = 0$*.*

Theorem 8.4 *In the special case that* $\mathbf{F}$ *and* $\mathbf{K} - \dot{\mathbf{\Omega}}$ *are time independent, a special, time independent, solution is*

$$\mathbf{q}_0 = (\mathbf{K} - \dot{\mathbf{\Omega}})^{-1}\mathbf{F}$$

(assuming that $\mathbf{K} - \dot{\mathbf{\Omega}}$ *is invertible.)*

This is the case where $\mathbf{F}$ displaces the equilibrium position to a new one.

8.3 Normal modes

Since the solutions of the homogeneous system play a central tole, lets study them in detail. Consider a system of N linear and homogeneous ODE written in matrix notation

$$\mathbf{M}\ddot{\mathbf{q}} = -\mathbf{K}\mathbf{q} + 2\mathbf{\Omega}\dot{\mathbf{q}} \tag{8.13}$$

($\mathbf{M}, \mathbf{K}, \mathbf{\Omega}$ may be time dependent). A basic fact in ODE is

Theorem 8.5 *The space of solutions of a system of* N *linear and homogeneous ODE is a* $2N$ *dimensional vector space: Solutions can be added and multiplied by a scalar. We can choose as a basis vectors the solutions with initial conditions*

$$\begin{aligned}
\mathbf{q}_1 = (1, 0, \ldots, 0), &\quad \dot{\mathbf{q}}_1 = 0 \\
&\ldots \\
\mathbf{q}_N = (0, 0, \ldots, 1), &\quad \dot{\mathbf{q}}_N = 0 \\
\mathbf{q}_{N+1} = 0, &\quad \dot{\mathbf{q}}_{N+1} = (1, 0 \ldots, 0) \\
&\ldots \\
\mathbf{q}_{2N} = 0, &\quad \dot{\mathbf{q}}_{2N} = (0, \ldots, 1)
\end{aligned}$$

In the case that $\mathbf{M}, \mathbf{K}, \mathbf{\Omega}$ are constant matrices we can say more and describe the time dependence explicitly:

Theorem 8.6 *When $\mathbf{M}, \mathbf{K}, \mathbf{\Omega}$ are time-independent, the basis for the vector space is given by harmonic functions in time, i.e.*

$$\mathbf{q}_\alpha e^{i\,\omega_\alpha t}, \quad \alpha = 1, \ldots, N$$

where $\mathbf{q}_\alpha$ satisfies

$$(\omega_\alpha^2 \mathbf{M} + 2i\omega_\alpha \mathbf{\Omega} - \mathbf{K})\mathbf{q}_\alpha = 0$$

The equation fixes $\mathbf{q}_\alpha$ up to overall normalization which we can set arbitrarily. For a non-trivial solution $\mathbf{q}_\alpha$ to exist ω_α must be a solution of

$$\det(\omega_\alpha^2 \mathbf{M} + 2i\omega_\alpha \mathbf{\Omega} - \mathbf{K}) = 0$$

These are the normal modes of the system. Since $\mathbf{M}, \mathbf{K}$ and $\mathbf{\Omega}$ are real matrices, if $\mathbf{q}_\alpha e^{i\omega t}$ is a normal mode, so is its complex conjugate $\mathbf{q}_\alpha^* e^{-i\omega_\alpha^* t}$. By energy conservation all the ω_α must be real (otherwise the energy will blow up at $\pm\infty$).

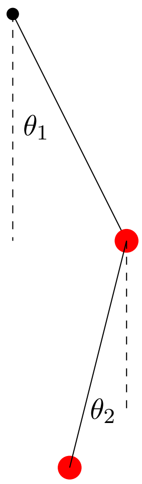


Figure 37: Double pendulum

8.4 Double pendulum

Two identical masses, arms of identical lengths. The kinetic energy for small oscillations is

$$T(\dot\theta_1\,\dot\theta_2) = \frac{m\ell^2}{2}\left(\dot\theta_1^2 + (\dot\theta_1 + \dot\theta_2)^2\right) = \frac{m\ell^2}{2}\left(2\dot\theta_1^2 + 2\dot\theta_1\dot\theta_2 + \dot\theta_2^2\right)$$

Exercise 8.3 *Show this.*

The matrix M is then

$$M = \frac{m\ell^2}{2}\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

The potential energy is

$$U(\theta_1, \theta_2) = -mg\ell(2\cos\theta_1 + \cos\theta_2) \approx -mg\ell\left(3 - \theta_1^2 - \frac{1}{2}\theta_2^2\right)$$

So the matrix K is

$$K = \frac{mg\ell}{2}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

The normal modes are determined by the matrix

$$-\omega^2 M + K = \frac{m\ell}{2}\begin{pmatrix} -2\omega^2\ell + 2g & -\omega^2\ell \\ -\omega^2\ell & -\omega^2\ell + g \end{pmatrix}$$

The normal modes have frequencies

$$\omega^2 = \frac{g}{\ell}(2 \pm \sqrt{2})$$

and the corresponding modes are

$$\left(\mp\frac{2+\sqrt{2}}{2\left(1+\sqrt{2}\right)}, 1\right) \approx (\mp .7, 1)$$

The lower frequency is when the co-move. The higher frequency is when they counter move.

8.5 Linear chain with periodic boundary conditions

Consider n identical equidistant particles on the circumference of a ring, connected by identical springs (that apply forces tangent to the ring). We identify the n+1-th particle with the 1-th particle, as if this was a circle. We call x_j the coordinate of the j-th particle and take the Lagrangian

$$L = \frac{m}{2}\sum_{j=1}^{n}\dot{x}_j^2 - \frac{k}{2}\sum_{j=1}^{n}(x_j - x_{j+1})^2, \quad x_{n+1} = x_1$$

The mass matrix is simple: $M = m\mathbb{1}$. The matrix K is tridiagonal

$$K = k\begin{pmatrix} 2 & -1 & 0 & \dots & -1 \\ -1 & 2 & -1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 2 & -1 \\ -1 & 0 & \dots & -1 & 2 \end{pmatrix}$$

To find the modes we need the eigenvalues and eigenvectors of K. Let

$$z_n = e^{2\pi i/n}$$

denote the n-th root of unity. I claim that for any j (counted $\mod n$), the vector

$$(z_n^j, z_n^{2j}, \dots, z_n^{nj} = 1)$$

is an eigenvector of K with eigenvalue

$$k(2 - z_n^j - z_n^{-j}) = k(2 - 2\cos 2\pi j/n)$$

It follows that the n frequencies of the normal modes are

$$\omega_j^2 = \frac{2k}{m}\left(1 - \cos\frac{2\pi j}{n}\right), \quad j = 0, 1, \dots, n-1$$

These modes are called phonons. In this case there is one zero mode, when $j = n$ with $\omega = 0$. This corresponds to the rigid rotation of the ring. A mode that has no restoring force. The normal modes are plane waves in the limit that the ring is infinite. Take

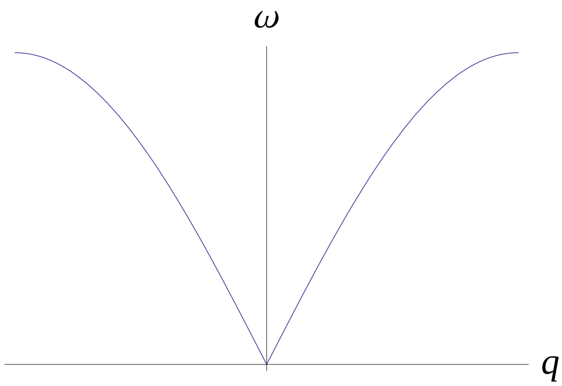


Figure 38: When n is large, the phonon dispersion approaches a continuous curve, $\omega(q)$ with $q = 2\pi j/n$ plotted for $|q|a \leq \pi$. qa is counted mod 2π and can be viewed as a torus (an angle). The circle qa is known as Brillouin zone.

$n \to \infty$ with average distance a between atoms. When n is large, and a small, we can replace the discrete mode variable j with a continuous variable q defined by $qa = 2\pi j/n$. q is a property of the mode, called wave number, and conventionally has dimensions $[cm{-}1]$. Since j is counted $\mod n$, we see that qa is counted $\mod 2\pi$. Conventionally one chooses $|qa| \leq \pi$. Small frequencies correspond to $q \to 0$ where

$$\omega(qa) \approx \sqrt{\frac{k}{m}}|qa|$$

The dispersion is linear.

The normal mode is now a huge vector with amplitudes (components):

$$(z_n^j, z_n^{2j}, \ldots, 1) = (e^{iqa}, \ldots, e^{iqam}, \ldots, 1)$$

The m-th position of the vector describes the amplitude of the $m-th$ mass whose which is oscillating near $x \leftrightarrow ma$ where a is the average distance between the atoms. In the limit $n \to \infty$ the entries of this long vector can be viewed as the entries of a function of a continuous coordinate, x. The m-th entry in this (long) vector, has time dependence

$$e^{iqam} e^{-i\omega(qa)t} = e^{i(qx-\omega(qa)t)}$$

This describes a plane wave propagating at (group) velocity

$$\frac{d\omega}{dq} = a\sqrt{\frac{k}{m}} \cos\left(\frac{qa}{2}\right)$$

8.6 Friction

An oscillator with friction is modeled by

$$\ddot{q} = -\omega^2 q - \gamma \dot{q}, \quad \gamma > 0 \tag{8.14}$$

Now there are two frequency scales: ω and γ.

So far I have avoided talking about friction. Can one put friction in Lagrangians? Since friction dissipates energy, the Lagrangian must be time dependent. Let us try

$$L = g(t)\big(T(\dot{q}) - U(q)\big)$$

The momentum and Euler-Lagrange are

$$p = g(t)\frac{\partial T}{\partial \dot{q}}$$

Euler-Lagrange are

$$\dot{p} = \dot{g}\frac{\partial T}{\partial \dot{q}} + g(t)\frac{d}{dt}\frac{\partial T}{\partial \dot{q}} = -g(t)\frac{\partial U}{\partial q} \Longrightarrow \frac{d}{dt}\frac{\partial T}{\partial \dot{q}} = -\frac{\partial U}{\partial q} - \frac{d\log g}{dt}\frac{\partial T}{\partial \dot{q}}$$

In particular, for kinetic energy $T = \frac{m}{2}\dot{q}^2$ we get

$$m\ddot{q} = -\frac{\partial U}{\partial q} - \frac{d\log g}{dt} m\dot{q}$$

We get friction if we choose

$$g(t) = e^{\gamma t}$$

Here is a cautionary remark. The energy, as we defined it, and assuming T homogeneous of degree n

$$E = \underbrace{p\dot{q} - L}_{Legendre\ of\ L} = g(t)\left(\frac{\partial T}{\partial \dot{q}}\dot{q} - T(\dot{q}) + U(q)\right) = g(t)\underbrace{\big((n-1)T(\dot{q}) + U(q)\big)}_{kinetic\ +\ potential\ energy}$$

The standard T has $n = 2$. So the notion of kinetic + potential energy and the notion of energy defined by the Legendre transform of the Lagrangian do not agree in this case. The standard rules do not apply here. One needs to be careful.

Solving the equation of motion Eq. 8.14

The ansatz $q = e^{\lambda t}$ gives the quadratic equation for λ

$$\lambda^2 + \gamma\lambda + \omega^2 = 0 \Longrightarrow \lambda = \underbrace{-\frac{\gamma}{2}}_{decay} \pm i\sqrt{\omega^2 - (\gamma/2)^2} = -\frac{\gamma}{2} \pm i\tilde{\omega}$$

If $\omega^2 > (\gamma/2)^2$ there is an oscillatory piece with shifted (lower) frequency $\tilde{\omega}$. The oscillator is weakly damped. If $\omega^2 < (\gamma/2)^2$ the $\tilde{\omega}$ is imaginary, there is no oscillatory part and the oscillator is over-damped.

Assuming under-damping the general solution is

$$q(t) = e^{-\gamma t/2} Re\big(Ae^{i\tilde{\omega}t}\big)$$

8.7 Dirac delta

Define $\delta(t)$ so that for any smooth function $f(t)$

$$\int_{-\infty}^{\infty} \delta(t) f(t) dt = f(0)$$

This implies that $\delta(t) = 0$ for all $t \neq 0$. For example

$$\delta(t) = \lim_{\varepsilon \to 0} \begin{cases} \frac{1}{\varepsilon}, & |t| < \varepsilon/2 \\ 0 & otherwise \end{cases}$$

It follows that any nice function $f(t)$ can be written as

$$f(t) = \int f(s)\delta(t-s)ds$$

It follows from the linearity of the equations of motions that a special solution for the driving force $f(t)$ can be re-constructed by superposing the special solutions to impulses $f(s)\delta(t-s)$. Now for an impulse the equations of motion are the homogeneous equations except for one time $t = s$. Integrating Eq. (8.9) over the impulse time says that p_α makes a jump at the time of the impulse whose size is:

$$\Delta p_\alpha(s) = f_\alpha(s) \tag{8.15}$$

It follows that if we have the general solution of the homogenous equations, we can also construct the solution for an impulse, and by superposition, a special solution for any $f(t)$, (see section 8.9 for an example).

8.8 Driven oscillator

The Lagrangian is

$$L = \frac{m}{2}\dot{q}^2 - \frac{k}{2}q^2 + F(t)q$$

and the Euler Lagrange equations are

$$\ddot{q} = -\omega^2 q + \frac{1}{m}F(t)$$

Since the equation is linear if q_0 is a solution, so is $q_0 + A\cos(\omega t + \alpha)$: We are free to add solutions of the homogeneous (not driven) equation. It also follows from linearity that if, q_1 and q_2 are solutions of

$$m\ddot{q}_1 = -kq_1 + F_1(t), \quad m\ddot{q}_2 = -kq_2 + F_2(t)$$

then $q_{12} = q_1 + q_2$ is a solution of

$$m\ddot{q}_{12} = -kq_{12} + F_1(t) + F_2(t)$$

We want to use this observation to solve the equation for general F be decomposing F as a superposition of elementary impulses.

8.9 Impulse

Any nice function $F(t)$ can be written as

$$F(t) = \int F(s)\delta(t-s)ds$$

We can find a special solution for general F if we know how to solve

$$\ddot{q} = -\omega^2 q + \gamma\delta(t)$$

γ is a parameter with dimensions $[q/t]$. (Note that $[\delta(t)] = [1/t]$). We know how to solve the equation for $t > 0$ and also for $t < 0$. What we need is to stitch the solutions. To see how this is done integrate the equation from $0-$ to $0+$

$$\dot{q}(0_+) - \dot{q}(0_-) = -\omega^2 \underbrace{\int_{0-}^{0+} dt\, q(t)}_{=0} + \gamma \underbrace{\int_{0-}^{0+} \delta(t)dt}_{=1} = \gamma \tag{8.16}$$

The pulse changes the momentum abruptly.

We call Green function the solution that is associated with the oscillator initially, before the impulse, at time $0-$, is at rest: $q(0-) = \dot{q}(0-) = 0$. From Eq. 8.16 we see that for $t > 0$ we look for solution with initial conditions at $t = 0+$

$$q(0+) = 0, \quad \dot{q}(+0) = \gamma$$

Hence

$$G(t) = \begin{cases} 0 & t < 0 \\ \gamma \frac{\sin \omega t}{\omega} & t > 0 \end{cases}$$

Hence a special solution to the driving $F(t)$ is a superposition of pulses:

$$q(t) = \int \frac{F(s)}{m\gamma} G(t-s) ds = \frac{1}{m\omega} \int_{-\infty}^{t} F(s) \sin \omega(t-s) ds$$

and the general solution is

$$q(t) = \underbrace{A\cos(\omega t + \alpha)}_{solution\ of\ homogeneous} + \frac{1}{m\omega} \int_{-\infty}^{t} F(s) \sin \omega(t-s) ds \tag{8.17}$$

Suppose for example that the pendulum is at rest and at time $t = 0$ I apply $F(t) = F \sin \nu t$. For $\omega \neq \nu$ one finds an oscillatory motion at the natural frequency of the oscillator ω and the frequency of the driving ν

$$q(t) = \frac{F}{m\omega} \frac{\omega \sin(t\nu) - \nu \sin(t\omega)}{\omega^2 - \nu^2}$$

For $\omega = \nu$ one finds instead an oscillating piece and a piece whose amplitude grows linearly

$$\frac{F}{2m\omega^2} \left(\sin(t\omega) - \underbrace{t\omega \cos(t\omega)}_{resonance} \right)$$

Exercise 8.4 *Derive the analog of Eq. 8.17 for the damped oscillator. What happens if you drive a damped oscillator at resonance?*

Tacoma narrows bridge

9 Hamiltonian Mechanics

9.1 Legendre transform: Geometry

In mechanics, Legendre transform relates the Lagrangian to the Hamiltonian. In thermodynamics and statistical mechanics it relate various notions of entropies and free energies where convexity plays a key role.

We recall first the notion of convex functions:

Definition 9.1 *A function $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ is convex if for all $\mathbf{x}, \mathbf{y}$ and $0 \le \lambda \le 1$*

$$f\big(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}\big) \le \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

see figure 39. In other words if the function of the average is below the average of the function.

Theorem 9.1 *A convex function is continuous and admits right and left derivatives.*

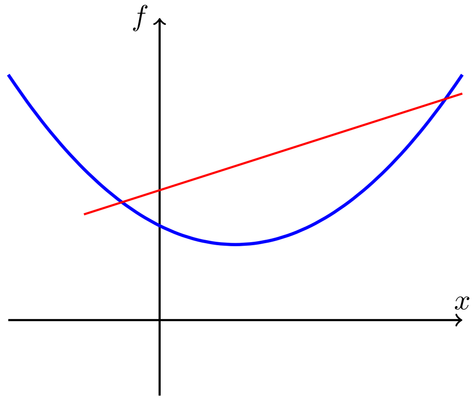


Figure 39: A convex function

The analytic definition of the Legendre transform is:

Definition 9.2 *The Legendre transform of a function $f : \mathbb{R}^n \to \mathbb{R}$ denoted by $\mathcal{L}f$ is also a function: $\mathcal{L}f \ : \mathbb{R}^n \to \mathbb{R}$, defined by*

$$(\mathcal{L}f)(\mathbf{p}) = \sup_{\mathbf{v}} \left(\mathbf{p} \cdot \mathbf{v} - f(\mathbf{v})\right)$$

Example 9.1 *Let $\mathbf{M}$ be a positive matrix and L the kinetic energy*[7]

$$L(\mathbf{v}) = \frac{1}{2} \mathbf{v}^t \cdot \mathbf{M} \mathbf{v}, \quad \mathbf{M} > 0$$

[7] A (real) matrix $\mathbf{M}$ is positive if $\mathbf{v}^t \cdot \mathbf{M}\mathbf{v} > 0$ for every (real) vector $\mathbf{v}$.

