Notes prepared by Prof. Ramsharan Rangarajan for his lecture to ME 256 students on Feb. 11, 2016.

Noether's theorem

February 10, 2016

Conservation laws, such as conservation of energy, linear and angular momentum are fundamental in mechanics. Pondering why these laws seems to appear so commonly in nature is as much a topic in philosophy as it is in mechanics. We will not attempt to pursue it. Instead, we will try to understand a useful result that helps us discover such conservation laws in a systematic way. Most if not all the examples we will discuss today are conservation laws you have seen in the past. But:

- You may have seen them derived in an ad hoc manner, using tools and equations specific to the subject. Hence conservation of electric charge in electromagnetics may have been derived using very different arguments than conservation of angular momentum in a mechanical system.
- Often, such conservation laws are derived *after computing the solution*, for instance after solving a (partial) differential equation resulting from a statement of force balance.

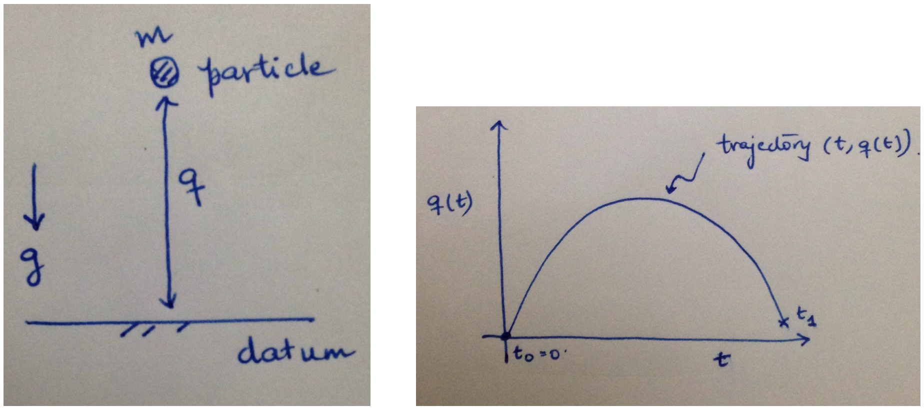
The subject of this lecture, Noether's theorem provides a unified way of discovering conservation laws for systems that follow variational principles. In essense, the theorem identifies a relationship between symmetries and conservation laws in a system.

1 Recap

We begin by recalling concepts learnt from previous lectures with the help of a simple example. Consider a particle of mass m moving under the influence of gravity and denote its height above the datum by q. That is, q will be the coordinate defining the position of the particle. We are interested in computing the trajectory of the particle as a function of time, i.e., the mapping $t \mapsto q(t)$.

1.1 The Lagrangian & the action

We can easily write down Newton's second law as $m\ddot{q} = mg$. But this is not the point of the exercise. As we have seen, this equation of motion also follows



(a) Coordinate q represents the height above a datuum

(b) Trajectory of the particle as a function of time, satisfying end conditions

Figure 1: Choice of coordinate for a particle moving in a gravitational field.

from Hamilton's principle of least action as well. To this end, we first define the Lagrangian for the system, in this case

$$L(t, q, \dot{q}) := \frac{1}{2}m(\dot{q})^2 - mgq, \tag{1}$$

which is the difference of the kinetic and potential energies. Then we have the action defined as

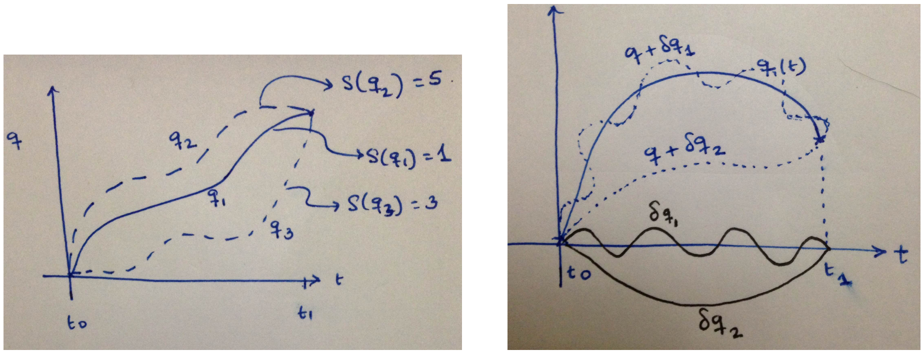
$$S(q) := \int_{t=t_0}^{t_1} L(t, q, \dot{q})\, dt. \tag{2}$$

