16 Applications to Physical Systems: Hamilton's Principle and Noether's Theorem

Nothing happens until something moves
—Albert Einstein

The force is strong with this one
—James Earl Jones

The universe is built on a plan the profound symmetry of which is somehow present in the inner structure of our intellect.
—Paul Valery

Don't confuse symmetry with balance
—Tom Robbins

The Calculus of Variations has a myriad of applications in the physical sciences. It is often the primary tool used in the derivation of many physical laws and forms the basis for many engineering principles. This is the foundation for what is often referred to as Lagrangian mechanics which is a description of physical laws often taken as parallel to Newtonian mechanics (remember Newton's Laws?). We focus on one of the most celebrated aspects of this development, as well as one of the most mathematically beautiful. In reality any principle that can be stated as an optimization principle has its roots in variational methods.

Hamilton's principle or the principle of least action dictates that the actual 'path' taken by a dynamical system will minimize the time integral of the difference between the kinetic and potential energy. This minimization principle can be applied to a variety of physical and mechanical systems to derive the equations of motion governing the evolution of the system. This is in fact often how we arrive at the equations governing the motion of certain systems. Intuitively we may think of this as nature's way of being lazy, i.e. everything will naturally try to minimize the amount of kinetic energy spent, while maximizing the potential energy saved.

Succinctly stated (without proof):

Proposition 16.0.1 (Hamilton's Principle). *The actual path of a dynamical system* $\mathbf{x}(t)$ *will minimize the time integral of the difference between kinetic and potential energies, i.e.* Nature is LAZY.

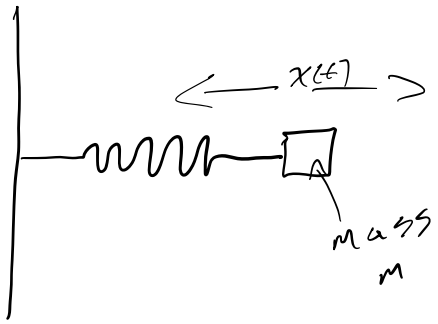


Figure 16.1: Motion of a simple spring.

This principle is far more general than one might suppose (or is over-used by physicists to try and explain everything depending on your perspective). In fact, this is one method of deriving the partial differential equations that govern the principles of general relativity which is one example that we do not go into great detail to derive here. The underlying principle that nature is lazy is far more reaching than relativity, and underlies definitive questions about human nature and the basic premise that philosophers use to discuss reality. So the next time you accused of laziness, just blame Hamilton, or say you are getting in touch with nature.

Remark 16.0.2. It is worth noting at this point that these notes are dedicated to developing a course in Mathematics and so we will not focus on particular problems in Physics, nor will we be overly concerned with how you derive the kinetic or potential energy for different complicated settings. Specifically, most of the examples that are contained in this chapter will deal with potential energy that is either due to gravity, or the compression of a spring. More sophisticated calculations of the potential energy e.g. magnetic potentials, will only be considered in specific examples and not in full generality.

16.1 Some physical examples

Hamilton's principle applied to a physical system indicates that we want to minimize

$$J[y] = \int_{t_0}^{t_1} L(t, x, x')dt$$

where $L = T - U$ (T is the kinetic energy and U is the potential energy). Usually the Lagrangian $L(t, x, x')$ has t (time) as the independent variable and $x(t)$, $x'(t) = \frac{dx}{dt}$, and $x''(t) = \frac{d^2x}{dt^2}$.

Example 16.1.1. Consider a spring attached to a rigid wall with a freely moving mass m on the other end, see Figure 16.1. If the mass is at equilibrium when $x = 0$ then the potential energy in the spring is given by Hooke's law, i.e. $U = \frac{1}{2}kx^2$ where k is a constant. The kinetic energy is $T = \frac{1}{2}mx'^2$, so the Lagrangian is

$$L = T - U = \frac{1}{2}m(x')^2 - \frac{1}{2}kx^2.$$

Remark 16.1.2. This example demonstrates that we can use Hamilton's principle (or more affectionately known as the 'lazy principle') to derive the same equations of motion that we originally derived using Newton's second law. This is a remarkable fact that brings to mind two different fundamental approaches to classical physics. The first is to use Newton's second law (sum of the forces = mass times acceleration), and the one introduced here is to note that lazy behavior is always the best option (Hamilton's principle).