The supremum is found by looking for the stationary point

$$0 = \mathbf{p} - \nabla_v L = \mathbf{p} - \mathbf{M}\mathbf{v}$$

It can be simply inverted to

$$\mathbf{v} = \mathbf{M}^{-1}\mathbf{p}$$

Substituting we find the Legendre transform to be a function of $\mathbf{p}$*:*

$$(\mathcal{L}L)(\mathbf{p}) = \frac{1}{2}\mathbf{p}^t \cdot \mathbf{M}^{-1}\mathbf{p}$$

The Legendre transform of convex function has a nice geometric meaning given in Fig. 40 and it works even if the function has different right and left derivatives, as in the figure. Such singularities are important in thermodynamics where they characterize phase transitions. Their role in mechanics, if any, is not so clear. However, the Legendre

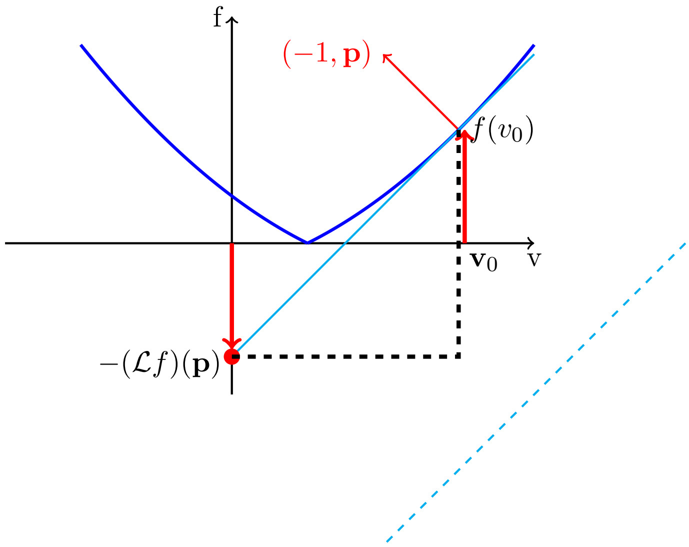


Figure 40: Legendre transform: Take a plane (line in the figure) with normal $(-1, \mathbf{p})$ as shown in the figure (dotted cyan line). Push the plane so it is tangent to the curve $f(\mathbf{v})$ (cyan line). The point of contact is $\mathbf{v}_0$. The plane intercepts the vertical axis at $-\mathcal{L}f$, the red dot. This gives the value of the Legendre transform at $\mathbf{p}$. To see that this agree with the formula note that the hight of the dashed triangle is pv_0. The coordinate on the f axis of the red dot is then $-pv_0 + f(v_0) = -\mathcal{L}f$. The procedure needs f to be convex but not necessarily smooth, as in the figure.

transform is well defined even if the function f is not convex. The geometric picture is given in Fig. 41

Theorem 9.2 *$\mathcal{L}f$ is a convex function.*

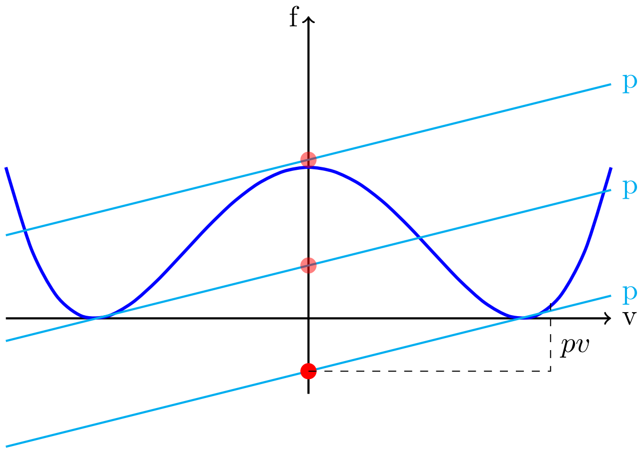


Figure 41: Legendre transform generates a function of the slope p by the following prescription: Draw tangents to the curve with slope p. The intercept with the f axis gives the value of $-\mathcal{L}f$. These are the red dots. The supremum picks the bottom one. If the function is convex, there is a single tangent.

Proof:

$$
\begin{aligned}
(\mathcal{L}f)\big(\lambda\mathbf{p} + (1-\lambda)\mathbf{p}'\big) &= \sup_{\mathbf{v}} \big(\big(\lambda\mathbf{p} + (1-\lambda)\mathbf{p}'\big)\mathbf{v} - f(\mathbf{v})\big) \\
&= \sup_{\mathbf{v}} \big(\lambda\big(\mathbf{p}\cdot\mathbf{v} - f(\mathbf{v})\big) + (1-\lambda)\big(\mathbf{p}'\cdot\mathbf{v} - f(\mathbf{v})\big)\big) \\
&\le \sup_{\mathbf{v}} \big(\lambda\big(\mathbf{p}\cdot\mathbf{v} - f(\mathbf{v})\big)\big) + \sup_{\mathbf{v}} \big((1-\lambda)\big(\mathbf{p}'\cdot\mathbf{v} - f(\mathbf{v})\big)\big) \\
&= \lambda(\mathcal{L}f)(\mathbf{p}) + (1-\lambda)(\mathcal{L}f)(\mathbf{p}')
\end{aligned}
\tag{9.1}
$$

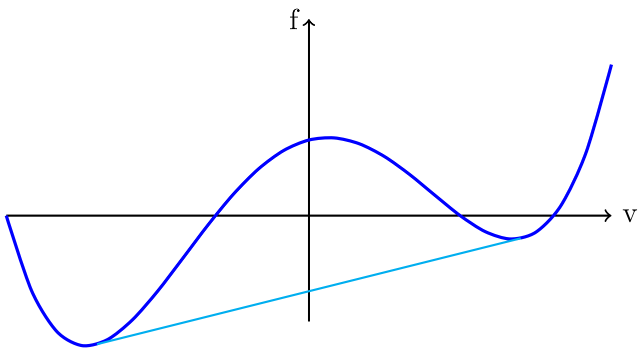


Figure 42: To find the Legendre transform of a function which is not convex, you may first replace the function by its convex envelope and then compute the Legendre transform of the envelope.

9.2 Legendre transform: Analysis

The Legendre transform makes sense for any function, if you allow for $\mathcal{L}f$ to take the value $+\infty$. For example, the trivial function $f(v) = 0$ has as Legendre transform the "function"

$$\sup_x xp = \begin{cases} 0 & p = 0 \\ \infty & p \neq 0 \end{cases}$$

to avoid these pathological examples let us restrict attention to functions f so that the Hessian is non-singular

$$\det \mathbf{J} \neq 0, \quad J_{jk} = \left(\frac{\partial^2 f}{\partial v_j \partial v_k}\right)$$

(except perhaps for some low dimensional sets).

The stationary points are candidates for the supremum:

$$0 = \nabla_{\mathbf{v}} \left(\mathbf{p} \cdot \mathbf{v} - f(\mathbf{v})\right) = \mathbf{p} - \nabla_{\mathbf{v}} f(\mathbf{v})$$

This gives n equations for the n unknown components of $\mathbf{v} = (v_1, \ldots, v_n)$. Suppose that for $\mathbf{p}_0 \in \mathbb{R}^n$ there is a point $\mathbf{v}_0$ so that

$$\nabla_{\mathbf{v}} f(\mathbf{v}_0) = \mathbf{p}_0 \tag{9.2}$$

In the neighborhood of this point we have

$$\nabla_{\mathbf{v}} f(\mathbf{v}_0 + \delta\mathbf{v}) = \mathbf{p}_0 + \delta\mathbf{p}$$

Taylor expansion then tells us how $\mathbf{p}$ changes with $\mathbf{v}$:

$$\mathbf{J}\delta\mathbf{v} = \delta\mathbf{p}, \quad J_{jk} = \left(\frac{\partial^2 f}{\partial v_j \partial v_k}\right)(\mathbf{v}_0)$$

If $\det \mathbf{J} \neq 0$ you can invert the relation to get $\mathbf{p}$ as a function of $\mathbf{v}$:

$$\mathbf{J}^{-1}\delta\mathbf{p} = \delta\mathbf{v}$$

This is the baby version of the implicit function theorem which guarantees that in the neighborhood of $\mathbf{v}_0$ there is a unique function $\mathbf{v}_0(\mathbf{p})$. You may then feed in the solution in

$$(\mathcal{L}f)(\mathbf{p}) = \mathbf{p} \cdot \mathbf{v}_0(\mathbf{p}) - f(\mathbf{v}_0(\mathbf{p}))$$

to get a function of $\mathbf{p}$ alone. If Eq. 9.2 has several solutions, you pick the largest.

Theorem 9.3 *Legendre is a duality between the coordinate* $\mathbf{v}$ *and* $\mathbf{p}$ *in the sense that*

$$df = \mathbf{p} \cdot d\mathbf{v}$$

while

$$d(\mathcal{L}f)(\mathbf{p}) = \mathbf{v} \cdot d\mathbf{p}$$

Proof:
$$d(\mathcal{L}f)(\mathbf{p}) = d(\mathbf{p}\cdot\mathbf{v} - f) = \mathbf{v}\cdot d\mathbf{p} + \mathbf{p}\cdot d\mathbf{v} - df = \mathbf{v}\cdot d\mathbf{p}$$

Theorem 9.4 *If f is a convex function then*
$$\mathcal{L}^2 f = f$$

Proof: This follows from the duality, Theorem 9.3.

Exercise 9.1 *Suppose $f(\mathbf{v})$ is homogeneous of degree β, i.e.*
$$f(\lambda\mathbf{v}) = \lambda^\beta f(\mathbf{v}), \quad \lambda > 0$$
Show that its Legendre transform is a homogeneous of degree α where
$$(\mathcal{L}f)(\lambda\mathbf{p}) = \lambda^\alpha(\mathcal{L}f)(\mathbf{p}), \quad \frac{1}{\beta} + \frac{1}{\alpha} = 1$$

Exercise 9.2 *Let $T_\mathbf{w}$ be a shift, i.e.*
$$(T_\mathbf{w} f)(\mathbf{v}) = f(\mathbf{v} - \mathbf{w})$$
Show that
$$(\mathcal{L}T_\mathbf{w} f)(\mathbf{p}) = \mathbf{p}\cdot\mathbf{w} + (\mathcal{L}f)(\mathbf{p})$$

Exercise 9.3 *What is the Legendre transform of $|\mathbf{v}|$?*

Answer:
$$(\mathcal{L}|\mathbf{v}|)(\mathbf{p}) = \begin{cases} 0 & |\mathbf{p}| \le 1 \\ \infty & |\mathbf{p}| > 1 \end{cases}$$
(This result is most easily seen from geometric construction.)

9.3 The Hamiltonian

In Lagrangian mechanics the players are the coordinates and the velocities $(\mathbf{q}, \mathbf{v})$ and the Lagrangian is a function $L(\mathbf{q}, \mathbf{v}, t)$. The momenta are then derived quantities
$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}}$$
or, the slope of the Lagrangian in $\mathbf{v}$. This leads to

Definition 9.3 *The Hamiltonian $H(\mathbf{q}, \mathbf{p}, t)$ is the Legendre transform of the Lagrangian $L(\mathbf{q}, \mathbf{v}, t)$. For L convex in $\mathbf{v}$ it is given by*
$$H(\mathbf{q}, \mathbf{p}, t) = (\mathcal{L}L)(\mathbf{q}, \mathbf{p}, t) = \mathbf{v}\cdot\mathbf{p} - L(\mathbf{q}, \mathbf{v}, t)$$

In the time independent case the Hamiltonian coincides with the energy as defined by Nöther theorem (=Jacobi integral).

9.4 Hamilton equations

The definition of the momentum and Euler-Lagrange equations are encoded in the differential

$$\begin{aligned} dL &= \nabla_{\mathbf{q}} L \cdot d\mathbf{q} + \nabla_{\mathbf{v}} L \cdot d\mathbf{v} + (\partial_t L) dt \\ &= \dot{\mathbf{p}} \cdot d\mathbf{q} + \mathbf{p} \cdot d\mathbf{v} + (\partial_t L) dt \end{aligned}$$

Let us now look at the differential of H

$$\begin{aligned} dH &= \nabla_{\mathbf{q}} H \cdot d\mathbf{q} + \nabla_{\mathbf{p}} H \cdot d\mathbf{p} + (\partial_t H) dt \\ &= d(\mathbf{p} \cdot \mathbf{v} - L) \\ &= d\mathbf{p} \cdot \mathbf{v} + \mathbf{p} \cdot d\mathbf{v} - dL \\ &= d\mathbf{p} \cdot \mathbf{v} + \mathbf{p} \cdot d\mathbf{v} - \dot{\mathbf{p}} \cdot d\mathbf{q} - \mathbf{p} \cdot d\mathbf{v} - (\partial_t L) dt \\ &= d\mathbf{p} \cdot \mathbf{v} - \dot{\mathbf{p}} \cdot d\mathbf{q} - (\partial_t L) dt \end{aligned}$$

Comparing the differentials we get Hamilton equations:

Theorem 9.5 *The Hamiltonian form of Euler-Lagrange equations is*

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

which determine the evolution of the point $(\mathbf{q}, \mathbf{p})$.

In addition we have

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}$$

What have we accomplished?

- $(\mathbf{q}, \mathbf{p})$ are treated symmetrically.
- The equations of motion are simple, symmetric and elegant.
- The equations of motion is a system of first order ODE's.

9.5 Gauge freedom

The gauge ambiguity in L allows us to add a complete derivative of $\Lambda(\mathbf{q}, t)$

$$L' = L + \frac{d\Lambda}{dt} = L + \frac{\partial \Lambda}{\partial \mathbf{q}} \cdot \mathbf{v} + \frac{\partial \Lambda}{\partial t}$$

This changes the momentum:

$$\mathbf{p}' = \mathbf{p} + \frac{\partial \Lambda}{\partial \mathbf{q}} = \mathbf{p} + \nabla_q \Lambda$$

and then changes H to

$$H'(\mathbf{q}', \mathbf{p}', t) = \mathbf{p}' \cdot \mathbf{v} - L' = \left(\mathbf{p} + \cancel{\frac{\partial \Lambda}{\partial \mathbf{q}}}\right) \cdot \mathbf{v} - \left(L + \cancel{\frac{\partial \Lambda}{\partial \mathbf{q}} \cdot \mathbf{v}} + \frac{\partial \Lambda}{\partial t}\right) = H(\mathbf{q}, \mathbf{p}, t) - \frac{\partial \Lambda}{\partial t}$$

It follows that

Theorem 9.6 *Under a gauge transformation* $(\mathbf{q}, \mathbf{p}, H)$ *are transformed to* $(\mathbf{q}', \mathbf{p}', H')$ *where*

$$\mathbf{q}' = \mathbf{q}, \quad \mathbf{p}' = \mathbf{p} + \mathbf{A}, \quad H'(\mathbf{q}', \mathbf{p}', t) = H(\mathbf{q}, \mathbf{p}, t) - \partial_t \Lambda(\mathbf{q}, t)$$

where $\mathbf{A} = \nabla_q \Lambda$. *Hamilton equations are unaffected by a gauge transformation.*

Proof: We only need to verify the last assertion. Indeed,

$$\dot{\mathbf{q}}' = \nabla_{p'} H' = \nabla_{p'}(H - \partial_t \Lambda) = \nabla_{p'} H(\mathbf{q}, \mathbf{p}, t) = \nabla_p H(\mathbf{q}, \mathbf{p}, t) = \dot{\mathbf{q}}$$

Similarly

$$\dot{\mathbf{p}}' = -\nabla_{q'} H'$$

is equivalent to the equation without primes. To see this note first that

$$\dot{\mathbf{p}}' = \dot{\mathbf{p}} + \frac{d\mathbf{A}}{dt}$$

Turning to the rhs

$$\begin{aligned}
\nabla_{q'} H'(\mathbf{q}', \mathbf{p}', t) &= \nabla_{q'} H(\mathbf{q}, \mathbf{p}, t) - \partial_t \mathbf{A} \\
&= \nabla_{q'} H(\mathbf{q}', \mathbf{p}' - \mathbf{A}, t) - \partial_t \mathbf{A} \\
&= \nabla_q H(\mathbf{q}, \mathbf{p}, t) - (\nabla_p H \cdot \nabla_q)\mathbf{A} - \partial_t \mathbf{A} \\
&= \nabla_q H(\mathbf{q}, \mathbf{p}, t) - (\dot{\mathbf{q}} \cdot \nabla_q)\mathbf{A} - \partial_t \mathbf{A} \\
&= \nabla_q H(\mathbf{q}, \mathbf{p}, t) - \frac{d\mathbf{A}}{dt}
\end{aligned}$$

This gives

$$\dot{\mathbf{p}} = -\nabla_q H$$

as claimed.

Gauge freedom is the ambiguity you also find in the scalar and vector potentials in electrodynamics $(A_0, \mathbf{A})$, where the name comes from.