A couple of points should be noted. In (1), since q and $\dot{q}$ are functions of t, L can be thought of as a function of t alone. But the notation $L(t, q, \dot{q})$ helps to specify the dependence of L on q and $\dot{q}$. Second, the action S in (2) is an example of a functional. It takes the function q as its argument and returns a number. The action assigns a "cost" with each admissible trajectory, where admissibility here requires that q be (i) sufficiently smooth to permit calculation of all derivatives that appear in our calculations, and (ii) that q satisfy end conditions $q(t_0) = q_0$ and $q(t_1) = q_1$ which fix the initial and final positions of the particle.

1.2 Extremality of the action

To extremize a function $f(x)$, we set the first derivative to zero $f'(x) = 0$. Similarly, to extremize the function $f(x, y)$ along a given direction $\mathbf{u} = (u_x, u_y)$, we set $\nabla f \cdot \mathbf{u} = 0$. Equivalently, we set

$$\frac{d}{d\varepsilon} f(x + \varepsilon u_x, y + \varepsilon u_y)\Big|_{\varepsilon=0} = 0. \tag{3}$$



(a) The action assigns a cost to every admissible trajectory

(b) Finding an extremal trajectory requires investigating the value of the functional at all nearby admissible trajectories

Figure 2: The action functional.

Eq. (3) is generalized to define the notion of extremality of functionals as well. In the case of the action S, the trajectory q plays the role of the point (x, y) and an admissible variation δq plays the role of the direction $\mathbf{u}$.

Definition: We say that an admissible trajectory $t \mapsto q(t)$ is an extremum of the action S if for all admissible variations δq,

$$\frac{d}{d\varepsilon} S(q + \varepsilon \delta q)\bigg|_{\varepsilon=0} = 0. \tag{4}$$

A commonly used notation for denoting the derivative in (4) is $\langle \delta S(q), \delta q \rangle$, which is to be read as the derivative of the functional S evaluated at the trajectory q along the variation δq.

The idea behind (4) is that for q to be an extremum, we should examine the value of the action at all nearby trajectories. We have done that by requiring (4) hold for every admissible δq. Compare (3) and (4) and make sure that you understand the parallels between the definitions of extremality for functions and functionals.

1.3 Hamilton's principle

Hamilton's principle of least action states that the actual trajectory of a system will be one that renders the action stationary for all admissible variations[1].

We have seen in previous lectures that by invoking the divergence theorem and the fundamental lemma of calculus of variations, Hamilton's principle ap-

[1] The principle is also commonly stated by requiring minimization of the action. We will however be computing stationary points, and not determine whether the computed trajectories indeed minimize the action.

plied to the action in (2) yields the familiar Euler-Lagrange equations

$$\langle \delta S(q), \delta q \rangle = 0 \ \forall \delta q \ \Rightarrow \ \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q}. \tag{5}$$

Generalized momentum: For the particle system with Lagrangian (1), the E-L equations yield Newton's second law. You may have noticed from examples in previous lectures as well, that the E-L equations are often just statements of force balance. Such an observation is made more evident by introducing generalized momentum variables.

Definition: The momentum p conjugate to the coordinate q for a system with Lagrangian $L(t, q, \dot{q})$ is defined as

$$p = \frac{\partial L}{\partial \dot{q}}. \tag{6}$$

Using the momentum p in the E-L equation (5) yields

$$\frac{dp}{dt} = \frac{\partial L}{\partial q}. \tag{7}$$

For the particle moving under gravity, $p = \partial L/\partial \dot{q} = m\dot{q}$, which is of course the linear momentum, and the E-L equation for the particle now reads $\dot{p} = mg$.

Hamiltonian: For a system with Lagrangian $L := L(t, q, \dot{q})$, the Hamiltonian[2] is defined as

$$H(t, q, p) := p\dot{q} - L(t, q, \dot{q}). \tag{8}$$

Notice that the arguments of H are t, q, p and unlike L, H does not depend explicitly on $\dot{q}$. For the particle moving under gravity, we have

$$H(t, q, \dot{p}) = \frac{\dot{p}^2}{2m} + mgq, \tag{9}$$

which is interpreted to be the energy of the system.