Often the real trick to figuring out the evolution equations for a physical system reduces to understanding how to compute that systems' kinetic and potential energies. In most settings, computation of the potential energy is closely related to finding the force acting on an object, so it really is all connected.

With this in mind, we return to one other example that we had previously considered. This time, we consider the potential energy of BB-8 rolling down a sand-dune; see Figure 15.2.

Example 16.1.3. Returning to BB-8 rolling down a sand dune on Jakku, the kinetic energy is made up of the translational and rotational kinetic energy (recognizing that this comes from the translational and rotational velocities...don't stress about the rotational velocity if you haven't seen it before) so that

$$T = \frac{1}{2}m(y')^2 + \frac{1}{4}mR^2(\theta')^2.$$

This is because the rotational kinetic energy is given by $\frac{1}{2}I(\theta')^2$ where $I = \frac{m}{2}R^2$ is the moment of inertia, and the potential energy is defined by the effects of gravity, i.e.

$$U = mg(y - l)\sin\phi.$$

This brings up the important point that Hamilton's principle is useful only insofar as you know how to compute the potential and kinetic energies. If we had omitted the rotational kinetic energy in this case, BB-8 would have had a very hard time getting anywhere, but would have had to rely on gravity to slide him down the sand dune. This is not at all what Disney had in mind, and would mean that Jakku had some very interesting sand.

Remark 16.1.4. It is important to point out that the kinetic and potential energies are *NOT* dependent on the choice of coordinates used to describe the system. Nevertheless it is often ideal to consider the system in terms of a particular set of coordinates because it will make the resultant calculation a bit simpler. In the following examples we will find that Cartesian coordinates are often not the ideal setting, dependent on the geometry of the problem we are trying to model. Even so, we will find that it is often easier to state the kinetic energy (at least for the physically simple examples we are considering) in terms of Cartesian coordinates first, and then make the change of coordinates to whatever system is amenable to the problem under consideration. This is illustrated in the following example.

Now for a far more practical consideration, we will consider Thor hanging from a chain when he is 'interrogating' Surtur. To fully understand why Thor keeps swinging around in a circle is actually just a product of conservation of angular momentum, a very clever comedic plot twist for the writers of the script. We won't focus on the actual conservation of angular momentum here, but we will consider a simplified model where we suppose that Thor is hanging from a chain of fixed length, and allowed to swing in a plane rather than three dimensional space.

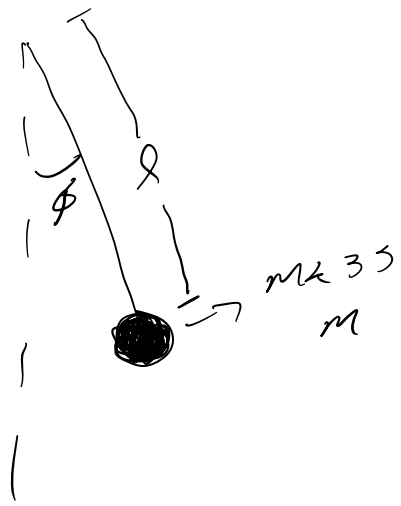


Figure 16.2: Motion of a simple pendulum (a practical model of Thor swinging from a chain.