9.6 Examples

Example 9.2 (Small oscillations) *Let* $\mathbf{M}, \mathbf{K}$ *be positive* $n \times n$ *matrices. The Hamiltonian for small oscillations is the quadratic*

$$H = \frac{1}{2}\left(\mathbf{p}^t \mathbf{M}^{-1} \mathbf{p} + \mathbf{q}^t \mathbf{K} \mathbf{q}\right)$$

The Hamilton equations are

$$\dot{\mathbf{q}} = \mathbf{M}^{-1}\mathbf{p}, \quad \dot{\mathbf{p}} = -\mathbf{K}\mathbf{q}$$

When $\mathbf{M}, \mathbf{K}$ *are time independent, we get a system of ODE with constant coefficients. These can be solved by making the ansatz*

$$\mathbf{q} = e^{i\omega t}\mathbf{q}_0, \quad \mathbf{p} = e^{i\omega t}\mathbf{p}_0$$

Substitution in Hamilton's equations we get the secular equation:

$$\left(\omega^2 \mathbb{1} - \mathbf{M}^{-1}\mathbf{K}\right)\mathbf{q}_0 = 0$$

Example 9.3 (Charged particle in electromagnetic field) *The Lagrangian is*

$$L(\mathbf{q}, \mathbf{v}) = \frac{m}{2}\mathbf{v}^2 + e\mathbf{v} \cdot \mathbf{A}(\mathbf{q}, t) - e\Phi(\mathbf{q}, t)$$

The momentum is then

$$\mathbf{p} = m\mathbf{v} + e\mathbf{A}$$

and the Hamiltonian is

$$H = \frac{(\mathbf{p} - e\mathbf{A})^2}{2m} + e\Phi(\mathbf{q}, t)$$

Hamilton equations are

$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}} = \frac{\mathbf{p} - e\mathbf{A}}{m} = \mathbf{v}$$

and

$$\begin{aligned}
\dot{\mathbf{p}} &= -\frac{\partial H}{\partial \mathbf{q}} \\
&= -m(\mathbf{v} \cdot \nabla_q)\mathbf{v} - m\mathbf{v} \times (\nabla_q \times \mathbf{v}) - e\nabla_q \Phi \\
&= e(\mathbf{v} \cdot \nabla_q)\mathbf{A} + e\mathbf{v} \times (\nabla_q \times \mathbf{A}) - e\nabla_q \Phi \\
&= e(\mathbf{v} \cdot \nabla_q)\mathbf{A} + e\mathbf{v} \times \mathbf{B} - e\nabla_q \Phi
\end{aligned}$$

where in the first line I used the vector identity

$$\nabla(\mathbf{v} \cdot \mathbf{v}) = 2(\mathbf{v} \cdot \nabla)\mathbf{a} + 2\mathbf{a} \times (\nabla \times \mathbf{a})$$

and in the third the expression of the magnetic field in terms of the potential $\mathbf{B} = \nabla \times \mathbf{A}$*. Similarly, using*

$$\mathbf{E} = -\partial_t \mathbf{A} - \nabla\Phi$$

the equation for $\dot{\mathbf{p}}$ *takes the form*

$$\begin{aligned}
\dot{\mathbf{p}} &= e\big((\mathbf{v} \cdot \nabla_q)\mathbf{A} + \partial_t \mathbf{A}\big) + e\mathbf{v} \times \mathbf{B} + e\mathbf{E} \\
&= e\frac{d\mathbf{A}}{dt} + e\mathbf{v} \times \mathbf{B} + e\mathbf{E}
\end{aligned}$$

We have recovered Lorentz-Coulomb force

$$m\dot{\mathbf{v}} = e\big(\mathbf{v} \times \mathbf{B} + \mathbf{E}\big)$$

9.7 Phase space

The space $(\mathbf{q}, \mathbf{p})$ is called phase space. Let me start with $q \in \mathbb{R}$ a single coordinate and $L = \frac{m}{2}v^2$. Then $p \in \mathbb{R}$ and phase space is $\mathbb{R}^2$ with coordinates $\zeta = (q, p)$. It is naturally a vector space, since multiplication by scalars and additions are naturally defined. However, there is no way to associate a length of the vector ζ or the scalar product of vectors[8]. One way to see this is that the two coordinates $\zeta_1 = q$ and $\zeta_2 = p$

[8] Without adding some additional structure.

have different dimensions, so there is no way to add them without having additional information, such as conversion unit between coordinates and momenta.

Now, even though we do not know how to measure lengths in phase space, we know how to measure areas. The (oriented) area of the parallelogram defined by $\boldsymbol{\zeta} = (\zeta_1, \zeta_2)$ and $\boldsymbol{\xi} = (\xi_1, \xi_2)$ is

$$Area(\boldsymbol{\xi}, \boldsymbol{\zeta}) = \boldsymbol{\xi} \times \boldsymbol{\zeta} = \xi_1\zeta_2 - \xi_2\zeta_1 = \boldsymbol{\xi} \cdot \Omega \boldsymbol{\zeta}^t \tag{9.3}$$

where Ω is the matrix[9]

$$\Omega = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Ω is called the symplectic structure.

Theorem 9.7

$$\Omega^2 = -\mathbb{1}$$

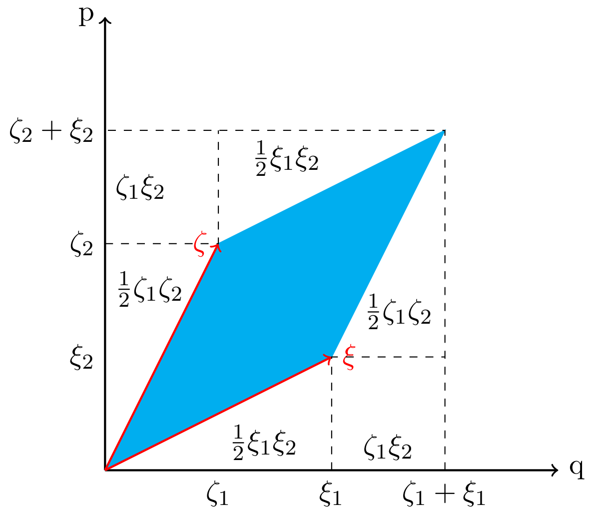


Figure 43: The formula area of the cyan parallelogram, given in Eq. 9.3, can be seen by subtracting from the big rectangle the 4 triangles and two small rectangles with areas indicated in the figure. The big rectangle has area $(\zeta_1 + \xi_1)(\zeta_2 + \xi_2)$. It follows that the cyan parallelogram has area $(\zeta_1 + \xi_1)(\zeta_2 + \xi_2) - \xi_1\xi_2 - \zeta_1\zeta_2 - 2\zeta_1\xi_2$ given in in Eq. 9.3.

The Hamiltonian H gives us a function on phase space whose gradient defines a vector field in phase space:

$$\nabla_\zeta H = (\partial_q H, \partial_p H)$$

The vector field of the flow given by Hamilton equations is rotated by $\pi/2$:

$$(\dot{\mathbf{q}}, \dot{\mathbf{p}}) = (\partial_p H, -\partial_q H)$$

[9]Actually just the Levi-Civita symbol in 2 dimensions.

Since Ω is a rotation matrix by 90 degrees, we can now write Hamilton equations as

$$\dot{\boldsymbol{\zeta}} = \Omega \nabla_\zeta H(\boldsymbol{\zeta}, t) \leftrightarrow \dot{\boldsymbol{\zeta}}_i = \Omega_{ij} \partial_j H$$

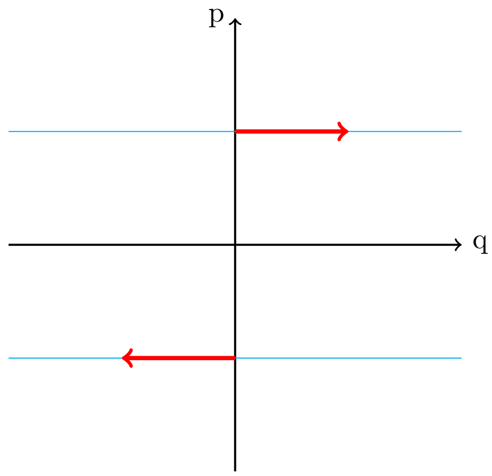


Figure 44: Phase portraits of free particle. $\nabla H = (0, p/m)$ so the line $p = 0$ is stationary. The rate at which points separate depend only on δp.

It is interesting to see how neighboring points in phase space separate. Taking the differential

$$\delta \dot{\zeta}_i = \Omega_{ij} \partial_j \delta H = \Omega_{ij} (\partial_{j\ell} H) \delta \zeta_\ell \tag{9.4}$$

In particular, for a free particle the rate of separation is:

$$\delta \dot{x} = \frac{\delta p}{m}, \qquad \delta \dot{p} = 0$$

9.8 The freedom of generalized coordinates: Phase space and the complex plane

Lagrangian mechanics gives us the freedom to choose the generalized coordinates. Consider the Harmonic oscillator with

$$L(x, v) = \frac{m}{2} v^2 - \frac{k}{2} x^2$$

Now choose a new generalized coordinate q

$$x = \lambda q$$

The Lagrangian in the new coordinate is

$$L(x, v) = L'(q, \dot{q}) = \frac{m\lambda^2}{2} \dot{q}^2 - \frac{k\lambda^2}{2} q^2$$

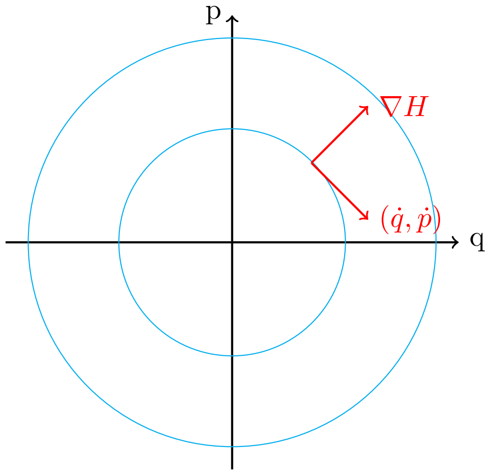


Figure 45: The gradient of H and the vector field of the flow for the Hamiltonian of the Harmonic oscillator $H = \frac{1}{2m}p^2 + \frac{k}{2}q^2$. Note that since q and p have different dimensions we can always choose scales so that the curves $H = const$ are circles rather than ellipses.

The new conjugate momentum is

$$p_q = m\lambda^2 \dot q$$

and the Hamiltonian in the new coordinates is

$$H'(q, p_q) = \frac{1}{2\lambda^2 m} p_q^2 + \frac{k\lambda^2}{2} q^2$$

This holds for any λ. Now, let me pick λ so that the the coefficients in front of p and q are the same:

$$\frac{1}{\lambda^2 m} = k\lambda^2$$

This choice associates the same dimension to the momentum and the coordinate: Both are measured in units of $[energy][sec]$ and the Hamiltonian has the elegant form

$$H' = \frac{\omega}{2}(p^2 + q^2), \quad \omega^2 = \frac{k}{m} \tag{9.5}$$

Hamilton equations are

$$\begin{pmatrix} \dot q \\ \dot p \end{pmatrix} = \omega \begin{pmatrix} p \\ -q \end{pmatrix}$$

Points in phase space separate at the rate

$$\delta \dot x = \omega \delta p, \quad \delta \dot p = -\omega \delta q$$

It follows that

$$0 = \delta q \delta \dot q + \delta p \delta \dot p = \frac{1}{2}\frac{d}{dt}\big((\delta q)^2 + (\delta p)^2\big)$$

It follows that the Euclidean distance:

$$(\delta q)^2 + (\delta p)^2$$

is conserved by the flow. In fact, the flow is a rigid rotation of phase space at rate ω.

Here is a nice way to see this. We can identify phase space with the complex plane by writing

$$z = q + ip$$

The symplectic structure is then simply the product by i

$$\Omega \Leftrightarrow i$$

and Hamilton equations take the form

$$\dot{z} = -i\omega z$$

This is immediately solved by

$$z = e^{-i\omega t} z_0$$

This corresponds to clockwise rigid rotation of phase space.

9.9 The cylinder as phase space: pendulum

Phase space need not always be a vector space. Here is an example where phase space is a cylinder: The Lagrangian and the generalized momentum of a mathematical pendulum are

$$L = \frac{m}{2}\ell^2\dot{\theta}^2 + mg\ell\cos\theta, \quad J = m\ell^2\dot{\theta}$$

Coordinate space is the circle $\theta \in \mathbb{S}$, which is not a vector space. Momentum space is still the line $J \in \mathbb{R}$ which is. Phase space is the cylinder:

$$\mathbb{S} \times \mathbb{R}$$

The Hamiltonian is

$$H = \frac{J^2}{2m\ell^2} - mg\ell\cos\theta$$

The contour plot of H is shown in the figure. By conservation of energy the motion is on level sets. If the energy is small, the level sets are topologically circles. The motion is oscillations. If the momentum is large enough the motion are the wavy horizontal lines: The pendulum makes rotations. The two types off motions are separated by a line called the separatrix.

Exercise 9.4 *Plot the vector field $(\dot{q}, \dot{p})$ near the the stable and unstable fixed point. How do they differ?*

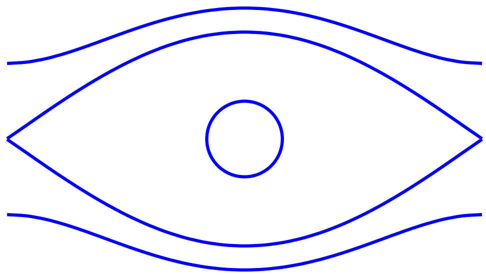


Figure 46: Contour plot for the Hamiltonian of mathematical pendulum. Phase space is a cylinder with $-\pi < \theta < \pi$ and $-\infty < J < \infty$. The stationary points are $(0,0)$ and $\pm\pi, 0)$. The first is stable the second is not. The integral curves meet at the unstable stationary point.

9.10 Liouville theorem

Theorem 9.8 *The Hamiltonian flow preserves the area (more generally, volume) in phase space*

Proof: Consider and infinitesimal area spanned by $(\delta\boldsymbol{\xi}, \delta\boldsymbol{\zeta})$. Denote the area by

$$\delta A = \delta\xi_j \Omega_{jk} \delta\zeta_k$$

Then, by Hamilton equation, the area evolves by

$$\begin{aligned}
\frac{d(\delta A)}{dt} &= \delta\dot{\xi}_j \Omega_{jk} \delta\zeta_k + \delta\xi_j \Omega_{jk} \delta\dot{\zeta}_k \\
&= \Omega_{jk} \left(\delta\dot{\xi}_j \delta\zeta_k + \delta\xi_j \delta\dot{\zeta}_k \right) \\
&= \Omega_{jk} \left(\Omega_{ja} \partial_{ab} H \delta\xi_b \, \delta\zeta_k + \Omega_{ka} \, \partial_{ab} H \delta\xi_j \delta\zeta_b \right) \\
&= \partial_{ab} H \left(\delta_{ka} \delta\xi_b \, \delta\zeta_k - \delta_{ja} \delta\xi_j \, \delta\zeta_b \right) \\
&= \partial_{ab} H \left(\delta\xi_b \, \delta\zeta_a - \delta\xi_a \, \delta\zeta_b \right) && (9.6) \\
&= 0 && (9.7)
\end{aligned}$$

In the third line we used Eq. (9.4). In the 4-th line that $j^2 = -\mathbb{1}$ and in the last line that the trace of the product of a symmetric matrix (the first term) with an anti-symmetric matrix (the second) vanishes.

9.11 Phase space in higher dimensions

Suppose the configuration space in $\mathbb{R}^n$ with Euclidean coordinates $\mathbf{q}$. This makes configuration space a vector space. Suppose the kinetic energy is quadratic in the velocities. The momenta $\mathbf{p}$ are then also Euclidean coordinates in $\mathbb{R}^n$. The coordinates and momenta are paired via

$$p_j = \frac{\partial L}{\partial \dot{q}_j}$$

Definition 9.4 $\mathbb{R}^{2n}$ *with coordinates* $\boldsymbol{\zeta} = (\mathbf{q}, \mathbf{p})$ *is called phase space. The pairing of the coordinates and momenta is called the symplectic structure and given by a* $2n \times 2n$ *matrix* Ω *with components*

$$\Omega_{q_j,p_k} = -\Omega_{p_j,q_k} = \delta_{j,k}$$

The choice of signs Ω agrees with the choice we made in 1 dimension.

The Hamiltonian H defines a vector field on phase space that allows to propagate any point in it in time

$$d\mathbf{q} = \frac{\partial H}{\partial \mathbf{p}} dt, \quad d\mathbf{p} = -\frac{\partial H}{\partial \mathbf{q}} dt$$

You may think of a point in phase space as representing the initial conditions of a particle. It then has a well defined future (and past) obtained by solving the equations of motion.

We can repeat the construction in 1 dimension. With H a function on phase space, its gradient defines a $2n$ vector field in phase space:

$$\nabla_\zeta H = (\nabla_q H, \nabla_p H)$$

Compare this vector field with the vector field of the flow given by Hamilton equations

$$(d\mathbf{q}, d\mathbf{p}) = (\nabla_{\mathbf{p}} H, -\nabla_{\mathbf{q}} H)\, dt$$

Clearly

$$(\nabla_{\mathbf{p}} H, -\nabla_{\mathbf{q}} H)^t = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} (\nabla_q H, \nabla_p H)^t = \Omega(\nabla_q H, \nabla_p H)^t$$

This allows to write Hamilton equations in a compact form

$$\dot{\boldsymbol{\zeta}} = \Omega \nabla_\zeta H(\boldsymbol{\zeta}, t)$$

A solution of this first order system of equations starting at some fixed point $\dot{\boldsymbol{\zeta}}$ gives a curve in phase space. The curve is called an "integral curve".

It follows:

- A point is a stationary point if $\nabla_\zeta H = 0$
- If $\nabla_\zeta H \neq 0$, there is a unique direction of motion
- Integral curves can not cross. They can meet at stationary points.

Phase space formulation is geometric and intuitive. The vector field

$$\nabla H = (\partial_x H, \partial_p H)$$

is orthogonal to the energy contours $H(x,p) = const$. The Hamiltonian vector field

$$(\partial_p H, -\partial_x H)$$

is tangent to the energy contours. In particular, if H is time independent the energy contours are time independent, and the motions stays on a fixed energy surface.

10 Action, Hamilton Jacobi equation

We introduced the action as a functional from the space of paths to $\mathbb{R}$

$$S : path \mapsto \mathbb{R}$$

with fixed starting point and time (q_i, t_i) and fixed end point and time (q_i, t_i). Explicitly,

$$S[q(t)] = \int_{q_i,t_i}^{q_f,t_f} L(q, \dot{q}, t)dt$$

We shall now reverse the point of view, and consider the action as a function of the end-

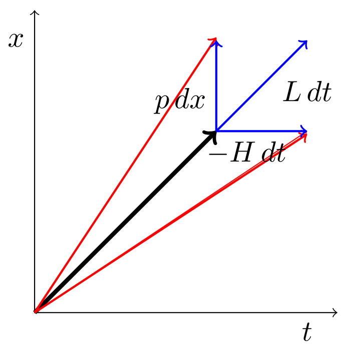


Figure 47: Variation of the action of the classical path due to variation of the end point. The thick black arrow is the classical path to the end point (q, t).

point, for the path that satisfies Euler-Lagrange, now S is an ordinary function $S(q, t)$ with (q, t) the end points.