Homework: Demonstrate that the E-L equation (7) can be written in terms of the Hamiltonian as[3]

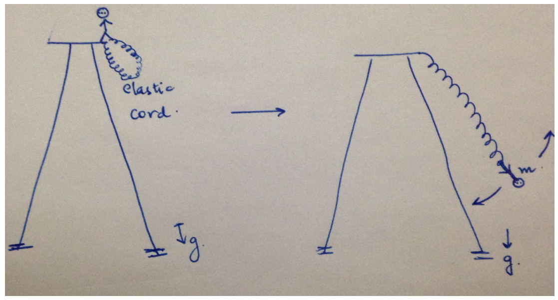
$$\frac{d\dot{p}}{dt} = -\frac{\partial H}{\partial q}. \tag{10}$$

1.4 Homework: Bungee jump

As was the case in the simple example with the particle moving in gravity, E-L equations are statements about the rate of change of generalized momenta conjugate to our choice of coordinates for the system. When we choose displacements as coordinates, we recover statements of force balance.

[2]You may recognize that H is the Legendre transformation of L.

[3]Rewriting the E-L equations in terms of the Hamiltonian is not just an academic exercise. It has profound consequences on understanding the dynamics of the system, made possible by exploiting results from a topic in mathematics called symplectic geometry.



It is fair to ask what is the point in learning a new technique to arrive at the same equations. Among others, one of the main benefits is the fact that we can write down equations of motion while using generalized coordinates. That is, the unknowns need not be restricted to just translational degrees of freedom. For example, when one of the chosen coordinates is an angle, the corresponding E-L equation can be interpreted as a statement of balance of torque. In this case, the force conjugate to the angular coordinate is a torque. In many real world situations, it is necessary to choose coordinates for which statements of force balance are not evident. On the otherhand, E-L equations yield these statements in a systematic way. The following problem explifies this point.

Say that you have decided to try a bungee jump, wherein you jump off a tall structure such as a bridge with an *elastic* cord bound to your feet. You are interested in computing your trajectory during a *safe* jump beforehand, so that you can be sure of what to expect. For simplicity, assume that

(a) your motion is restricted to a plane,

(b) as a first approximation, consider yourself to be a point particle so that your moment of inertia can be ignored,

(c) the cord is linearly elastic (a spring),

(d) ignore all dissipative processes like air drag, friction, etc.

The following steps will help you solve the problem:

(i) Choose a proper system of coordinates. How many coordinates are needed?

(ii) Write down the set of admissible trajectories and variations.

(iii) Write down the Lagrangian and the action.

(iv) Write down the E-L equations. Interpret the terms in each equation.

(v) Integrating these equations numerically (Matlab/Mathematica).

(vi) To improve the model, replace the approximation of the body as a point particle by that of a rigid bar. Again, assume that the motion is restricted to a plane. What would be a proper coordinate system for the new system?

2 Conservation laws and symmetry

The main topic of today's lecture is Noether's theorem, which provides an elegant connection between conservation laws and coordinate symmetries in a system. Let us first see the statement and then understand what it says using examples. To avoid indicial complications, we state the theorem for the case in which q is a scalar-valued function, $L := L(t, q, \dot{q})$ and $S := S(q)$.

Theorem: Consider a system with Lagrangian $L := L(t, q, \dot{q})$, Hamiltonian $H(t, q, p) := p\dot{q} - L(t, q, \dot{q})$ and action $S(q) := \int_{t=t_0}^{t_1} L(t, q, \dot{q})\, dt$. Assume that:

(i) the action is extremized at the trajectory q, and that

(ii) the action is invariant under the continuous coordinate transformations

$$\begin{cases} \hat{t} & = T(t, q, \varepsilon), \\ \hat{q} & = Q(t, q, \varepsilon). \end{cases} \tag{11}$$

Then, with

$$\xi = \left.\frac{\partial Q}{\partial \varepsilon}\right|_{\varepsilon=0} \text{ and } \tau = \left.\frac{\partial T}{\partial \varepsilon}\right|_{\varepsilon=0}, \tag{12}$$

the quantity $(p\xi - H\tau)$ is conserved along the extremal trajectory.