Example 16.1.5. Consider a simple pendulum of length l (the chain Thor is hanging from) as depicted in Figure 16.2 with the mass m (Thor's weight which we should probably not extrapolate on) suspended from a frictionless support (Surtur is a fire demon after all, so friction should naturally not be an issue). Let $x = l\sin\theta$, $y = l\cos\theta$ denote the displacement of the mass along the arc of the motion where l is the length of the pendulum (which we will suppose is a stiff rod for simplicity). Then the kinetic energy is given by

$$T = \frac{1}{2}m\left((x')^2 + (y')^2\right) = \frac{1}{2}ml^2\phi'^2.$$

The potential energy is

$$U = mg(l - l\cos\phi),$$

i.e. the gravitational acceleration multiplied by the vertical distance the bob is above its equilibrium position (it doesn't matter how far above the ground the bob is hung as the string will not allow the bob to fall all the way to the ground). This implies that the 'action integral' (the term beloved by physicists everywhere) is

$$J[\phi] = \int_{t_0}^{t_1}(T-U)dt = \int_{t_0}^{t_1}\frac{1}{2}ml^2(\phi')^2 - mg(l - l\cos\phi)dt.$$

Thus, the motion of the pendulum is governed by

$$\phi'' - \frac{g}{l}\sin\phi = 0.$$

If ϕ is small this is approximated by

$$\phi'' - \frac{g}{l}\phi = 0.$$

This simple pendulum example where Thor is interrogating Surtur is of course a little simplified (Thor can't rotate if he is swinging in a plane), but it does illustrate one of the most canonical examples in all of classical physics. The planar pendulum is quite simple to describe, but surprisingly exhibits some of the most complicated dynamics of any such simple a system. If the angle ϕ is small then the pendulum will simply sway back and forth periodically from one side to the other. If ϕ is larger, and if we have some frictional effects, or if we include any type of forcing then the pendulum can actually have some quite complicated evolution.

Remark 16.1.6. For the purposes of these notes being used in a Mathematics course, we will not focus on the derivation of the potential and kinetic energy in different settings. In all that follows we will only use two different types of potential energy (unless noted explicitly otherwise):

- gravitational potential energy.
- Hooke's law for a spring.

Similarly we will only require knowledge of how to compute the Cartesian kinetic energy that is

$$T = \frac{m}{2}\left[(x')^2 + (y')^2 + (z')^2\right],$$

in three dimensions.

Now to see the real power of Hamilton's principle, and why we really do get excited about it, we expand on this simple pendulum to demonstrate what can happen if we have several different interactions going on simultaneously.

Example 16.1.7 (Pirates on a cruise). The cruise liner 'Adventurous' has an acrobatic circus act as part of its Caribbean cruise. While the show is going on, a group of pirates approach the ship. Without warning, the captain is forced to begin accelerating away from the pirates at a very rapid pace. At this point in time, one of the trapeze artists from the acrobatic group was swinging from the high trapeze (in this case they are swinging from a stiff rod rather than a rope). Now, his/her colleagues must determine the swinging motion of this performer in order to safely adjust the nets in case he/she falls (they had not planned on the rapid acceleration of the boat before). See Figure 16.3 for some idea of the picture here.

Let positive values of x be the direction of acceleration of the boat with constant acceleration a and specifically let $x(t)$ be the position of the trapeze artist with respect to the initial position i.e. $x(0) = 0$. We will assume for now that this is in the same direction as the swinging of the trapeze artist (this would be quite lucky actually, although the general case could be worked out without to much trouble), i.e. this could be the case if the performer was at rest before the ship started to accelerate. Let y be the vertical direction in which gravity is acting. We further assume that $x'(0) = v_0$ is the initial velocity of the boat for a fixed cartesian coordinate system, and we suppose that the trapeze artist is connected to a wire of length l, and is hanging at an angle $\theta(t)$ relative to the direction of gravity. Then it follows that the horizontal position of the boat $x_b(t)$ will satisfy:

$$x_b'' = a, \quad x_b(0) = 0, \quad x_b'(0) = v_0,$$

which has the explicit solution:

$$x_b(t) = v_0 t + \frac{1}{2}at^2.$$

To get the horizontal position of the trapeze artist, we just add on the distance he is swinging,