Example 10.1 *For a free particle $L = \frac{m}{2}\dot{q}^2$. Suppose the particle starts at (q_0, t_0) and ends at (q, t). The solution of Euler-Lagrange has constant speed so*

$$\dot{q} = \frac{q - q_0}{t - t_0}$$

and then the action is simply

$$S(q, t) = \frac{m(q - q_0)^2}{2(t - t_0)}$$

Theorem 10.1

$$dS = \partial_q S\, dq + \partial_t S\, dt = pdq - Hdt$$

Proof: For a variation of the path with fixed initial and final end points we found

$$\delta S = p\,\delta q\big|_{t_i}^{t_f} - \int_{t_i}^{t_f} dt \left(\dot{p} - \frac{\partial L}{\partial q}\right)\delta q$$

On the classical path the brackets on the right vanish and a variation of q_f (while keeping t_f fixed) gives the first part of the theorem

$$\delta S = p\,\delta q\big|_{t_i}^{t_f} \Longrightarrow \frac{\partial S}{\partial q} = p$$

What about a variation of t_f? By inspecting the figure you can see that

$$dS(q_f, t_f) = Ldt = p\,dq + (\partial_t S)\,dt$$

Hence

$$-\partial_t S = p\dot{q} - L = H$$

Exercise 10.1 *Compute S for the Harmonic oscillator.*

The statements about the partial derivatives of S can also be rephrased as a single, first order, non-linear partial differential equation for S

Theorem 10.2 *$S(q,t)$ satisfy the Hamilton-Jacobi equation*

$$H\left(q, \partial_q S, t\right) = -\partial_t S$$

One way to think about this is actually in reverse: If you have a non-linear first order partial differential equation, then you should be able to solve it by Hamilton equations which are ODE's rather than PDE. This is known as the method of characteristics.

11 Liouville theorem, Poisson brackets

11.1 Incompressible fluids

Here is a different way to see that classical evolution preserves the volume in phase space. Consider a fluid with density $\rho(\mathbf{x},t) > 0$ and current density $\mathbf{J}(\mathbf{x},t) = \rho(\mathbf{x},t)\mathbf{v}(\mathbf{x},t)$. Conservation of mass leads to the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0, \quad \mathbf{J} = \rho \mathbf{v} \tag{11.1}$$

Definition 11.1 *A homogenous fluid is said to be incompressible if $\rho(\mathbf{x},t) = \rho_0 > 0$ independent of $\mathbf{x}$ and t.*

Water is (approximately) incompressible (at low frequencies the small pressure gradients).

Theorem 11.1 *The velocity field $\mathbf{v}(\mathbf{x},t)$ of an incompressible fluid is divergence-less, $\nabla \cdot \mathbf{v} = 0$.*

This follows immediately from $\rho = \rho_0 \neq 0$ being a constant and the continuity equation

$$\underbrace{\frac{\partial \rho}{\partial t}}_{=0} + \nabla \cdot \mathbf{J} = \rho \nabla \cdot \mathbf{v} = 0$$

Theorem 11.2 (Liouville) *Let phase space be $\mathbb{R}^{2n}$. It behaves like an incompressible homogeneous fluid in the sense that the Hamiltonian flow is divergence-less.*

Proof: Let $\mathbf{v} = (\dot{\mathbf{x}}, \dot{\mathbf{p}})$. By Hamilton equations

$$\nabla \cdot \mathbf{v} = \sum \left(\frac{\partial \dot{x}_j}{\partial x_j} + \frac{\partial \dot{p}_j}{\partial p_j} \right) = \sum \left(\frac{\partial}{\partial x_j} \frac{\partial H}{\partial p_j} - \frac{\partial}{\partial p_j} \frac{\partial H}{\partial x_j} \right) = 0$$

by the commutativity of mixed derivatives.

This is perhaps the most important result of Hamiltonian mechanics. Note the generality. No assumption on H need be made. It holds for arbitrary number of particles with arbitrary interactions and H may be time dependent.

A classical demonstration of some surprising features of incompressibe flows is given here: Youtube: Reversible Low Reynolds number flow.

11.2 Poisson brackets

I shall call a function on phase space an observable. $e^{-\beta H}$ is an examples.

Definition 11.2 (Poisson brackets) *The Poisson brackets of the observables f and g are defined by*

$$\{f, g\} = \sum_{j=1}^{n} \left(\frac{\partial f}{\partial p_j} \frac{\partial g}{\partial q_j} - \frac{\partial g}{\partial p_j} \frac{\partial f}{\partial q_j} \right) = - \sum_{j,k=1}^{2n} \Omega_{jk} \left(\frac{\partial f}{\partial \xi_j} \right) \left(\frac{\partial g}{\partial \xi_k} \right)$$

Theorem 11.3

$$\{\xi_j, \xi_k\} = -\Omega_{jk} \tag{11.2}$$

and more explicitly

$$\{q_j, q_k\} = \{p_j, p_k\} = 0, \quad \{p_j, q_k\} = \delta_{jk},$$

Theorem 11.4 *Poisson brackets are anti-symmetric*

$$\{f, g\} = -\{g, f\}$$

Linear

$$\{f, g_1 + g_2\} = \{f, g_1\} + \{f, g_2\}$$

and satisfy Leibnitz rule

$$\{f, g_1 g_2\} = \{f, g_1\} g_2 + g_1 \{f, g_2\}$$

and the chain rule

$$\{f(g), h\} = f'(g) \{g, h\}$$

Proof: By inspection

Theorem 11.5 *Poisson brackets satisfy the Jacobi identity*

$$\{f,\{g,h\}\}+\{g,\{h,f\}\}+\{h,\{f,g\}\}=0 \tag{11.3}$$

Proof: You can try and prove this by brute force: You will then have to write many terms (30) and not make any sign mistake. A better strategy its to be clever. Consider the first term

$$\begin{aligned}
\{f,\{g,h\}\} &= -\sum \Omega_{jk}(\partial_j f)\partial_k(\{g,h\}) \\
&= \sum \Omega_{jk}\Omega_{ab}(\partial_j f)\partial_k\Big((\partial_a g)(\partial_b h)\Big) \\
&= \sum \Omega_{jk}\Omega_{ab}(\partial_j f)\Big((\partial_{ak} g)(\partial_b h)+(\partial_a g)(\partial_{bk} h)\Big)
\end{aligned} \tag{11.4}$$

Focus on the term proportional to $\partial_{ak}g$. You can get terms with two derivatives of g only from the first and third terms in Eq. (11.3). Since the first and last terms are related by anti-cyclic permutations the second derivative of g will come from

$$\begin{aligned}
&\sum \Omega_{jk}\Omega_{ab}\Big((\partial_j f)(\partial_{ak} g)(\partial_b h)+(\partial_j h)(\partial_a f)(\partial_{bk} g)\Big) \\
&= \sum \Omega_{jk}\Omega_{ab}(\partial_{ak} g)\Big((\partial_j f)(\partial_b h)-(\partial_j h)(\partial_b f)\Big) \\
&= \frac{1}{2}\sum \underbrace{(\Omega_{jk}\Omega_{ab}+\Omega_{ja}\Omega_{kb})}_{jb-symmetric}(\partial_{ak} g)\underbrace{\Big((\partial_j f)(\partial_b h)-(\partial_j h)(\partial_b f)\Big)}_{jb-anti} \\
&= 0
\end{aligned}$$

In the second line I interchanged the summation indexes $a \leftrightarrow b$ in the second term. In the third line I used the symmetry in $a \leftrightarrow k$ of the second derivative, to symmetrize also the first term in $a \leftrightarrow b$. Alternatively, you could feed this into Mathematica and let it sweat.

11.3 Single particle observables: Scalars and vectors

Consider a single particle (scalar) observable $f(\mathbf{x})$. Under rotations of the coordinates $\mathbf{x}' = \mathbf{R}\,\mathbf{x}$ with $\mathbf{R}$ and orthogonal matrix the rotated observable is

$$f'(\mathbf{x}') = f(\mathbf{x}) = f(\mathbf{R}^t\mathbf{x}')$$

For a rotation about the z axis

$$x' = x\cos\theta + y\sin\theta, \quad y' = -x\sin\theta + y\cos\theta, \quad z' = z \tag{11.5}$$

we then have

$$f'(x',y',z') = f(x'\cos\theta - y'\sin\theta, x'\sin\theta + y'\cos\theta, z')$$

The rate of change of f' at $\theta = 0$ is

$$\begin{aligned}
\frac{df'}{d\theta} &= -y'\partial_x f + x'\partial_y f \\
&= -y\partial_x f + x\partial_y f \\
&= -y\{p_x, f\} + x\{p_y, f\} \\
&= \{-yp_x, f\} + \{xp_y, f\} \\
&= \{xp_y - yp_x, f\} \\
&= \{J_z, f\}
\end{aligned} \tag{11.6}$$

In the second line I used the fact that when $\theta = 0$ the primed and unprimed coordinates coincide. In the third I replaced derivatives by Poisson brackets and then used linearity and Leibnitz.

We see something we already saw before: The intimate relation between angular momentum and rotations.

The basic single particle vector observables are

$$\mathbf{x} = (x, y, z), \quad \mathbf{p} = (p_x, p_y, p_z), \quad \mathbf{J} = \mathbf{x} \times \mathbf{p} \tag{11.7}$$

It is a feature of vectors that all vectors behave under rotation in the same way as the coordinate behave under rotation. From the computation above the following remarkable observation comes:

Theorem 11.6 *Let* $\mathbf{A} = \{A_x, A_y, A_z\}$ *be any one of the vectors in Eq. (11.7). The Poisson brackets with the angular momentum is the same:*

$$\{J_j, A_k\} = -\varepsilon_{ijk} A_i \tag{11.8}$$

Proof: Compute.

Exercise 11.1 *Show that* $\mathbf{x}, \mathbf{p}$ *and* $\mathbf{J}$ *have the Poisson brackets of vectors.*

Theorem 11.7 *Let* $\mathbf{A}, \mathbf{B}$ *be one particle observables that have the same Poisson brackets as the basic vectors then* $\mathbf{A} \cdot \mathbf{B}$ *vanishing Poisson brackets with the angular momentum.*

Proof:

$$\begin{aligned}
\{A_a B_a, J_k\} &= A_a\{B_a, J_k\} + \{A_a, J_k\}B_a \\
&= \varepsilon_{ak\ell} A_a B_\ell + \varepsilon_{ak\ell} A_\ell B_a \\
&= 0
\end{aligned}$$

11.4 Poisson's Equations of motion

Hamilton equations take the concise form

$$\dot{p} = -\frac{\partial H}{\partial q} = \{H, p\}, \quad \dot{q} = \frac{\partial H}{\partial p} = \{H, q\} \tag{11.9}$$

Remark 11.1 *Heisenberg created quantum mechanics from these equation and by giving Poisson brackets a different interpretation.*

For the observable f we have

$$\begin{aligned}\frac{df}{dt} &= \partial_t f + \dot{p}\partial_p f + \dot{x}\partial_x f \\ &= \partial_t f + \underbrace{\{H,p\}\partial_p f + \{H,x\}\partial_x f}_{By\ Eq.(11.9)} \\ &= \underbrace{\{H,f\}}_{chain\ rule} + \frac{\partial f}{\partial t}\end{aligned} \tag{11.10}$$

11.5 Constants of motion–Poisson theorem

Theorem 11.8

$$\lim \frac{f(\xi + \varepsilon\Omega\nabla H, t+\varepsilon) - f(\xi,t)}{\varepsilon} = \{H,f\} + \partial_t f \tag{11.11}$$

Proof:

$$f(\xi + \varepsilon\Omega\nabla H, t+\varepsilon) = f(\xi,t) + \varepsilon\Omega_{jk}\partial_k H\partial_j f = f(\xi) + \varepsilon\{H,f\} + \varepsilon\partial_t f$$

It follows from this that

Definition 11.3 *An observable f is a constant of motion if*

$$\dot{f} = 0 \Longrightarrow \{H,f\} + \partial_t f = 0$$

This means that f is preserved by the flow: The flow lines in phase space lie on the surface of constant values of f.

Since

$$\{H,H\} = 0 \longrightarrow \dot{H} = \partial_t H$$

a time independent Hamiltonian is a constant of motion. More generally, by the chain rule for Poisson brackets any function of H is a constant of motion in this case

$$f(H) \longrightarrow \dot{f} = f'(H)\partial_t H = 0$$

is a constant of motion if H is time independent.

Theorem 11.9 (Poisson) *Suppose f and g are constants of motion then their Poisson brackets is also constants of motion.*

By assumption

$$0 = \dot{f} = \partial_t f + \{H,f\} = \dot{g} = \partial_t g + \{H,g\}, \tag{11.12}$$

Using this, and Jacobi

$$
\begin{aligned}
\frac{d\{f,g\}}{dt} &= \underbrace{\partial_t\{f,g\} + \{H,\{f,g\}\}}_{by\ definition} \\
&= \underbrace{\{\partial_t f, g\} + \{f, \partial_t g\} - \{g,\{H,f\} - \{f,\{g,H\}\}}_{Leibnitz+Jacobi} \\
&= \{\partial_t f, g\} + \{f, \partial_t g\} - \{g,\{H,f\}\} \underbrace{+\{f,\{H,g\}}_{antisymmetry}\} \\
&= \{\partial_t f, g\} + \{f, \partial_t g\} + \underbrace{\{g, \partial_t f\} - \{f, \partial_t g\}}_{by\ 11.12} \\
&= 0
\end{aligned}
\tag{11.13}
$$

You would think that Poisson theorem would allow you to generate infinitely many constant of motions. It normally does not. The Poission bracket may vanish, or may give you a function of the constants you already had.

11.6 Canonical transformations

The Lagrangian formulation gave us the freedom to chose generalized coordinates. No matter which coordinates we pick, the equations of motions are the same

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

The analogous freedom in Hamiltonian mechanics are called canonical transformations

Definition 11.4 *A transformation $\eta_j = \eta_j(\xi_1, \dots \xi_{2n})$ is canonical if it preserves Poisson brackets*

$$
\{\xi_j, \xi_k\} = \{\eta_j, \eta_k\} = -\Omega_{jk}
$$

Example 11.1 (Scalinng) *Consider the freedom to scale the coordinate*

$$
q' = \lambda q, \quad L'(q', \dot q', t) = L(q, \dot q, t)
$$

Under scaling $q' = \lambda q$ the momenta scale $p' = \frac{1}{\lambda}p$. This follows from

$$
p' = \frac{\partial L'}{\partial \dot q'} = \frac{\partial L}{\partial \dot q}\frac{d\dot q}{d\dot q'} = \frac{1}{\lambda}p
$$

We see that scaling acts on phase space by

$$
\begin{pmatrix} q' \\ p' \end{pmatrix} = \underbrace{\begin{pmatrix} \lambda & 0 \\ 0 & \frac{1}{\lambda} \end{pmatrix}}_{C} \begin{pmatrix} q \\ p \end{pmatrix} \Longleftrightarrow \xi'_j = C_{jk}\xi_k
$$

Since

$$C^t \Omega C = \Omega$$

we have

$$\Omega'_{jk}\xi'_j\eta'_k = \Omega'_{jk}C_{ja}\xi_a C_{kb}\eta_b = (C^t\Omega' C)_{ab}\xi_a\eta_b = \Omega'_{ab}\xi_a\eta_b = \Omega_{ab}\xi_a\eta_b$$

Hence, Poisson brackets and the area are conserved.

Theorem 11.10 *If ξ and η are canonical then*

$$\{f, g\}_\xi = \{f, g\}_\eta$$

where

$$\{f, g\}_\xi = -\sum_{jk} \Omega_{jk} \frac{\partial f}{\partial \xi_j} \frac{\partial f}{\partial \xi_k}$$

Proof: It is enough to show this for the coordinates

$$\{\xi_j, \xi_k\}_\xi = \{\xi_j, \xi_k\}_\eta$$

This follows from

$$\{\xi_j, \xi_k\}_\xi = \{\eta_j, \eta_k\}_\xi = \Omega_{jk} = \{\eta_j, \eta_k\}_\eta$$

Just as the Euler Lagrange equations do not care which generalized coordinates you pick, the Hamiltonian equations do not care which canonical coordinates you take. Formally

Theorem 11.11 *Time independent canonical transformations preserve Poisson's equation i.e. with $H'(\eta) = H(\xi)$ we have*

$$\dot\eta_j = \{H', \eta_j\}_\eta \Longleftrightarrow \dot\xi_k = \{H, \xi_k\}_\xi$$

where $H'(\eta) = H(\xi)$, and the Poisson brackets can be computed in any canonical coordinate system.

Proof: The left hand side the two equations are related by

$$\dot\eta_j = \frac{\partial \eta_j}{\partial \xi_k}\dot\xi_k$$

The rhs are related by

$$\{H', \eta_j\}_\eta = \{H', \eta_j\}_\xi = \{H, \eta_j\}_\xi = \Omega_{ak}\frac{\partial H}{\partial \xi_a}\frac{\partial \eta_j}{\partial \xi_k} = \dot\xi_k \frac{\partial \eta_j}{\partial \xi_k}$$

Remark 11.2 *The case that the canonical transformation is time dependent is more complicated.*

Theorem 11.12 *Let $\xi_t = \xi_t(\xi)$ be the image of the point ξ under the Hamiltonian flow after some time t. The coordinates ξ_t are canonical.*

We need to show that

$$\{\xi_{tj,}\xi_{tk}\} = -\Omega_{jk}$$

Denote the lhs $f_t = \{\xi_{tj}, \xi_{tk}\}$. The time evolution of f_t is given by the linear first order differential equation

$$\dot{f}_t = \{H, f_t\}_\xi$$

By the theory of ODE, initial data for f_t, the solution is unique. The initial data is $f_{t=0} = -\Omega_{jk}$. The constant solution $f_t = -\Omega$ satisfies the differential equation, since Poisson brackets of H with Ω vanish and it also satisfies the initial condition at $t = 0$ and so the unique solution.