We have stated a simplified version of a more general statement, and to avoid indicial nuisances, assumed that q is a scalar-valued function. The theorem says that a certain quantity remains a constant along the real trajectory of the system, provided that assumptions (i) and (ii) are satisfied. To fully understand the statement, we need to understand what is meant by a *continuous coordinate transformation*, and what is meant by the *action remaining invariant* under such transformations.

2.1 Coordinate transformations

Given a function $f(x, y)$, we can rewrite f using a different pair of coordinates (u, v), so that $\hat{f}(u, v) = f(x, y)$. For example, the function

$$f(x, y) = \sqrt{x^2 + y^2}(x + y)$$

can be written in a polar coordinate system (r, θ) as

$$\hat{f}(r, \theta) = r^2(\cos\theta + \sin\theta).$$

The functional forms of f and $\hat{f}$ are very different, although they represent the same function. This means for example, that if we plot f and $\hat{f}$ over the plane, they will represented by identical surfaces in 3D.

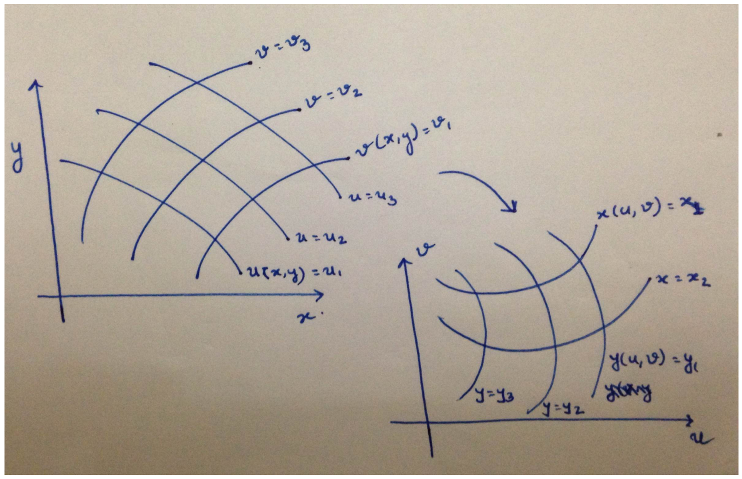


Figure 3: Straight lines in one coordinate systems can be curved in a second coordinate system.

Is any rule relating (x, y) and a new coordinate system (u, v) permissible? No. We have to insist that $(u(x, y), v(x, y))$ uniquely identify a point in the domain of the function. For example, the choice $u = x^2$ would be inadmissible because the points x and $-x$ will both be mapped to the same value. The coordinate transformation $(x, y) \mapsto (u(x, y), v(x, y))$ is admissible if it is one-to-one and onto, i.e., a bijection. This allows us to go back and forth between the two coordinate systems (x, y) and (u, v).

The statement of Noether's theorem refers to a change of coordinates from (t, q) to a new one $(\hat{t}, \hat{q})$ via the rules $\hat{t} = T(t, q, \varepsilon)$ and $\hat{q} = Q(t, q, \varepsilon)$. We have made these transformations dependent on a parameter ε. For each value of ε, we have a different coordinate system. For example, we may have

$$\begin{aligned} \hat{t} &= t \cos \varepsilon + q \sin \varepsilon, \\ \hat{q} &= q \cos \varepsilon - t \sin \varepsilon, \end{aligned}$$

which defines a coordinate system rotated by angle ε.

2.2 Invariance of a functional

We understand the word "invariant" to mean something that is unchanged. Then, it is quite clear what it means for a function to be "invariant" under a certain coordinate transformation. For example:

(a) The function $f(x, y) = \sqrt{x^2 + y^2}$ is invariant under rotation of coordinates

$$\begin{aligned} \hat{x} &= x \cos \theta + y \sin \theta, \\ \hat{y} &= y \cos \theta - x \sin \theta, \end{aligned}$$

because we find that $f(\hat{x}, \hat{y}) = f(x, y)$. That is, replacing the arguments (x, y) by new values $(\hat{x}(x, y), \hat{y}(x, y))$ leaves the value of the function unchanged.