$$x = v_0 t + \frac{1}{2}at^2 + l\sin\theta.$$

It follows that the trapeze artist's horizontal velocity is given by:

$$x'(t) = v_0 + at + \theta' l\cos\theta.$$

In a similar manner, we can describe the vertical motion (we will ignore the effects of waves and swells on the boat's motion) by:

$$y = -l\cos\theta$$
$$y'(t) = l\theta'\sin\theta.$$

At this point we can compute the kinetic energy as

$$T = \frac{1}{2}m(x'^2 + y'^2),$$

and the potential energy $U = mg(y + l)$, and hence the Lagrangian is given by

$$L = T - U = \frac{1}{2}m(v_0 + at + l\theta'\cos\theta)^2 + \frac{1}{2}m(l\theta'\sin\theta)^2 + mgl(\cos\theta - 1).$$

After some work, the (EL) will become

$$\theta'' = \frac{-g}{l}\sin\theta - \frac{a}{l}\cos\theta.$$

To catch the trapeze artist, the net should be placed directly below the performer when he is in equilibrium. In the absence of the boat accelerating, equilibrium occurs when $\theta'' = 0$, so that if the angle θ_e is the new angle of equilibrium (taking the accelerating boat into account), then it turns out that $\tan\theta_e = \frac{-a}{g}$ describes this angle (as shown in the exercises). This would allow the other performers to determine (knowing the length of the line) a good approximate location to place the nets by making the approximation that $\theta = \theta_e + \phi$ where θ_e is fixed, i.e. $\theta' = \phi'$ and allowing for ϕ to be 'small'.

A less physically motivated, but nevertheless interesting example is when we consider a particle constrained to live on a certain surface under the action of a particularly symmetric force field. This actually has very many interesting applications in nuclear physics and quantum mechanics, but we don't want to take to much space focusing on the details here. Instead, consider the very simplified following example.

Example 16.1.8 (Central Force Field). Now consider a central force field (an example may be something like gravity that depends only on the distance but not location between two masses) in 2 dimensions with a force proportional to the distance from the origin, i.e.

$$F = -\frac{k}{r^2}$$

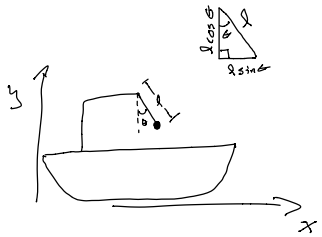


Figure 16.3: Magnificently sketched image of an acrobat on an accelerating boat.

where

$$r = \sqrt{x^2 + y^2}.$$

It follows that the potential energy is given by

$$U = -\int F(r)dr = -\frac{k}{r}.$$

In addition, in two dimensional Cartesian coordinates, the kinetic energy is going to be given by

$$T = \frac{1}{2}\left[(x')^2 + (y')^2\right] = \frac{1}{2}(r')^2 + \frac{1}{2}r^2(\theta')^2.$$

where $x = r\cos\theta$ and $y = r\sin\theta$. This leads to a Lagrangian given by

$$L = T - U = \frac{1}{2}(r')^2 + \frac{1}{2}r^2(\theta')^2 + \frac{k}{r}.$$

16.2 Hamiltonian description

Up to this point we have not been terribly precise in our use of Hamilton's principle. Even so, it has worked remarkably well, and provided us with a host of interesting examples. Before we convince ourselves that we are armed with all of the tools necessary to conquer all of classical physics, we need to formalize our description of these physical systems better. To begin, we require the following definition.

Definition 16.2.1. *The generalized coordinates of a dynamical system, $y_1(t), y_2(t), ..., y_n(t)$, are coordinates that are functions of time and completely describe the motion of the system at any time t. In addition the $y_i(t)$ are completely independent (not just linearly independent). In this context, $y_i'(t)$ is usually referred to as the generalized velocity.*

Example 16.2.2. The motion of the pendulum can be governed by its Cartesian coordinates $(x(t), y(t))$, but in the case of a pendulum with constant length, these are *not* independent coordinates because $x(t)^2 + y(t)^2 = l^2$ where l is the length of the pendulum. That is why we prefer to use the angle $\theta(t)$ from the vertical for the pendulum, because this is a single coordinate, it is clearly independent.