11.7 Action angle

Action angles are special kind of canonical coordinates. which we denote (S_j, θ_j) with the property that S_j have dimensions of action $[S_j] = [Action]$ and θ_j are angles, i.e. θ_j is identified with $\theta_j + 2\pi$. For example, consider the Harmonic oscillator Eq. 9.5 in (almost) polar coordinates

$$2S = p^2 + q^2 = \frac{H'}{\omega}, \quad \tan\theta = \frac{p}{q}, \quad [S] = [action] = [\hbar]$$

The area in (q, p) and (S, θ) are related by

$$Area = |dqdp| = rdrd\theta = \frac{d(p^2 + q^2)}{2} d\theta = |dSd\theta|$$

(If you think of the area as a signed quantity you will need to worry about the signs too.) The Jacobian of the transformation $(q, p) \leftrightarrow (\theta, S)$ is area preserving. This means that (θ, S) are canonical coordinates.

Example 11.2 *To verify that the pair (S, θ) is canonical compute*

$$\begin{aligned}
-\{S, \theta\}_{pq} &= \frac{\partial S}{\partial q}\frac{\partial \theta}{\partial p} - \frac{\partial S}{\partial p}\frac{\partial \theta}{\partial q} \\
&= \left(q\frac{\partial \tan\theta}{\partial p} - p\frac{\partial \tan\theta}{\partial q}\right)\left(\frac{d\tan\theta}{d\theta}\right)^{-1} \\
&= \left(1 + \frac{p^2}{q^2}\right)\left(\frac{d\tan\theta}{d\theta}\right)^{-1} \\
&= 1
\end{aligned} \tag{11.14}$$

Action angle for Pendulum

In case you got a little tired from the Harmonic oscillator, lets us look at a more complicated problem namely, the pendulum. Pick the energy low enough so that the motion is an oscillation, but not necessarily so low that it is harmonic. I want to introduce new canonical coordinate, action angle, coordinates. First define

$$S(E) = \int_\gamma p dq$$

The integration is on a fixed energy contour. Clearly $S(E)$ is a 1-1 function in the relevant phase space. Its geometric meaning is that of the phase space area enclosed in γ. dS is the area between two neighboring contours. Now choose $d\theta$ so that equal areas are covered in equal angles:

$$|dSd\theta| = |dqdp|$$

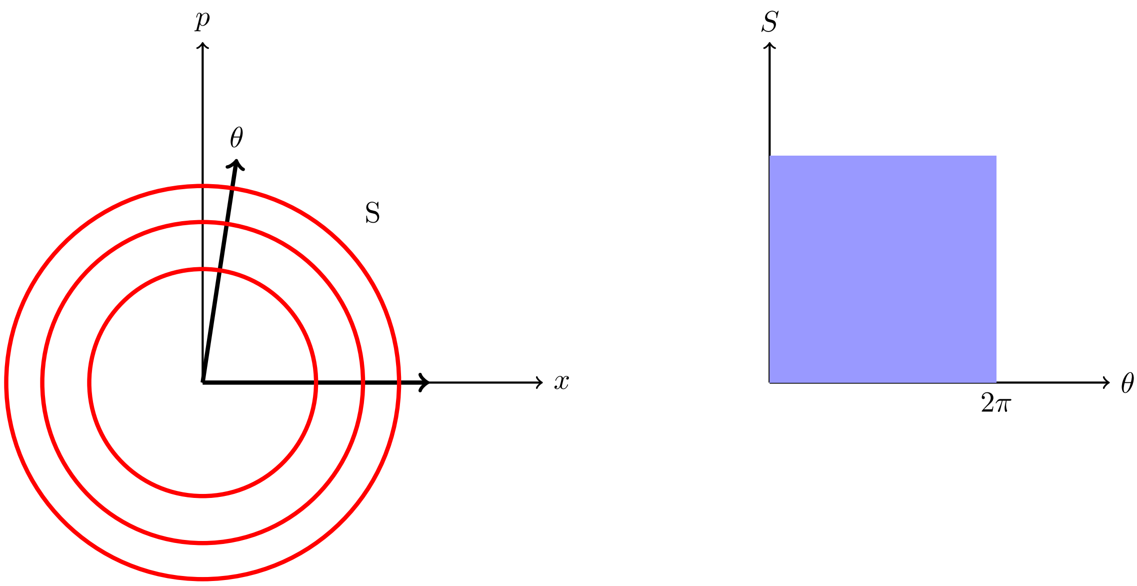


Figure 48: Action angles: The circles are curves of constant action for $S(E) \in 1, 2, 3$. The distance between the circles decreases with energy since S measures area. The angles are the conjugate variables. The areas agree if $2\pi S = 2\frac{H}{\omega}\pi$

Remark 11.3 *The case when the energy is large so that the motion is a libration is more complicated. For a given energy E there are two orbits γ_1 and γ_2, one turning to the right and one turning to the left with the same energies and actions.*

11.8 Integrable Hamiltonians

Definition 11.5 *A time independent Hamiltonian H is called (completely) integrable, if there are action-angels coordinates so that*

$$H = H(S_1 \ldots, S_n)$$

where

$$\{S_j, S_k\} = \{\theta_j, \theta_k\} = 0, \quad \{S_j, \theta_k\} = \delta_{jk}, \quad \theta_j \equiv \theta_j + 2\pi$$

It is called so because we can readily integrate the equations of motion. All the actions are constants of motion

$$\dot{S}_j = -\frac{\partial H}{\partial \theta_j} = 0 \Longrightarrow S_j(t) = S_j(0)$$

And all the angels move at uniform velocities that depend on actions S

$$\dot{\theta}_j = \frac{\partial H}{\partial S_j} = \omega_j(S) \Longrightarrow \theta_j(t) = \theta_j(0) + \underbrace{\omega_j(S)}_{constants}\ t$$

The period depends on S, in general which is another way to say that the period depends on the amplitude. This is the case for the pendulum..

I shall not proof the following fact

Theorem 11.13 *Quadratic Hamiltonians, or equivalently, the Hamiltonians of small oscillations admit action angle coordinates and a linear function of the actions:*

$$H = \sum \omega_j S_j, \quad , \{S_j, \theta_k\} = \delta_{jk}, \quad \{\theta_j, \theta_k\} = \{S_j, S_k\} = 0,$$

Not all Hamiltonians are integrable. If H is integrable, $H = H(S)$ the motion takes place on an n dimensional torus of the angle coordinates, while keeping all the actions S_j constant. The frequencies are then functions of the actions $S = (S_1, \dots, S_n)$

$$\omega_j(S) = \frac{\partial H}{\partial S_j}$$

Definition 11.6 *The periods are said to be rationally related if there are integers n_j, not all zero, so that*

$$\sum n_j \omega_j = 0 \tag{11.15}$$

If the periods are not rationality related the orbit will eventually cover all the torus of angles uniformly. (The orbit will close on itself if there are $n-1$ independent rational relations.)

Lissajous figures

11.9 Types of orbits in phase space

Two extreme types and some intermediate types of possible behaviors in mechanics are:

- The classical orbit closes on itself: A one dimensional closed circle in phase space.
- The classical orbit never closes and covers densely all of phase space. The (closure) of the orbit is $2n$ dimensional.

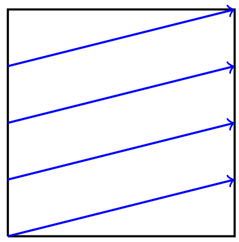


Figure 49: The motion on the torus in angle coordinates. If the ratio of periods is rational the orbit closes. It the ration of the periods is irrational it does not. The orbit will then cover the torus of angles uniformly in time. If $\omega_j(S)$ is

- The time independent Hamiltonian is integrable and the motion is restricted to the n dimensional angle torus.
- The Hamiltonian is time dependents and the energy is conserved. The motion is restricted to the constant energy surface $H(x,p) = E$ in phase space which is $2n-1$ dimensional.
- There are $0 \le m \le n$ independent constants of motion K_j, with mutually vanishing Poisson brackets. We can then view these m conserved quantities as momenta. The motion is restricted to the $2n-m$ dimensional shell in phase space.

11.10 Ergodicity

Consider the trace of a point in phase space in time

$$t : \xi = (\mathbf{q}, \mathbf{p}) \mapsto \xi(t) = (\mathbf{q}(t), \mathbf{p}(t))$$

The motion may take place in high dimension and can be complicated.

If you take a photograph of phase space with long exposure time T the photo will show a curve in phase space which is described by the function (actually distribution)

$$P_T(\xi) = \frac{1}{T}\int_0^T \delta\Big(\xi(t) - \xi\Big)\, dt$$

Theorem 11.14 *$P_T(\xi)$ is a probability density on phase space, since $P_t(\xi) \ge 0$ and*

$$\int P_T(\xi)\, dVol(\xi) = 1$$

The probability P_T gives the fraction of the time the orbit $\xi(t)$ visited the point ξ.

Theorem 11.15 *Suppose P_T has a limit P as $T \to \infty$*[10] *then the observable P on phase space is a constant of motion:*

$$0 = \dot{P}\big(\xi\big) = \{H, P\}$$

Proof. If the limit exists, then

$$P(\xi) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \delta\Big(\xi(u) - \xi\Big)\, du = \lim_{T\to\infty} \frac{1}{T} \int_0^T \delta\Big(\xi(u+s) - \xi\Big)\, du$$

is independent of s. Now take the derivative wrt s of both sides:

$$0 = \frac{dP(\xi)}{ds} = \lim_{T\to\infty} \frac{1}{T} \int_0^T \frac{d}{ds}\delta\Big(\xi(u+s) - \xi\Big)\, du$$

But, the derivative at $s = 0$ is

$$\begin{aligned}
\frac{d}{ds}\delta\Big(\xi(u+s) - \xi\Big) &= \partial_j \delta\Big(\xi(u+s) - \xi\Big)\dot{\xi}_j(u) \\
&= \partial_j \delta\Big(\xi(u+s) - \xi\Big)\{H, \xi_j(u)\}_\xi \\
&= \partial_j \delta\Big(\xi(u+s) - \xi\Big)\{H, \xi_j(u)\}_{\xi(u)} \\
&= \partial_j \delta\Big(\xi(u+s) - \xi\Big)\Omega_{aj}\partial_a H \\
&= -\delta\Big(\xi(u+s) - \xi\Big)\Omega_{aj}\partial_{aj} H \\
&= 0
\end{aligned}$$

In the penultimate line I integrated by parts. It follows that the $P(\xi)$ is a constant of motion.

Definition 11.7 (Standard, slightly vague) *A time-independent Hamiltonian H is called ergodic if for most initial data ξ in phase space time-average is given by phase space average:*

$$P(\xi) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \delta\Big(\xi(t) - \xi\Big)\, dt = \frac{\delta\big(H(\xi) - E\big)}{Z}, \quad E = H\big(\xi(0)\big)$$

with Z a normalization which counts that volume of an energy shell.

The 1 dimensional Harmonic oscillator is ergodic in this sense and $Z = \frac{4\pi}{\omega}$.

It is believed that:

The Ergodic Hypothesis *A mechanical system with a time-independent Hamiltonian, and no symmetries is generically ergodic.*

[10] The limit may exist only in the "weak sense". This roughly means that you first want to replace the delta function by a narrow Gaussian, and after taking the $T \to \infty$ limit take the limit where the Gaussian goes to a delta.

Genericity protects the hypothesis from annoying counter-examples where the parameters are fine tuned. For example, the hypothesis will survive counter examples if ergodicity fails for a countable set of initial data.

The main application of this is in statistical physics where one assumes ergodicity and define

Definition 11.8 (Entropy) *The Boltzman entropy S is defined by*

$$S = k_B \log Z$$

An isolated mechanical system with no special symmetries time averages can be replaced by phase space average over the energy shell of the system.

This should be contrasted with the integrable case where

$$P(\xi) = N \prod_j \delta(S_j - S_j(0))$$

Integrable Hamiltonians with $n \geq 2$ are not ergodic.

Exercise 11.2 *Use*

$$\int \delta(g(p))dp = \sum \frac{1}{|g'(p_j)|}, \quad \forall p_j \textit{ such that } g(p_j) = 0$$

to show that

$$\int \delta\left(\frac{p^2}{2m} + \frac{kx^2}{2} - E\right) dp = \frac{4m}{|p(x,E)|}, \quad p(x,E) = \pm\sqrt{2m\left(E - \frac{kx^2}{2}\right)} \tag{11.16}$$

Exercise 11.3 *Show that the Harmonic oscillator in one dimension is ergodic.*

It is a little disappointing that although one believes that "generic" Hamiltonians are ergodic, I do not know of any simple non-trivial and non-contrived example (i.e. not in in one dimension).

11.11 Adiabatic theorem

Theorem 11.16 *If H is slowly time dependent with rate ε and for each fixed time admits action angle variables, then the actions are adiabatic invariants: They change by at most $O(\varepsilon)$ (and often by $O\left(e^{-1/\varepsilon}\right)$ for any n) if the rate of change is slow, even if the total change of the Hamiltonian is $O(1)$.*

Let us see how this works for the time dependent oscillator

$$H(t) = \frac{p^2 + \overbrace{\omega^2(s)}^{\textit{time dependent}} x^2}{2}, \quad s = \frac{t}{T}$$

T is the adiabatic time scale, i.e. $T\omega \gg 1$.

The (frozen) action S expressed in terms of the canonical coordinates (x,p) is

$$P=\frac{p}{\sqrt{\omega(s)}},\quad X=\sqrt{\omega(s)}x\quad S(x,p,s)=\frac{1}{2}(P^2+X^2)=\frac{1}{2}\left(\frac{p^2}{\omega(s)}+\omega(s)x^2\right)$$

Since S is proportional to $H(x,p)$ its evolution is given by

$$\begin{aligned}\frac{dS}{dt}&=\underbrace{\{H,S\}_{xp}}_{=0}+\frac{\partial S}{\partial t}\\&=\frac{\omega'(s)}{2T}\left(-\frac{p^2}{\omega^2(s)}+x^2\right)\\&=-\frac{\omega'(s)}{2T\omega(s)}\left(P^2-X^2\right)\\&=-\frac{\big(\log\omega(s)\big)'}{2T}\left(P^2-X^2\right)\\&=-\frac{\big(\log\omega(s)\big)'}{T}S(\sin^2\theta-\cos^2\theta)\\&=\frac{S}{T}\big(\log\omega(s)\big)'\cos(2\theta)\end{aligned}$$

Let us see how much $\log S$ changes in one period of the oscillator, from t_n to t_{n+1}. Since ω is slowly varying in one period

$$\Delta t=t_{n+1}-t_n\approx\frac{2\pi}{\omega(t_n/T)}$$

and

$$\begin{aligned}\Delta\log S&=-\int_{t_n}^{t_{n+1}}\frac{d\log\omega(t/T)}{dt}\cos 2\theta(t)dt\\&=-\frac{1}{T}\int_{t_n}^{t_{n+1}}\frac{\omega'(t/T)}{\omega(t/T)}\cos 2\theta(t)dt\\&=-\frac{1}{T}\frac{\omega'(t_n/T)}{\omega(t_n/T)}\int_{t_n}^{t_{n+1}}\cos 2\theta(t)dt+O(\Delta t/T^2)\\&=-\frac{1}{T}\frac{\omega'(t_n/T)}{\omega(t_n/T)}\underbrace{\int_{t_n}^{t_{n+1}}\cos\big(2\omega(t_n/T)t\big)dt}_{=0}+O(\Delta t/T)^2\\&=O(\Delta t/T)^2\end{aligned}$$

This means that even after long time intervals, $O(T)$, when ω changed substantially, S changed little, $O(1/T)$. This is what is meant by saying that S is an adiabatic invariant.

11.12 Lyapunov exponents

Let me look at the Jacobian matrix

$$J_{jk}(t) = \frac{\partial \xi_j(t)}{\partial \xi_k}$$

We know that $\det J = 1$. This means that the product of its eigenvalues is unity. Let us look at its largest eigenvalue

$$\lambda(t) : \textit{Largest eigenvalue of} \, M_{jk}(t)$$

λ says something about the maximal rate at which different directions in phase separate. If points separate at exponential speed we shay that the Hamiltonian is chaotic:

Definition 11.9 *H is (uniformly) chaotic if*

$$\lim_{t \to \infty} \frac{1}{t} \log \lambda(t) > 0$$

Exercise 11.4 *Find λ for the Hamiltonians*

$$H = \frac{\omega}{2}(p^2 - x^2), \quad \tilde{H} = \omega p x$$

Are they chaotic? Explain.

12 Rigid body motions

12.1 Rigid rotations

Rotations are a symmetry of Euclidean space $\mathbb{R}^n$. Rotations preserve preserve the distances between any two points. Rotations are described by orthogonal matrices:

Definition 12.1 *A real $n \times n$ matrix $\mathbf{R}$ is orthogonal if*

$$\mathbf{R}^t \mathbf{R} = \mathbb{1}$$

Equivalently

$$\mathbf{R}^{-1} = \mathbf{R}^t$$

Theorem 12.1 *An orthogonal matrix $\mathbf{R}$ preserves the scalar product of Euclidean vectors*

$$\mathbf{v}^t \mathbf{w} = \mathbf{v} \cdot \mathbf{w} = \mathbf{R}\mathbf{v} \cdot \mathbf{R}\mathbf{w}$$

In particular, it preserves the length of vectors and the distance between any two points in Euclidean space.