(b) Similarly, the function $f(x, y) = |x|$ is invariant under the coordinate transformation $\hat{x} = -x, \hat{y} = y + \varepsilon$ for any ε because $f(\hat{x}, \hat{y}) = |\hat{x}| = |x| = f(x, y)$.

We have seen a couple of examples where the value of a function $f(x, y)$ remains unchanged by certain coordinate transformations. Such coordinate transformations are called symmetries of the function f[4].

We can now generalize the idea of invariance of functions to invariance of *functionals*, and in particular the action $S(q) = \int_{t=t_0}^{t_1} L(t, q, \dot{q})\, dt$. We say that $S(q)$ is invariant under the coordinate transformation $(t, q) \mapsto (\hat{t}, \hat{q})$ if $S(\hat{q}(\hat{t})) = S(q)$. More explicitly, this requires that the action associate the same cost with the trajectory $\hat{t} \mapsto \hat{q}(\hat{t})$ as it does with $t \mapsto q(t)$. That is,

$$S(\hat{q}) = \int_{\hat{t}=\hat{t}_0}^{\hat{t}_1} L(\hat{t}, \hat{q}, \frac{d\hat{q}}{d\hat{t}})\, d\hat{t}. \tag{13}$$

Observe that:

- We are simultaneously changing both coordinates t and q. Hence the curve $(t, q(t))$ over $[t_0, t_1]$ can look completely different that the transformed curve $(\hat{t}, \hat{q}(\hat{t})$ over the interval $[\hat{t}_0, \hat{t}_1]$.

- We are not saying that the action assign the same cost to the curves $t \mapsto q(t)$ and $t \mapsto \hat{q}(t)$, although this may be possible too. In general however, it may be necessary to simultaneously transform both t as well as q.

- The invariance of the action S does not require that the function L be invariant under the transformation $(t, q) \mapsto (\hat{t}, \hat{q})$. It may however be the case in some systems. In general, S can be invariant even though L is not.

Let us look at some examples:

(i) Consider the example of the particle moving in gravity with action

$$S(q) = \int_{t=t_0}^{t_1} \left(\frac{1}{2} m\dot{q}^2 - mgq \right)\, dt,$$

[4]Symmetries are usefully studies using the language of group theory rather than just as change of coordinates.

and the coordinate transformations $\hat{t} = t + \varepsilon, \hat{q} = q$. We have

$$\begin{aligned} S(\hat{q}) &= \int_{\hat{t}=\hat{t}_0}^{\hat{t}_1} \left(\frac{1}{2} m \left(\frac{d\hat{q}}{d\hat{t}} \right)^2 - mg\hat{q} \right) d\hat{t} \\ &= \int_{t_0}^{t_1} \left(\frac{1}{2} m \left(\frac{d\hat{q}}{dt} \frac{dt}{d\hat{t}} \right)^2 - mg\hat{q} \right) \frac{d\hat{t}}{dt} dt \\ &= \int_{t_0}^{t_1} \left(\frac{1}{2} m\dot{q}^2 - mgq \right) dt \\ &= S(q) \end{aligned}$$

In this particular case, it turned out that the function L was also invariant under the coordinate transformation that represents a translation in time.

(ii) Consider an example of the length functional

$$S(q) = \int_a^b \sqrt{1 + \dot{q}^2}\, dt.$$

Show that it is invariant under the coordinate transformation (with actual computations)

$$\begin{cases} \hat{t} &= t\cos\varepsilon + q\sin\varepsilon, \\ \hat{q} &= q\cos\varepsilon - t\sin\varepsilon. \end{cases} \tag{14}$$

It is clear why the arc-length should be conserved under the transformation (14), irrespective of the value of ε. The functional $S(q)$ measures the length of the curve $t \mapsto (t, q(t))$ for $t \in [a, b]$. Inspecting (14), we realize that for each ε, the transformation (14) is a rotation of the axes by an angle ε. In particular, (14) is an isometry that preserves all angles and lengths. Therefore, changing coordinates according to (14) simultaneously rotates the axes and the curve. This is equivalent to simply rotating the page on which the curve is drawn, and therefore does not alter the length of the curve.

We now understand what the second assumption in statement of Noether's theorem means— we need to find, by a lucky guess or otherwise, some coordinate transformations $(t, q) \mapsto (\hat{t}, \hat{q})$ such that $S(q) = S(\hat{q})$.