Remark 16.2.3. This last example illustrates an important point. We may have a dynamical system that can be described by coordinates $y_1(t)$ and $y_2(t)$ for example, but we make use of a Lagrange multiplier to enforce some type of relationship between $y_1(t)$ and $y_2(t)$ which would mean that these coordinates are NOT generalized coordinates. If we can choose a set of coordinates that do NOT require the use of a Lagrange multiplier to fully describe the system then we very likely have a set of generalized coordinates.

Generically, T (the kinetic energy) is quadratic in the $y_i'(t)$ (think of all the examples of the previous section), and U is independent of the $y_i'(t)$, but likely depends on $y_i(t)$ and t. With these variables, the (EL) become

$$L_{y_i} - \frac{d}{dt} L_{y_i'} = 0$$

for each i. If there is no explicit time dependence in L then the first integral described in Section 13.2.1 indicates that

$$L - \sum_{i=1}^{n} y_i' L_{y_i'} = -H,$$

where H is a constant that is defined as the Hamiltonian of the system which in most cases is equal to the total energy of the system (up to a sign). Oftentimes you may see the Hamiltonian defined with the opposite sign to this definition, i.e.

$$\sum_{i=1}^{n} y_i' L_{y_i'} - L = \tilde{H},$$

but since it is a constant if L is independent of t then this is just a matter of convention. We have chosen this particular form of the Hamiltonian because it will pop up again in our discussion on optimal control (spoiler alert!!).

16.2.1 Canonical Variables

Now that we have the definition of the Hamiltonian in place, we want to work with not just generalized coordinates, but a more specific set of generalized coordinates that we refer to as canonical variables. Although the formulas that are produced are rather nice, there are far better reasons to investigate Hamiltonian systems (systems that can be described completely by Hamilton's equations derived below). Most of the arguments explaining why this is the case rely on the physical nature of the underlying system that is being modeled, so we don't dwell on that here. For now, just believe that Hamiltonian systems occur frequently in nature and are studied extensively by physicists (just ask the physicist in the room, or if there isn't one, go find one).

Definition 16.2.4. *For a Lagrangian $L(t, y_1, \ldots, y_n, y_i', \ldots, y_n')$, the canonical momenta p_i are defined as $p_i = L_{y_i'}$.*

So long as the Hessian $\frac{D^2 L}{D(\mathbf{y}')^2}$ matrix is invertible, then the Implicit function theorem guarantees that there is a function ζ_i such that $y_i' = \zeta_i(t, y_1, \ldots, y_n, p_1, \ldots, p_n)$ so that we can easily transition between the canonical momenta and the $y_i'(t)$. If the explicit time dependence goes away, then the Hamiltonian is defined as

$$H = -L(y_1, \ldots, y_n, \zeta_1, \ldots, \zeta_n) + \sum_{i=1}^{n} \zeta_i p_i.$$

Example 16.2.5. Recall that for the spring with mass m,

$$L = T - U = \frac{1}{2}mx'^2 - \frac{1}{2}kx^2,$$

so the canonical momentum is

$$p = L_{x'} = mx'$$

which is the physical momentum of the system. Thus in this case the physical and canonical momenta agree. We aren't always so lucky, and often the canonical momentum has no physical interpretation.