Proof:

$$\mathbf{v}^t \mathbf{w} = \mathbf{v}^t \mathbb{1} \mathbf{w} = \mathbf{v}^t \mathbf{R}^t \mathbf{R} \mathbf{w} = (\mathbf{R}\,\mathbf{v})^t \mathbf{R}\mathbf{w}$$

It follows from the definition that

Theorem 12.2 *Orthogonal transformations preserve volumes:*

$$(\det \mathbf{R})^2 = 1 \Longrightarrow \det \mathbf{R} = \pm 1$$

Definition 12.2 *A rotation is an orthogonal transformation* $\mathbf{R}$ *with* $\det \mathbf{R} = 1$*. Orthogonal transformations with* $\det \mathbf{R} = -1$ *are called reflections.*

For example, a rotation by $\pi/2$ around the z axis is

$$R_z(\pi/2) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Theorem 12.3 *The product of two rotations is a rotation.*

Proof:

$$(\mathbf{R}_1\mathbf{R}_2)(\mathbf{R}_1\mathbf{R}_2)^t = \mathbf{R}_1\mathbf{R}_2\mathbf{R}_2^t\mathbf{R}_1^t = \mathbf{R}_1\mathbf{R}_1^t = \mathbb{1}$$

In general, rotations in more than 2 dimensions do not commute. For example,

$$R_z(\pi/2)R_x(\pi/2) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

while

$$R_x(\pi/2)R_z(\pi/2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}$$

A rotation in 3 dimensions leaves the axis of rotation in place. Hence

Theorem 12.4 *The axis of a rotations* $\mathbf{R}$ *in three dimensions is given by the direction of the eigenvector* $\mathbf{n}$ *with unit eigenvalue*

$$\mathbf{R}\,\mathbf{n} = \mathbf{n}$$

Choosing the z-axis to be $\mathbf{n}$ *the rotation matrix is*

$$\mathbf{R} = \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{12.1}$$

The angle of rotation ψ *is*

$$Tr\,\mathbf{R} = 1 + 2\cos\psi \tag{12.2}$$

It follows from this that

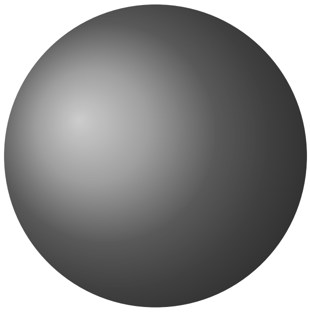


Figure 50: A rotation can be specified by three angles: The two standard spherical angles $0 \leq \theta \leq \pi$ and $0 \leq \phi < 2\pi$ specify the direction of the axis of rotation, and a third angle $0 \leq \psi < \pi$ gives the angle of rotation. The identity $\mathbb{1}$ is at the center of the ball. Since a rotation by 2π is no rotation a rotation by π about $\mathbf{n}$ is the same as a rotation by $-\pi$ about $-\mathbf{n}$. This means that antipodal points on the surface with radius π on the ball are the same rotation: $(\theta, \phi, \pi) \cong (\pi - \theta, \phi \pm \pi, \pi)$

Theorem 12.5 *If $\mathbf{R}_1$ and $\mathbf{R}_2$ are rotations then the angle of rotation ϕ of*

$$\mathbf{R} = \mathbf{R}_1 \mathbf{R}_2 \mathbf{R}_1^t$$

is the same as the angle of rotation ψ_2 of $\mathbf{R}_2$ and the axis of rotations are related by $\mathbf{R}_1$.

Proof:

$$Tr\,\mathbf{R} = Tr(\mathbf{R}_1 \mathbf{R}_2 \mathbf{R}_1^t) = Tr(\mathbf{R}_1^t \mathbf{R}_1 \mathbf{R}_2) = Tr\,\mathbf{R}_2$$

Theorem 12.6 *The group of rotations in 3 dimensions can be identified with a ball of radius π with antipodal points on the surface identified.*

Proof: Explained in the figure.

12.2 Infinitesimal rotations and angular velocity

A special feature of 3 dimensions is that there is a 1-1 correspondence between vectors and anti-symmetric matrices.

Definition 12.3 *In 3 dimensions a correspondence between a vector $\mathbf{a}$ and an antisymmetric matrix $\mathbf{A}$ is given by (summation over repeated indexes implied)*

$$A_{ij} = -\varepsilon_{ijk} a_k \Leftrightarrow a_k = -\frac{1}{2} \varepsilon_{ijk} A_{ij}$$

and explicitly

$$\mathbf{A} = -\begin{pmatrix} 0 & a_3 & -a_2 \\ -a_3 & 0 & a_1 \\ a_2 & -a_1 & 0 \end{pmatrix} \tag{12.3}$$

The correspondence clearly respects the the rules of vector addition and multiplication by scalars. Moreover one has

Theorem 12.7

$$\mathbf{A}\mathbf{x} = \mathbf{a} \times \mathbf{x}$$

Proof:

$$(\mathbf{A}\mathbf{x})_i = -A_{ij}x_j = -\varepsilon_{ijk}a_k x_j = \varepsilon_{ijk}a_j x_k = (\mathbf{a} \times \mathbf{x})_i$$

One consequence of this is that we can identify infinitesimal rotations with vectors.

Theorem 12.8 *An infinitesimal rotation in three dimensions has the form*

$$\mathbf{R} = \mathbb{1} + \mathbf{N}\delta\psi - \frac{1}{2}\mathbf{N}\mathbf{N}^t(\delta\psi)^2 + O\big((\delta\psi)^3\big)$$

where the anti-symmetric matrix, $\mathbf{N}\delta\psi$ *corresponds to the unit vector vector* $\mathbf{n}$ *of the axis and* $\delta\psi$ *the infinitesimal angle.*

Proof: Substitute

$$\mathbf{R} = \mathbb{1} + \mathbf{N}\delta\psi - \frac{1}{2}\mathbf{N}\mathbf{N}^t(\delta\psi)^2$$

in

$$\mathbb{1} = \mathbf{R}\mathbf{R}^t$$

gives to second order

$$\big(\mathbf{N} + \mathbf{N}^t\big)\delta\psi + \left(\mathbf{N}\mathbf{N}^t - \frac{2}{2}\mathbf{N}\mathbf{N}^t\right)(\delta\psi)^2 = O((\delta\psi)^3)$$

Each term has to vanish separately. The first term imposes the condition that $\mathbf{N}$ is anti-symmetric and the second is automatic. $\mathbf{n}$ is the rotation axis since

$$\mathbf{R}\,\mathbf{n} = \mathbf{n} + \mathbf{N}\,\mathbf{n}\delta\psi = \mathbf{n} + \mathbf{n} \times \mathbf{n}\delta\psi = \mathbf{n}$$

as it must. The angle of rotation ψ is small and can be computed from the trace

$$Tr\,\mathbf{R} = 3 - \frac{1}{2}Tr\,(\mathbf{N}\mathbf{N}^t)\,(\delta\psi)^2 + O((\delta\psi)^3) = 3 - (\delta\psi)^2 + O((\delta\psi)^3)$$

where I used the fact that $\mathbf{n}$ is a unit vector:

$$Tr\,(\mathbf{N}\mathbf{N}^t) = 2\,\mathbf{n}\cdot\mathbf{n} = 2$$

It follows from Eq. (12.2) that the angle of rotation is $\delta\psi$.

Suppose that $\mathbf{R}(t)$ is time dependent. How shall we define the angular velocity? Fix a vector $\mathbf{x}$, the vector will then rotate by

$$\mathbf{x}(t) = \mathbf{R}(t)\mathbf{x}$$

and its velocity is then

$$\mathbf{v}(t) = \dot{\mathbf{x}}(t) = \dot{\mathbf{R}}(t)\mathbf{x} = \big(\dot{\mathbf{R}}(t)\mathbf{R}^t(t)\big)\,\mathbf{x}(t) = \boldsymbol{\Omega}(t)\mathbf{x}$$

The instantaneous position and velocity are linearly related by the matrix $\boldsymbol{\Omega}(t)$.

Theorem 12.9 *The matrix $\mathbf{\Omega}(t)$ is anti-symmetric*

Proof:

$$\mathbf{\Omega}^t(t) = \left(\dot{\mathbf{R}}(t)\mathbf{R}^t(t)\right)^t = \mathbf{R}(t)\dot{\mathbf{R}}^t(t) = -\dot{\mathbf{R}}(t)\mathbf{R}^t(t) = -\mathbf{\Omega}(t)$$

We can now use the 1-1 correspondence between 3×3 anti-symmetric matrices and vectors to associate with $\mathbf{\Omega}$ the 3 vector $\boldsymbol{\omega}$ recalling that the action of the matrix is equivalent to cross product of the vector:

$$\mathbf{\Omega}\,\mathbf{x} = \boldsymbol{\omega} \times \mathbf{x}$$

Definition 12.4 *The vector $\boldsymbol{\omega}(t)$ associated to the anti-symmetric matrix $\mathbf{\Omega}(t) = \dot{\mathbf{R}}(t)\mathbf{R}^t(t)$ according to Eq. (12.3), is the angular velocity of the rotation $R(t)$.*

12.3 Tensor of inertia

To construct the Lagrangian we can use any coordinate we please to express the kinetic energy in an inertial frame. For the sake of simplicity, let us consider a rigid body whose kinetic energy is pure rotation. This is the case if there is a point, call it the origin $\mathbf{x} = 0$, which is at rest and the body is rotating rigidly about this point. For example, in free space (no external forces) choose the inertial frame so that the center of mass is at rest and take the origin of the the inertial Euclidean frame so that the origin is at the center of mass. Then the entire kinetic energy is the rotational kinetic energy.

Lets think if $\mathbf{v}, \mathbf{x}$ and $\boldsymbol{\omega}$ as column vectors in an inertial frame. The kinetic energy is related to the scalar

$$\begin{aligned} \mathbf{v}\cdot\mathbf{v} &= (\boldsymbol{\omega}\cdot\boldsymbol{\omega})(\mathbf{x}\cdot\mathbf{x}) - (\omega\cdot\mathbf{x})^2 \\ &= \boldsymbol{\omega}^t\,(\mathbf{x}\cdot\mathbf{x})\mathbb{1}\,\boldsymbol{\omega} - (\boldsymbol{\omega}^t\mathbf{x})\,(\mathbf{x}^t\,\boldsymbol{\omega}) \\ &= \boldsymbol{\omega}^t\left((\mathbf{x}\cdot\mathbf{x})\mathbb{1} - \mathbf{x}\,\mathbf{x}^t\right)\boldsymbol{\omega} \end{aligned} \tag{12.4}$$

This is quadratic in $\boldsymbol{\omega}$ and the middle brackets is a 3×3 matrix. The kinetic energy of a mass m located at $\mathbf{x}$ is then

$$T = \frac{m}{2}\mathbf{v}\cdot\mathbf{v} = \frac{1}{2}\boldsymbol{\omega}^t\,\mathbf{I}\,\boldsymbol{\omega}$$

with $\mathbf{I}$ a 3×3 symmetric matrix

$$\mathbf{I} = m\Big((\mathbf{x}\cdot\mathbf{x})\mathbb{1} - \mathbf{x}\,\mathbf{x}^t\Big), \qquad I_{jk} = m(\mathbf{x}\cdot\mathbf{x}\delta_{jk} - x_j x_k)$$

Since the kinetic energy of a collection of masses m_a is additive and $\boldsymbol{\omega}$ is common to all the masses moving like a rigid body we have:

Theorem 12.10 *The kinetic energy of a rigidly rotation body with* $\mathbf{x} = 0$ *at rest in an inertial frame and* $\boldsymbol{\omega}$ *the instantaneous angular velocity of the body in an inertial frame is*

$$T = \frac{1}{2}\boldsymbol{\omega}^t \mathbf{I}\boldsymbol{\omega}$$

where $\mathbf{I}$*, the tensor of inertia, is the* 3×3 *symmetric matrix*

$$\mathbf{I} = \sum_a m_a \Big((\mathbf{x_a} \cdot \mathbf{x_a})\mathbb{1} - \mathbf{x_a}\,\mathbf{x_a}^t \Big) = \sum_a m_a \begin{pmatrix} y_a^2 + z_a^2 & -x_a y_a & -x_a z_a \\ -x_a y_a & x_a^2 + z_a^2 & -y_a z_a \\ -z_a x_a & -z_a y_a & x_a^2 + y_a^2 \end{pmatrix}$$

in the inertial frame.

Note that for a rotating body $\mathbf{I}$ and $\boldsymbol{\omega}$ would, in general, be functions of time t.

By the standard definition of conjugate momentum in Lagrangian mechanics the angular momentum is

$$\mathbf{J} = \frac{\partial T}{\partial \boldsymbol{\omega}} = \mathbf{I}\,\boldsymbol{\omega}$$

Exercise 12.1 *Show that the angular momentum* $\mathbf{J}$ *defined above agrees with the Newtonian definition.*

Unless $\mathbf{I} \propto \mathbb{1}$, the angular momentum and the angular velocity are not parallel: $\mathbf{J} \not\parallel \Omega$.

Exercise 12.2 *Let* $\mathbf{I}_{cm}$ *be the inertia tensor relative to the center of mass. Show that* $\mathbf{I}_a$ *relative to point* $\mathbf{b}$ *is*

$$\mathbf{I}_a = \mathbf{I}_{cm} + M \begin{pmatrix} b_y^2 + b_z^2 & -b_x b_y & -b_x b_z \\ -b_y b_x & b_x^2 + b_z^2 & -b_z b_y \\ -b_z b_x & -b_z b_y & b_x^2 + b_y^2 \end{pmatrix}$$

We picked the origin of the inertial Euclidean frame assuming that $\mathbf{x} = 0$ is at rest. This still leaves the freedom to choose the directions of the three axis. Let $\mathbf{R}(t)$ denote a time dependent rotation to a time dependent frame. The angular velocity in the inertial frame is represented by the vector $\boldsymbol{\omega}$. The same vector has in the moving frame different components

$$\boldsymbol{\omega}' = \mathbf{R}(t)\boldsymbol{\omega}$$

I want to write the kinetic energy in the inertial frame in terms of $\boldsymbol{\omega}'$ and $\mathbf{I}'$ in the rotating frame. The transformation rules are then

$$\boldsymbol{\omega}^t\,\mathbf{I}\,\boldsymbol{\omega} = \boldsymbol{\omega}'^t\,\mathbf{I}'\boldsymbol{\omega}' = (\mathbf{R}(t)\boldsymbol{\omega})^t\,\mathbf{I}'(\mathbf{R}(t)\boldsymbol{\omega}) = \boldsymbol{\omega}^t(\mathbf{R}^t(t)\mathbf{I}'\mathbf{R}(t)R)\boldsymbol{\omega}$$

From this we conclude

Theorem 12.11 *Under time dependent rotation* $\mathbf{R}(t)$ *the vector* $\boldsymbol{\omega}$ *and the second rank tensor* $\mathbf{I}$ *of the inertial frame transform to the body frame by*

$$\boldsymbol{\omega}' = \mathbf{R}(t)\boldsymbol{\omega}, \quad \mathbf{I}' = \mathbf{R}(t)\mathbf{I}\mathbf{R}(t)^t \tag{12.5}$$

and then the kinetic energy in an inertial frame can be written in terms of the body frame quantities. In particular, we can make $\mathbf{I}'$ *time independent provided we write the angular velocity* $\boldsymbol{\omega}'$ *in the body frame.*

Exercise 12.3 *Show that* $\mathbf{I}$ *is a non-negative matrix.*

12.4 Principal axis

Since $\mathbf{I}$ is a real symmetric matrix there is an orthogonal transformation (=a rotation) that bring it to a diagonal form. This frame is called the frame of principal axes. The inertia tensor is diagonal matrix in this frame with entries

$$\mathbf{I}_{jk} = \delta_{jk} I_j, \quad I_j \geq 0$$

I_j are fixed positive constants in the body frame. Positivity follows from the positivity of the kinetic energy. They are an analog of the mass except that the mass is a scalar while $\mathbf{I}$ is a tensor. Often you can guess the principal frame by symmetry consideration. The principal frame is unique if $I_1 \neq I_2 \neq I_3$. If $I_1 = I_2$ only the x_3 axis is uniquely determined. If $I_1 = I_2 = I_3$ then any frame is principal.

Theorem 12.12 *The rotation vector* $\boldsymbol{\omega}$ *is parallel to the angular momentum* $\mathbf{J}$ *if oriented along the principal axis.*

We may compute the kinetic energy in any frame we wish, provided it is an inertial frame. We are allowed to take the inertial frame to be so that it agrees with the instantaneous frame of principle axes

$$T = \frac{1}{2}\sum_{j=1}^{3} I_j \omega_j^2$$

$\boldsymbol{\omega}$ is, of course, the angular velocity in this inertial frame.

Definition 12.5 *A top is called symmetric if* $I_1 = I_2$*.*

Theorem 12.13 *If a body has a* π/n *symmetry of rotation about the* 3 *axis, with* $n \geq 2$ *then* $I_1 = I_2$*.*

Proof: For a rotation by angle α about the 3 axis the 11 component of Eq. 12.5 gives

$$\cos^2\alpha I_1 + \sin^2\alpha I_2 = I_1$$

The result follows if $\sin\alpha \neq 0$.

Theorem 12.14 *If a body has a π/n symmetry of rotation about two perpendicular axes, $n \geq 2$ then $I_1 = I_2 = I_3$.*

It follows that all the Platonic solids, the tetrahedron, cube,octahedron, and dodecahedron and icosahedron and the sphere have as tensor of inertia about their center which is proportional to $\mathbb{1}$.

Example 12.1 *One dimensional objects*

$$I_1 = I_2 \qquad I_3 = 0$$

I shall call the principal frame also the body frame. I is a fixed matrix in the frame of the body

$$\mathbf{x} = (x_1, x_2, x_3)$$

12.5 Euler angles

Euler angles demo (and another demo.)

To write a Lagrangian for a rotating rigid body we need coordinates. Three coordinates tell us how two frames with the same origin are related

$$\underbrace{(X, Y, Z)}_{lab} \Leftrightarrow \underbrace{(x_1, x_2, x_3)}_{body}$$

Different sources define Euler angles differently and denote them by different letters. We pick

$$(\theta, \phi, \psi)$$

$$\cos\theta = \hat{\mathbf{Z}} \cdot \hat{\mathbf{x}}_3, \qquad 0 \leq \theta \leq \pi$$

Line of nodes: The intersection of the plane $Z = 0$ with the plane $x_3 = 0$. It is indeed line provided $\theta \neq 0, \pi$ and a plane in these exceptional cases. We can associate a unit vector $\mathbf{n}$ with the line of nodes so that is points as right handed screw does when the Z axis is rotated towards towards x_3.

We can now define the two angles $0 \leq \phi, \psi \leq 2\pi$ by

$$\cos\phi = \mathbf{n} \cdot \mathbf{X}, \quad \cos\psi = \mathbf{n} \cdot \mathbf{x}_1,$$

and orientation fixed in the figure.