2.3 Why do we even expect conservation laws to be related to symmetry

Before looking at what Noether's theorem means for some examples, let us see why we even expect such a theorem to be true. At the outset, it is not at all clear what is the relation between conservation laws and symmetries. We resort to a simple example. Consider Newton's second law relating the rate of change of linear momentum equals the net external force acting on a body, i.e.,

$$\frac{d\mathbf{p}}{dt} = \mathbf{F}. \tag{15}$$

Suppose that the force field $\mathbf{F}$ is derived from an external potential [5], that is, $\mathbf{F} = -\nabla U$. Then we may write

$$\frac{d\mathbf{p}}{dt} = -\frac{\partial U}{\partial x} = \lim_{\varepsilon \to 0} \frac{U(x+\varepsilon) - U(x)}{\varepsilon}, \tag{16}$$

where for simplicity, we have assumed the problem to be one dimensional. Notice that if the potential U was a constant, then $U(x+\varepsilon) = U(x)$ irrespective of x and ε. Then we get that $\mathbf{F} = 0$ and hence $d\mathbf{p}/dt = 0$. Now $U(x) = U(y)$ is the same as saying that U is invariant under translations $x \mapsto x'$, i.e., that translation in space is a symmetry of the function U. In this way, we have deduced that translational symmetry of U implies conservation of linear momentum.

Of course, the above example is trivial. Noether's theorem systematically establishes such relationships between symmetries of the action and quantities conserved along the trajectory. Often, identifying conserved quantities from first principles is difficult, especially if the conserved quantity is not a physically relevant quantity such as a momentum, or an energy. For example, the J-integral is a widely used concept in fracture mechanics. It has the interpretation of a configurational force acting on a crack, not a physical one. You may see the path-independence of the J-integral derived in a fracture mechanics class starting from first principles. However, we can show that such a conservation law is the consequence of a specific symmetry of the action. Infact, we can show such a result for any hyperelastic, spatially homogeneous material.

2.4 Interpreting Noether's theorem: examples

(i) Consider a particle moving in space, under the action of an external potential $U(x_3)$. The action is given by

$$S(x_1, x_2, x_3) = \int_a^b \left(\frac{1}{2} m \dot{x}_i \dot{x}_i - U(x_3) \right) dt$$

where x_1, x_2, x_3 are all functions of t. Observe that the change of coordinate $(x_1, x_2, x_3, t) \mapsto (x_1, x_2, x_3, t + \varepsilon)$ leaves the action invariant. In terms of the notation used in (11), we have $Q_i(x_1, x_2, x_3, t, \varepsilon) = x_i$ and $T(x_1, x_2, x_3, t, \varepsilon) = t + \varepsilon$. Then Noether's theorem tells us that

$$p_i \xi_i - H\tau = \text{constant}, \tag{17}$$

[5] Recall that in a conservative force field, the amount of work done depends only on the end points and not the path taken. Force fields of the form $\mathbf{F} = -\nabla U$ are necessarily conservative.

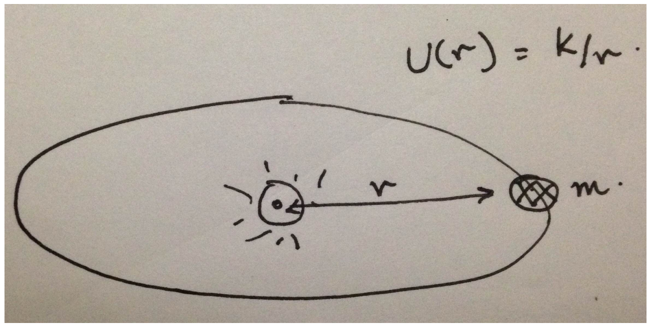


Figure 4: Motion of a particle in a central field

where $\xi_i := \frac{\partial Q_i}{\partial \varepsilon}(\varepsilon = 0)$ and $\tau = \frac{\partial T}{\partial \varepsilon}(\varepsilon = 0)$. Since $\xi_i = 0$ and $\tau = 1$, we get that $H =$ a constant. However,

$$H = p_i \dot{x}_i - L = \frac{1}{2} m \dot{x}_i \dot{x}_i + U(x_3) \tag{18}$$

is in fact the energy. Hence we have seen that the symmetry of the action corresponding to translation in time yields the statement of conservation of energy.