Now we return to the Hamiltonian. We rewrite it in terms of the canonical coordinates as is done above, then taking the derivative with respect to p_k we obtain

$$\frac{\partial H}{\partial p_k} = -\sum_{i=1}^{n}\left[\frac{\partial L}{\partial y_i'}\frac{\partial \zeta_i}{\partial p_k}\right] + \zeta_k + \sum_{i=1}^{n}\frac{\partial L}{\partial y_i'}\frac{\partial \zeta_i}{\partial p_k} = \zeta_k = y_k'.$$

For $L = L[y_1, \ldots, y_n, \zeta_1(y_1, \ldots, y_n, p_1, \ldots, p_n), \zeta_2, \ldots, \zeta_n]$ then we can take the derivative of the Lagrangian with respect to y_k as

$$\frac{dL}{dy_k} = \frac{\partial L}{\partial y_k} + \sum_{i=1}^{n}\frac{\partial L}{\partial \zeta_i}\frac{\partial \zeta_i}{\partial y_k} = \frac{\partial L}{\partial y_k} + \sum_{i=1}^{n}\frac{\partial L}{\partial y_i'}\frac{\partial \zeta_i}{\partial y_k}.$$

Using this, we find that

$$\frac{\partial H}{\partial y_k} = -\frac{\partial L}{\partial y_k} - \sum_{i=1}^{n}\frac{\partial L}{\partial y_i'}\frac{\partial \zeta_i}{\partial y_k} + \sum_{i=1}^{n}\frac{\partial \zeta_i}{\partial y_k}\frac{\partial L}{\partial y_i'} = -L_{y_k}.$$

If L does not depend on time (t) explicitly then

$$L_{y_k} - \frac{d}{dt}L_{y_k'} = 0.$$

Thus

$$\frac{\partial H}{\partial y_k} = -L_{y_k} = -\frac{d}{dt}L_{y_k'} = -p_k'.$$

Hence we end up with the system

$$\frac{\partial H}{\partial p_k} = y_k' \qquad \text{and} \qquad -\frac{\partial H}{\partial y_k} = p_k'. \tag{16.1}$$

These are referred to as Hamilton's equations of motion. Originally (in terms of y_k and y_k') we had a 2nd-order nth-dimensional system of ODEs, Now Hamilton's equations give us a 1st-order, 2nth-dimensional system of ODEs. This is one of the primary mathematical advantages of Hamiltonian systems, they can be written as 1st-order systems. Another mathematical advantage will be better understood once we begin working in optimal control wherein we find that the canonical momentum can also be thought of as the adjoint of the canonical position vector and that it satisfies the adjoint equation. There are host of reasons why this is mathematically 'nice', but we don't have the energy (or time) to expound on these all in this space. Of course if you want to be expounded to, then feel free to Google Hamiltonian systems and see what you come up with.

Remark 16.2.6. The real power of Hamiltonian systems comes from the fact that if the kinetic energy T is truly a function of the $y_i'(t)$ but not of the $y_i(t)$ and correspondingly, the potential energy U is a function of the $y_i(t)$ only, then the Hamiltonian $H = T + U$ is the total energy of the system. This arises from Hamilton's principle that states we are hoping to minimize $L = T - U$.

Remark 16.2.7. For the overly curious mind who is trying to figure out why there is an entire chapter of this text dedicated to Physics problems when the pretense of the course designed around this material is Mathematics, we note that Hamiltonian mechanics have a variety of application and use far beyond just the description of physical systems. In fact, Hamiltonian Monte Carlo (google please) is a stunning combination of Hamiltonian dynamics and Markov Chain Monte Carlo sampling methods.

Now we will move on beyond realistic problems, and enter the world of physics designed exercises/examples. Rather than considering something realistic, we will instead move to a more traditional example.

Example 16.2.8. Consider a particle of mass m, constrained to move on a cylinder of radius ρ (hence we can anticipate that cylindrical variables are going to be a good choice), acted on by a force $F = -k\mathbf{r}$ where $\mathbf{r}$ is the vector from the origin to the particle's location. Note that $F = -\nabla U$ where U is the potential energy, i.e.