The Euler angles have a coordinate singularity at $\theta = 0, \pi$ where the line of nodes is not defined. As we shall see this means that where $\theta = 0$ the assignment of ϕ and ψ is not unique–only their sum matters:

$$(0, \phi, \psi) \cong (0, \phi + \psi, 0) \cong (0, 0, \phi + \psi)$$

This is analogous to the coordinate singularity in spherical coordinates where when $\theta = 0$ all values ϕ give the same point in the sphere.

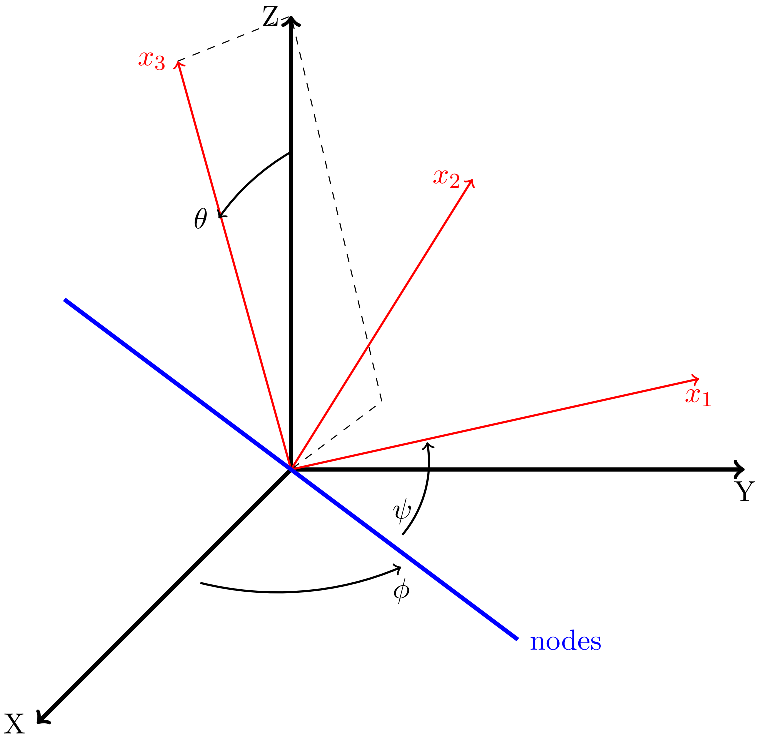


Figure 51: Two orthogonal frames, with a common origin, one black and one red. Euler angles specify how two frames are related. The blue line, the line of nodes, is the intersection of the $X - Y$ plane with the $x_1 - x_2$ plane.

12.6 Euler angles: Lab frame and the body frame

The lab frame (X, Y, Z) is related to the body frame (x_1, x_2, x_3) by three consecutive rotations[11]:

$$\underbrace{R_Z(\phi)}_{lab\ Z} \mapsto \underbrace{R_N(\theta)}_{Nodes} \mapsto \underbrace{R_3(\psi)}_{body\ 3} \Longleftrightarrow R_3(\psi)R_N(\theta)R_Z(\phi) \tag{12.6}$$

Since R_Z is the first rotation, it stands on the right on the rhs: It hits the vector $\mathbf{x}$ first.

We want to translate the rotations about the (moving) axes x_3 and the line of nodes N, to the rotations about the fixed lab frame:

Theorem 12.15

$$R_3(\psi)R_N(\theta)R_Z(\phi) = R_Z(\phi)R_X(\theta)R_Z(\psi) \tag{12.7}$$

Note that on the left we rotated about 3 different axis. On the right we rotate twice about Z. The order has been flipped.

[11] Here we find the active rotation relating the two frames. Goldstein studies passive rotations, so the formulas are different.

Proof:

$$R_X(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \tag{12.8}$$

is related to $R_N(\theta)$ by conjugation. The conjugation is by $R_Z(\phi)$ since

$$\hat{\mathbf{N}} = R_Z(\phi)\hat{\mathbf{X}}$$

Hence:

$$R_N(\theta) = R_Z(\phi) R_X(\theta) R^t_Z(\phi)$$

Similarly, since

$$\hat{\mathbf{x}}_{\mathbf{3}} = R_N(\theta)\hat{\mathbf{Z}}$$

$R_3(\psi)$ is related to the lab rotation $R_Z(\psi)$ by conjugation by $R_N(\theta)$:

$$R_3(\psi) = R_N(\theta) R_Z(\psi)\, R^t_N(\theta)$$

Hence

$$\begin{aligned} R_3(\psi) R_N(\theta) R_Z(\phi) &= \Big(R_N(\theta) R_Z(\psi)\, R^t_N(\theta)\Big)\Big(R_N(\theta)\Big) R_Z(\phi) \\ &= R_N(\theta) R_Z(\psi) R_Z(\phi) \\ &= \Big(R_Z(\phi) R_X(\theta) R^t_Z(\phi)\Big)\, R_Z(\psi) R_Z(\phi) \end{aligned} \tag{12.9}$$

But, rotation about the same axis commute and so

$$R^t_Z(\phi)\, R_Z(\psi) R_Z(\phi) = R_Z(\psi)\, \underbrace{R^t_Z(\phi) R_Z(\phi)}_{=\mathbb{1}} = R_Z(\psi)$$

We have then shown that

$$R_3(\psi) R_N(\theta) R_Z(\phi) = R_Z(\phi) R_X(\theta) R_Z(\psi) \tag{12.10}$$

Exercise 12.4 *Find the transformation from the body frame to the lab in terms of R_1 and R_3.*

12.7 ω and the velocities of Euler angles

Now think of the body frame rotating in the inertial lab frame. Clearly Ω depends linearly on the angular velocities $(\dot\theta, \dot\phi, \dot\psi)$.

- $\dot\theta$: angular velocity about the line of nodes
- $\dot\phi$: Rate of precession of the line of nodes about $\hat{\mathbf{Z}}$ (definition of precession)

Since the direction of the three axis, line of nodes, the x_3 and Z axis are not orthogonal, the relation between Ω and $(\dot\theta, \dot\phi, \dot\psi)$ involves geometric factors that we need to figure out.

Theorem 12.16

$$\begin{aligned}\omega_3 &= \dot\psi + \dot\phi\cos\theta \\ \omega_1 &= \dot\theta\cos\psi + \dot\phi\sin\theta\sin\psi \\ \omega_2 &= -\dot\theta\sin\psi + \dot\phi\sin\theta\cos\psi\end{aligned}$$

Proof: Since the dependence of $\mathbf{\Omega}$ on the velocities of the Euler angles in linear, we can figure out the contribution of $\dot\psi$ to $\mathbf{\Omega}$ when $\dot\phi = \dot\theta = 0$ etc.

Consider $\dot\psi \neq 0$ while $\dot\theta = \dot\phi = 0$. Clearly

$$\boldsymbol{\omega} = \dot\psi\mathbf{x}_3$$

This proves the red term in the equation. Next, consider $\dot\theta \neq 0$ while $\dot\psi = \dot\phi = 0$. In this case

$$\boldsymbol{\omega} = \dot\theta\mathbf{n}$$

It remains to express $\mathbf{n}$ in the body frame. Since $\mathbf{n}\cdot\mathbf{x}_3 = 0$ we have

$$\mathbf{n} = \hat{\mathbf{x}}_1\cos\psi - \hat{\mathbf{x}}_2\sin\psi$$

This proves the blue parts of the equation. It remains to consider the case $\dot\phi \neq 0$. The corresponding angular velocity is

$$\boldsymbol{\omega} = \dot\phi\,\hat{\mathbf{z}}$$

It therefor remains to express $\mathbf{z}$ in the body frame:

$$\mathbf{z} = \hat{\mathbf{x}}_3\cos\theta + \sin\theta(\hat{\mathbf{x}}_1\sin\psi + \hat{\mathbf{x}}_2\cos\psi)$$

This concludes the proof.

12.8 Free symmetric top

Theorem 12.17 *The kinetic energy of a symmetric top is*

$$T(\dot\theta, \dot\phi, \dot\psi, \theta) = \frac{I_1}{2}\left(\dot\theta^2 + \dot\phi^2\sin^2\theta\right) + \frac{I_3}{2}(\dot\psi + \dot\phi\cos\theta)^2 \tag{12.11}$$

We are allowed to use the body axis, provided we use $\boldsymbol{\omega}$ as seen in an inertial frame. From theorem 12.16 we get

$$\omega_1^2 + \omega_2^2 = (\dot\theta\cos\psi + \dot\phi\sin\theta\sin\psi)^2 + (-\dot\theta\sin\psi + \dot\phi\sin\theta\cos\psi)^2 = \dot\theta^2 + \dot\phi^2\sin^2\theta$$

The three conjugate momenta are

$$p_\theta = I_1\dot\theta, \quad p_\psi = I_3(\dot\psi + \dot\phi\cos\theta), \quad p_\phi = I_1\dot\phi\sin^2\theta + I_3(\dot\psi + \dot\phi\cos\theta)\cos\theta$$

For a free symmetric top thee kinetic energy is also the Lagrangian. Since T is a functions of all velocities, but only of one coordinate θ the two coordinates ψ and ϕ are cyclic coordinates,

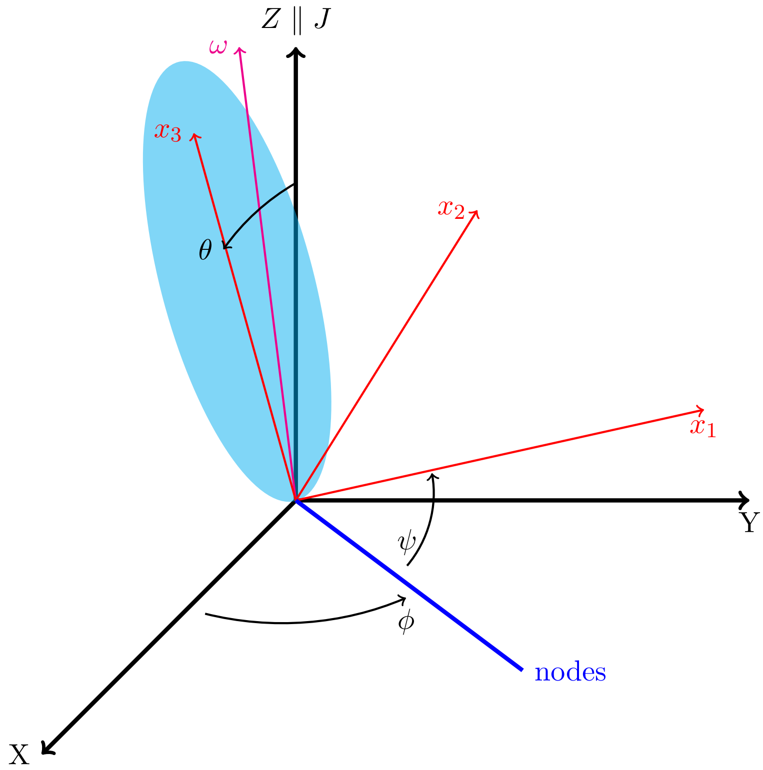


Figure 52: Symmetric top with **J** along the lab Z axis. ω x_3 and J are coplanar.

Theorem 12.18 *The free symmetric top:*

- *Has three constants of motion: p_ψ, p_ϕ and the energy.*
- *The energy is T.*
- *The pairs (p_ψ, ψ) and (p_ϕ, ϕ) are action angle coordinates.*
- *$p_\psi = J_3$ is the angular momentum about the body axis*

$$p_\psi = I_3(\dot{\psi} + \dot{\phi}\cos\theta) = I_3\omega_3 \tag{12.12}$$

- *$p_\phi = J_Z$ is the angular momentum about Z in the lab.*

$$p_\phi = I_1\dot{\phi}\sin^2\theta + \underbrace{I_3(\dot{\psi} + \dot{\phi}\cos\theta)}_{p_\psi}\cos\theta \tag{12.13}$$

- *The Hamiltonian of a free symmetric top is*

$$H = \frac{1}{2I_1}p_\theta^2 + \frac{1}{2I_1}\frac{(p_\phi - p_\psi\cos\theta)^2}{\sin^2\theta} + \frac{{p_\psi}^2}{2I_3}$$

The middle term can be viewed as the effective potential for the one dimenssional motion of the coordinate θ

- *The symmetric top is completely integrable by the methods of action angles.*

Proof:

- The three conservation laws come from Nöther.
- The energy is defined, by Nöther, by

$$H = p_\theta \dot{\theta} + p_\psi \dot{\psi} + p_\phi \dot{\phi} - T = T$$

  We actually did not need to compute that $H = T$ since we have shown, in general, that this is the case if the kinetic energy is homogeneous and quadratic in the velocities.

- The action angle statement follows from the conservation laws and dimension analysis

$$[p_\psi] = [p_\phi] = [action]$$

- This follows from Theorem 12.16
- Follows from Nöther
- Expressing the energy in terms of the momenta gives H
- Follows from the three conservation laws

The effective potential that protects the top away from $\sin\theta = 0$. This reduces the problem of a free symmetric top to a one dimensional Hamiltonian system in a potential.

The potential diverges when $\sin\theta = 0$ and takes its minimum at

$$\cos\theta = \begin{cases} \frac{p_\phi}{p_\psi} & |p_\phi| \le |p_\psi| \\ \frac{p_\psi}{p_\phi} & |p_\psi| \le |p_\phi| \end{cases}$$

In general, θ will oscillate about the minimum and then the precession rate $\dot{\phi}$ will not be a constant. However, as we shall see in the next section, by choosing the lab frame wisely, we can get rid of the motion in θ and the precession rate will then turn out to be a constant.

12.9 Choosing a good lab frame: J parallel to the Z axis

Let us start with a simple geometric property of the motion of symmetric top:

Theorem 12.19 *The three vectors ω, $\mathbf{J}$ and $\hat{\mathbf{x}}_3$ are coplanar*

That the vectors are coplanar follows from a computation of the volume in the body frame

$$\hat{\mathbf{x}}_3 \times \mathbf{J} \cdot \omega = \begin{vmatrix} \omega_1 & \omega_2 & \omega_3 \\ I_1\omega_1 & I_1\omega_2 & I_3\omega_3 \\ 0 & 0 & 1 \end{vmatrix} = I_1 \begin{vmatrix} \omega_1 & \omega_2 \\ \omega_1 & \omega_2 \end{vmatrix} = 0$$

We have chosen the body coordinates to be the principal frame. So far the Lab frame orientation was arbitrary. For a free top the total angular momentum $\mathbf{J}$ is conserved. We may then choose the lab fame so that $\mathbf{J} \parallel \hat{\mathbf{Z}}$. Since $\hat{\mathbf{Z}}$ and $\mathbf{x}_3$ are both perpendicular to the line of nodes we have

Theorem 12.20 *If the lab frame is oriented with $\mathbf{J} \parallel \mathbf{Z}$ then the three vectors ω, $\mathbf{J}$ and $\hat{\mathbf{x}}_3$ are all perpendicular to the line of nodes. Moreover, the angle θ is a constant of motion:*

$$\cos\theta = \frac{J_3}{|\mathbf{J}|} \tag{12.14}$$

Proof:

$$|\mathbf{J}|\cos\theta = J_3 = I_3\omega_3 = p_\psi$$

Since both $|\mathbf{J}|$ and p_ψ are conserved, the angle θ is constant.

Theorem 12.21 *If the lab frame is oriented with $\mathbf{J} \parallel \mathbf{Z}$ the precession rate is constant*

$$\dot\phi = \frac{|\mathbf{J}|}{I_2}$$

Proof;: Since θ is a constant, it follows from the constancy of p_ψ and p_ϕ and Eq. 12.13 that the precession rate $\dot\phi$ is a constant.

Since the precession is constant we can compute it in any time we wish. Let us choose the time so when the lab and top coordinates are especially simply related. This is the case when $\psi = 0$ so $\hat{\mathbf{x}}_1$ lies in the $X - Y$ plane and is orthogonal to Z. $\mathbf{J}$ lies in the $x_2 - x_3$ plane and

$$|\mathbf{J}|\sin\theta = \mathbf{J}\cdot\hat{\mathbf{x}}_2 = J_2 = I_2\omega_2 = I_2\dot\phi\sin\theta, \quad (\psi = 0)$$

and in the rightmost identity we used $\psi = 0$ in Theorem (12.16). This completes the proof.

Theorem 12.22 *The rotation rate about the top axis is*

$$\dot\psi = \left(1 - \frac{I_3}{I_1}\right)\omega_3$$

Proof:

$$J\cos\theta = J_3 = I_3\omega_3 = I_3(\dot\psi + \dot\phi\cos\theta)$$

Substituting the precession rate

$$\dot\psi = \left(\frac{J}{I_3} - \dot\phi\right)\cos\theta = \left(\frac{J}{I_3} - \frac{J}{I_1}\right)\cos\theta = J\left(\frac{1}{I_3} - \frac{1}{I_1}\right)\frac{J_3}{J} = \left(\frac{1}{I_3} - \frac{1}{I_1}\right)J_3$$

Animation.

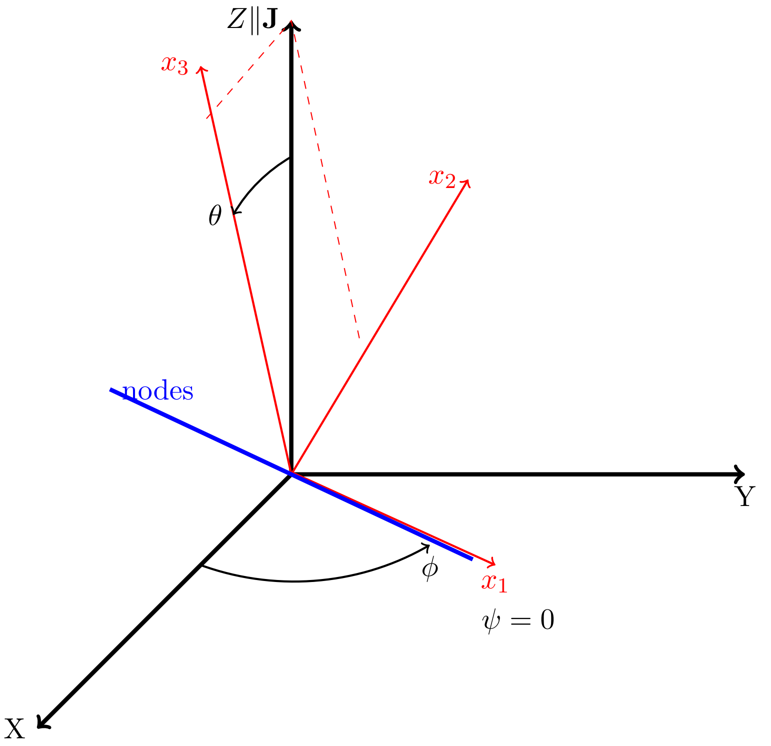


Figure 53: $\mathbf{J}$ points along the lab Z axis. When $\psi = 0$ the line of nodes is along the x_1 axis and total angular momentum $\mathbf{J}$ lies in the $x_3 - x_2$ plane.