(ii) Let us continue with the previous example. This time, observe that the transformation $(x_1, x_2, x_3, t) \mapsto (x_1 + \varepsilon, x_2, x_3, t)$ is a symmetry. This time, $\xi_1 = 1$ while $\xi_{2,3} = \tau = 0$. We therefore get the conserved quantity to be

$$p_i \frac{\partial Q_i}{\partial \varepsilon} - H\tau = p_1 = \text{constant}. \tag{19}$$

We have found that translational invariance of the action along the x_1 coordinate yields the statement of conservation of momentum $p_1 = m\dot{x}_1$. In an identical manner, we get conservation of momentum $p_2 = m\dot{x}_2$. Notice also that such translational symmetry does not exist along x_3 because of the presence of the external potential $U(x_3)$. As expected therefore, we do not get a conservation law for p_3.

2.5 Conservation law in a central force field

Consider a planet of mass m moving in a plane and attracted to the origin due to a potential of the form k/r, where r is the distance of the particle from the origin.

(i) Choose a proper system of coordinates for the system.

(ii) Write down the Lagrangian, the momenta conjugate to your choice of coordinates, the Hamiltonian and the action for the system.

(iii) Verify that the action is invariant under transltion of time. What conservation law does this yield?

(iv) Verify that the action is invariant under rotation of coordinates.

(v) What is the corresponding conservation law predicted by Noether's theorem?

(vi) Interpret what the conservation law says.

Solution:

(i) It is convenient to choose a polar coordinate system (r, θ) for the planet. The trajectory in the play of motion is hence of the form $t \mapsto (r(t), \theta(t))$. The mapping between Cartesian coordinates and (r, θ) is of course the usual one: $x = r\cos\theta, y = r\sin\theta$

(ii) The magnitude of the particle's velocity is

$$\begin{aligned}|v|^2 &= \dot{x}^2 + \dot{y}^2 \\ &= (\dot{r}\cos\theta - r\dot{\theta}\sin\theta)^2 + (\dot{r}\sin\theta + r\dot{\theta}\cos\theta)^2 \\ &= \dot{r}^2 + r^2\dot{\theta}^2.\end{aligned}$$

The Lagrangian follows as

$$L(t, r, \theta, \dot{r}, \dot{\theta}) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - \frac{K}{r}.$$

The action is given by

$$S(r, \theta) = \int_{t=0}^{T} L(t, r, \theta, \dot{r}, \dot{\theta}).$$

The momentum conjugate to the coordinates r and θ are

$$\begin{aligned}p_r &= \frac{\partial L}{\partial \dot{r}} = m\dot{r}, \\ p_\theta &= \frac{\partial L}{\partial \dot{r}} = mr^2\dot{\theta}.\end{aligned}$$

The Hamiltonian is then given by

$$\begin{aligned}H(t, r, \theta, p_r, p_\theta) &= p_r\dot{r} + p_\theta\dot{\theta} - L \\ &= \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{K}{r}.\end{aligned}$$

(iii) The Lagrangian is independent of time, it is trivial to verify that $\hat{t} = t + \varepsilon$ is a symmetry of the action. As we have seen before, this yields the conservation law $H =$ constant, which is the statement of conservation of total energy.

(iv) To check if rotation within the plane of motion is a symmetry of the action, we need to check that $\hat{r} = r, \hat{\theta} = \theta + \varepsilon, \hat{t} = t$ leaves the action invariant. Convince yourself that this is trivial.

(v) The corresponding conservation law is simply

$$p_\theta = \text{constant} \Rightarrow mr^2\dot{\theta} = \text{constant} \Rightarrow r^2\dot{\theta} = \text{constant}$$

(vi) What does the conservation law $r^2\dot{\theta} = \text{constant}$ mean? We claim that it is exactly Kepler's third law of planetary motion, namely, that equal areas are swept by the line joining the planet and the sun in equal times. To see this, let us compute the rate at which area is swept by the line joining the planet and the sun:

$$\frac{dA}{dt} = \int_{\xi=0}^{r} 2\pi\xi\dot{\theta}d\xi = \pi r^2\dot{\theta}, \tag{20}$$

from where out conclusion follows.