12.10 Symmetric top in a gravitational field

Symmetric top in a gravitational field is also an integrable problem. Since the gravitational filed applies a torque, $\mathbf{J}$ is not conserved and we need to choose the lab frame wisely again. It is now convenient to choose the direction of $\mathbf{g}$ as the lab $\mathbf{Z}$ axis. The Lagrangian in this frame is

$$L(\dot{\theta}, \dot{\phi}, \dot{\psi}, \theta) = T(\dot{\theta}, \dot{\phi}, \dot{\psi}, \theta) - mg\ell \cos\theta$$

where T is the kinetic energy. Since ψ and ϕ are cyclic p_ψ and p_ϕ, defined as before,

$$p_\psi = I_3(\dot{\psi} + \dot{\phi}\cos\theta) = I_3\omega_3 \tag{12.15}$$

$$p_\phi = I_1\dot{\phi}\sin^2\theta + \underbrace{I_3(\dot{\psi} + \dot{\phi}\cos\theta)}_{p_\psi}\cos\theta \tag{12.16}$$

are constants of motion, and so is the energy

$$H = \frac{1}{2I_1}p_\theta^2 + \underbrace{\frac{1}{2I_1}\frac{(p_\phi - p_\psi\cos\theta)^2}{\sin^2\theta} + mg\ell\cos\theta}_{effective\ potential} + \underbrace{\frac{p_\psi^2}{2I_3}}_{const}$$

The effective potential looks like the figure: This means that the θ oscillated between two limiting values. This motion is called nutation.

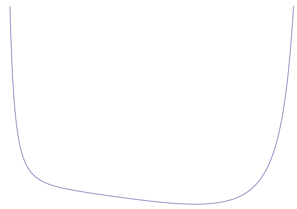


Figure 54: The effective potential prevents θ from reaching $\theta = 0$ and $\theta = \pi$. It limits the nutation.

Example 12.2 *Consider the bicycle gyro. In this case we are interested in a solution where $\theta \approx \pi/2$. Expanding the effective potential near $\pi/2$ gives*

$$V_{effective} = \frac{p_\phi^2}{2I_1} + \left(-mg\ell + \frac{p_\psi p_\phi}{I_1}\right)\left(\theta - \frac{\pi}{2}\right) + \frac{p_\psi^2 + p_\phi^2}{2I_1}\left(\theta - \frac{\pi}{2}\right)^2 + \dots$$

In particular, $\theta = \pi/2$ is a stable equilibrium if the brackets in the middle vanish

$$mg\ell = \frac{p_\psi p_\phi}{I_1} = I_3 \dot{\psi}\dot{\phi}$$

and I have used $\cos\theta = 0$ in p_ψ and p_ϕ. The faster the top spins about its axis, the slower the precession rate.

Example 12.3 *let us look at the stability of rotations about the vertical axis: $\theta = 0$*

1. *The effective potential near $\theta \approx 0$ is (up to a constant term)*

$$V_{effective} = \underbrace{\frac{(p_\psi - p_\phi)^2}{2I_1}}_{A} \frac{1}{\theta^2} + \underbrace{\frac{4(p_\psi + p_\phi)^2 - p_\psi p_\phi - 60 I_1 mg\ell}{120 I_1}}_{B} \theta^2 + O(\theta^3)$$

2. *$\theta = 0$ is an equilibrium point if $p_\psi = p_\phi$ and then*

$$\begin{aligned} V_{effective} &= \frac{4(p_\psi + p_\phi)^2 - p_\psi p_\phi - 60 I_1 mg\ell}{120 I_1} \theta^2 + O(\theta^3) \\ &= \frac{p_\psi^2 - 4 I_1 mg\ell}{8 I_1} \theta^2 + O(\theta^3) \end{aligned}$$

   The equilibrium point is stable if it spins fast enough:

$$p_\psi^2 > 4 I_1 mg\ell \tag{12.17}$$

   For a top of size of 1 [cm], $mg\ell/2I_1 = O(1000)[sec^{-2}]$, $\dot{\psi} = O(5)[Hz]$.

3. *Setting $p_\psi = p_\phi$ requires infinite precision and is unphysical. If $p_\psi \approx p_\phi$ and $B > 0$ the effective potential has a minimum near $\theta = 0$ and more precisely at*

$$\theta = \left(\frac{A}{B}\right)^{1/4} \approx \left(\frac{8(p_\psi - p_\phi)^2}{p_\psi^2 - 4I_1 mg\ell}\right)^{1/4}$$

This is the case if the top is fast enough in the sense of Eq. **??**. *If however, the top is lazy, $B < 0$ and the minimum of the potential is not near $\theta \approx 0$. This what you see in Hanukka tops ans is illustrated in the figure.*

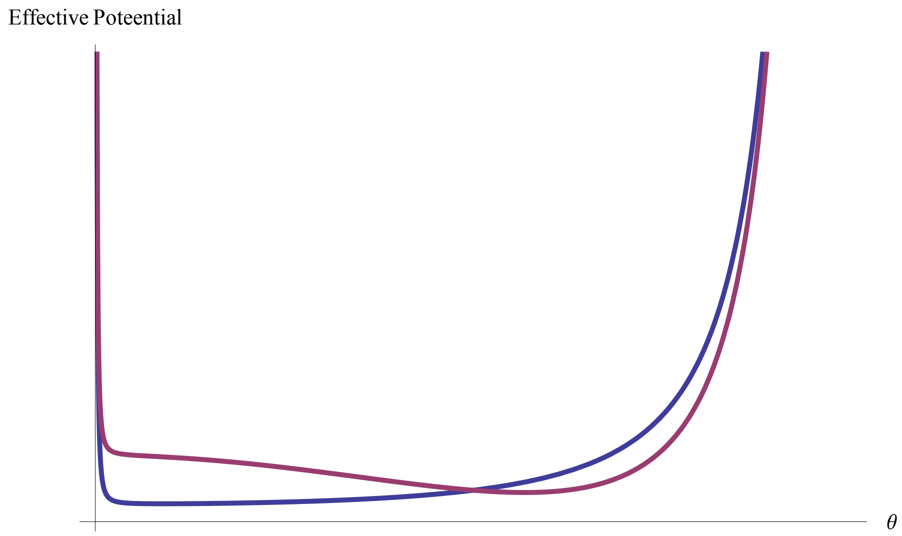


Figure 55: The effective potential for $p_\psi \approx p_\phi$ in the case, $B > 0$, where $\theta \approx 0$ is stable and in the case $B < 0$ where it is unstable.

13 Euler equations

So far we only talked about the Lagrangian formulation of rotation rigid bodies. There is also an interesting formula that comes directly from Newton's law for rotating bodies

$$\dot{\mathbf{J}} = \mathbf{N}$$

Theorem 13.1 *Euler equations*

$$\begin{aligned} N_1 &= I_1\dot\omega_1 - \omega_2\omega_3(I_2 - I_3) \\ N_2 &= I_2\dot\omega_2 - \omega_3\omega_1(I_3 - I_1) \\ N_3 &= I_3\dot\omega_3 - \omega_1\omega_2(I_1 - I_2) \end{aligned} \tag{13.1}$$

For a rigid body

$$\mathbf{J} = I_1\omega_1\hat{\mathbf{x}}_1 + I_2\omega_2\hat{\mathbf{x}}_2 + I_3\omega_3\hat{\mathbf{x}}_3 = \sum_{j=1}^{3} I_j\omega_j\hat{\mathbf{x}}_j$$

The vectors $\hat{\mathbf{x}}_j$ are fixed in the rotating body. Evidently

$$\frac{d\hat{\mathbf{x}}_j}{dt} = \boldsymbol{\omega} \times \hat{\mathbf{x}}_j$$

and so

$$\frac{d\mathbf{J}}{dt} = \sum I_j\left(\dot{\omega}_j\hat{\mathbf{x}}_j + \omega_j\boldsymbol{\omega} \times \hat{\mathbf{x}}_j\right)$$

Hence, in the body frame,

$$\begin{aligned}
N_k &= \mathbf{N} \cdot \hat{\mathbf{x}}_k \\
&= \frac{d\mathbf{J}}{dt} \cdot \hat{\mathbf{x}}_k \\
&= I_k\dot{\omega}_k + \sum_j I_j\omega_j\left(\boldsymbol{\omega} \times \hat{\mathbf{x}}_j \cdot \hat{\mathbf{x}}_k\right) \\
&= I_k\dot{\omega}_k + \sum_j I_j\omega_j\left(\boldsymbol{\omega} \cdot \hat{\mathbf{x}}_j \times \hat{\mathbf{x}}_k\right) \\
&= I_k\dot{\omega}_k + \sum_{ji} \varepsilon_{ijk} I_j\omega_j\omega_i
\end{aligned}$$

This gives Euler equations.

13.1 A geometric description of the motion of free top

When the top is free $\mathbf{N} = 0$ and the energy and $\mathbf{J}^2$ are both conserved:

$$\sum_j I_j\omega_j^2 = 2E, \quad \sum I_j^2\omega_j^2 = \mathbf{J}^2$$

The three $\omega_j(t)$ are time dependent, but satisfy two constraints. Each describes an ellipsoid in the three dimensional space $(\omega_1, \omega_2, \omega_3)$. The solution of Euler equations for a free top therefore must lie on the intersection of the ellipsoids.

13.2 Stable and unstable rotations about principal axis

A nice application of Euler equations is to the stability of motion about the principal axis. Suppose $\mathbf{N} = 0$ and the rotation starts with ω_1 large and $\omega_2, \omega_3 = O(\epsilon)$. The first Euler equations says that ω_1 is almost constant:

$$I_1\dot{\omega}_1 = \omega_2\omega_3(I_2 - I_3) = 0(\epsilon^2)$$

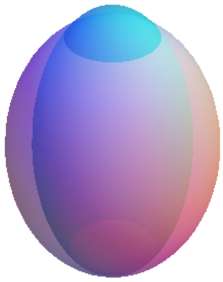


Figure 56: If the z-axis is the short moment rapid rotation bout z is stable the intersection of the two ellipses and is stable

Now let us see how $\omega_{2,3}$ evolves. By Euler

$$I_2\dot{\omega}_2 = \omega_3\omega_1(I_3 - I_1), \quad I_3\dot{\omega}_3 = \omega_1\omega_2(I_1 - I_2)$$

Using the fact that ω_1 is almost constant we get by differentiating and substituting one equation into the other

$$I_3 I_2\ddot{\omega}_2 = I_3\dot{\omega}_3\omega_1(I_3 - I_1) = \omega_2{\omega_1}^2(I_3 - I_1)(I_1 - I_2)$$

This is a second order, homogeneous and linear ODE for ω_2 with (almost) constant coefficients and the solutions is either an exponential or trigonometric function depending on the sign. If I_1 is the middle moment, $I_3 > I_1 > I_2$, then

$$(I_3 - I_1)(I_1 - I_2) > 0$$

and ω_2 will blow up exponentially. A rotation about the middle moment of inertia is unstable. If, however, I_1 is the smallest (or largest) moment, then

$$(I_3 - I_1)(I_1 - I_2) < 0$$

and ω_2 will be oscillatory.

The result can be understood geometrically by looking at how two ellipsoids intersect. There are two cases: One looks like the level lines near a maximum or a minimum in 2-d, namely and ellipse. The other looks like the level lines near a saddle, namely two crossing paths[12].

[12] I owe this insight to Amos Ori.

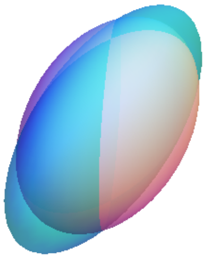


Figure 57: The intersection of two ellipsoids in the case of the rotation about the unstable middle axis.

A Quaternions

A.1 Intro

Quaternions were invented by Hamilton. They are useful for vector analysis and rotations. They were displaced by the vector analysis and Euler angles because Gibbs and Heaviside succeeded in giving a more elegant presentation than Hamilton. Quaternions then had a comeback with quantum mechanics, spin and computer graphics.

A.2 Quaternions

Consider the real vector space with (real) elements

$$q = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$$

We want to turn this into a (non-commutative) algebra, so we can also multiply elements. Do do this we set

$$\mathbf{i}\,\mathbf{i} = -1, \quad \mathbf{i}\,\mathbf{j} = \mathbf{k}$$

and cyclic permutations. We can realize the algebra with Pauli matrices. You will meet them in QM. They are defined by

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and satisfy the algebra

$$\sigma_i \sigma_j = i\varepsilon_{ijk}\sigma_k, \quad i, j, k \in 1, 2, 3$$

Now make the identification

$$\mathbf{i} = i\sigma_x, \quad \mathbf{j} = i\sigma_y, \quad \mathbf{k} = i\sigma_z$$

so that

$$q = q_0\sigma_0 + i\mathbf{q}\cdot\sigma, \quad \mathbf{q}\cdot\sigma = \sum q_j\sigma_j$$

and check that everything fits.

A.3 Vectors

$\mathbf{q}$ is the part of the quaternion that represents a vector. We can recover the components of the vector from the (matrix) representation by

$$q_j = \frac{1}{2}Tr(\mathbf{q}\cdot\sigma\,\sigma_j)$$

where I used

$$Tr\sigma_j\sigma_k = 2\delta_{jk}$$

This gives a way to represent vectors in term of Hermitian, traceless, 2×2 matrices. For example, if we have two vectors $\mathbf{v}$ and $\mathbf{w}$ then one can check

$$\mathbf{v}\cdot\mathbf{w} = \frac{1}{2}Tr(\mathbf{v}\cdot\sigma)(\mathbf{w}\cdot\sigma)$$

A.4 Conjugate

A natural notion of conjugate comes from the Pauli matrix representation using the notion of matrix adjoint:

$$q^* = q_0\sigma_0 - i\mathbf{q}\cdot\sigma = q_0 - q_1\mathbf{i} - q_2\mathbf{j} - q_3\mathbf{k}$$

In particular, the norm of a quaternion is Here we find

$$\|q\|^2 = q^*q = q_0^2 + \mathbf{v}\cdot\mathbf{v} \geq 0 \geq 0$$

Quaternions are a field: Every non-zero element has an inverse

$$q^{-1} = \frac{q^*}{\|q\|}$$

A.5 Rotations and Unitary matrices

Quaternions of length 1 correspond to unitary matrices. Indeed

$$U = x_0\sigma_0 + i\mathbf{x}\cdot\sigma = \begin{pmatrix} x_0 + ix_3 & ix_1 + x_2 \\ ix_1 - x_2 & x_0 - ix_3 \end{pmatrix}, \quad x_0^2 + \mathbf{x}\cdot\mathbf{x} = 1$$

$$U^* = x_0\sigma_0 - i\mathbf{x}\cdot\sigma = U^{-1}$$

To see how this is related to rotation in 3 dimensions let us look at the matrix representation of a vector

$$\mathbf{v} \Longleftrightarrow \mathbf{v}\cdot\sigma$$

Consider the action of conjugation by unitary

$$\mathbf{v}\cdot\sigma \mapsto U\mathbf{v}\cdot\sigma U^*$$

It is easy to see that conjugation maps a traceless matrix to a traceless matrix, so it maps a vector to a vector. Moreover, it preserves the length of a vector by the unitary invariance of the trace

$$\mathbf{v}\cdot\mathbf{v} = \frac{1}{2}Tr(\mathbf{v}\cdot\sigma)^2 = \frac{1}{2}TrU(\mathbf{v}\cdot\sigma)^2U^* = \frac{1}{2}Tr\big((U\mathbf{v}\cdot\sigma U^*)^2\big)$$

The representation of rotations in 3 dimensions by unitary 2×2 matrices is faithful except for one point which turns out to be important in QM: U and $-U$ give the same rotation of vectors.

If we write $U = q$ as a unit quaternion then

$$\frac{1}{2}TrU = q_0 = \cos\theta/2$$

since $|q_0| \le 1$. We shall now see that θ is the angle of rotation in 3 dimensions. To see why $\theta/2$ recall that $U = \pm\sigma_0$ both represent no rotation, or rotation by a multiple of 2π when acting on vectors. This corresponds to $\cos\theta/2 = \pm 1$ i.e. $\theta/2 = 0, \pi$ and θ is then a multiple of 2π rotation.

A.6 Rotation

Consider a unit quaternion (unitary)

$$u = u_0\sigma_0 + i\mathbf{u}\cdot\sigma, \quad u_0^2 + \mathbf{u}\cdot\mathbf{u} = 1$$

It is convenient to rewrite this as

$$u = \cos\phi\,\sigma_0 + i\sin\phi\,\hat{\mathbf{u}}\cdot\sigma, \quad \hat{\mathbf{u}}\cdot\hat{\mathbf{u}} = 1$$

Now consider its (adjoint) on the vector $\mathbf{v}$ represented as a traceless matrix

$$\mathbf{v}\cdot\sigma \mapsto u\mathbf{v}\cdot\sigma u^* = \mathbf{v}'\cdot\sigma$$

A computation gives

$$\begin{aligned}\mathbf{v}' &= (\cos^2\phi - \sin^2\phi)\mathbf{v} - 2\sin\phi\cos\phi\,\hat{\mathbf{u}}\times\mathbf{v}\\ &= \cos 2\phi\mathbf{v} - 2\sin 2\phi\,\hat{\mathbf{u}}\times\mathbf{v}\end{aligned} \tag{A.1}$$

a rotation of $\mathbf{v}$ by 2ϕ around the $\hat{\mathbf{u}}$ axis.