$$U = \frac{1}{2}kr^2 = \frac{1}{2}k(x^2 + y^2 + z^2) = \frac{1}{2}k(\rho^2 + z^2),$$

which is independent of all components of the momentum, and the kinetic energy is given by

$$T = \frac{1}{2}m(x'^2 + y'^2 + z'^2) = \frac{1}{2}m(\rho^2\theta'^2 \sin^2\theta + \rho^2\theta'^2\cos^2\theta + z'^2) = \frac{1}{2}m(\rho^2\theta'^2 + z'^2).$$

Then the Lagrangian is

$$L = T - U = \frac{1}{2}m(\rho^2\theta'^2 + z'^2) - \frac{1}{2}k(\rho^2 + z^2).$$

The generalized coordinates we have chosen are θ and z so the canonical momenta are:

$$p_\theta = \frac{\partial L}{\partial \theta'} = m\rho^2\theta'$$
$$p_z = \frac{\partial L}{\partial z'} = mz'.$$

The Hamiltonian is given by

$$H(z, p_\theta, p_z) = T + U = \frac{p_\theta^2}{2m\rho^2} + \frac{p_z^2}{2m} + \frac{1}{2}kz^2 + \frac{1}{2}k\rho^2.$$

Note that the final constant term $\frac{1}{2}k\rho^2$ will not influence the actual motion of the particle. This leads to the canonical equations of motion

$$\begin{aligned} p_\theta' &= -\frac{\partial H}{\partial \theta} = 0 \\ p_z' &= -\frac{\partial H}{\partial z} = -kz \\ \theta' &= \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{m\rho^2} \\ z' &= \frac{\partial H}{\partial p_z} = \frac{p_z}{m}. \end{aligned}$$

The first of these equations implies that the angular momentum $p_\theta = m\rho^2\theta'$ is a conserved quantity of the system. This conservation of momentum is guaranteed because the system is symmetric about the z-axis (see the next couple of sections on Noether's Theorem).

The natural symmetry that appears in the previous problem occurs for several other situations such as gravitational forces, electric and magnetic fields, and a variety of other physical settings. Setting the problem up in terms of the Hamiltonian structure with the proper choice of the canonical variables yields better insight into the underlying symmetry of the problem that is not readily apparent from any other approach. There are a variety of reasons for this, but one of them is that if we recall that the definition of the Hamiltonian H is actually the total energy of the system, then we can recognize that the Hamiltonian evolution equations are nothing less than guaranteeing conservation of the total energy.

Example 16.2.9 (electric charges in a magnetic field). For this particular example, don't worry if some of the terminology is rather specific. It should give you an idea of how powerful these ideas are.

Consider a Pym particle (better than just some ordinary particle right?) with constant electric charge e at the location $(x_1, x_2, x_3)^T$ in a medium with electric scalar potential $\phi(\mathbf{x})$, under the influence of a magnetic field (A_1, A_2, A_3). The kinetic energy in this case is

$$T = \frac{m}{2}\sum_{i=1}^{3}(x_i')^2 + e\sum x_i' A_i,$$

where the second term is a result of the magnetic field's interference, and the potential energy is simply $U = e\phi$. The Hamiltonian is then given by

$$H = \sum_{i=1}^{3}\frac{(p_i - eA_i)^2}{2m} + e\phi,$$

with the Lagrangian specified as

$$L = \frac{m}{2}\sum_{i=1}^{3}(x_i')^2 + e\sum_{i=1}^{3} x_i' A_i - e\phi,$$

where

$$p_i = \frac{\partial L}{\partial x_i'} = mx_i' + eA_i.$$

This allows us to completely describe the evolution of the system.

Example 16.2.10 (Poor man's relativity). When developing the Ant Man suit, Hank Pym realized that he needed to consider the effects of relativity. In the relativistic setting, a charged particle will have kinetic energy

$$T = -mc^2\sqrt{1 - \frac{|\mathbf{x}'|^2}{c^2}},$$

and potential energy (similar to the previous example)

$$U = e\phi(\mathbf{x}) - e\mathbf{x}' \cdot \mathbf{A}.$$

Using $L = T - U$, this indicates that the canonical momenta will be given by the vector

$$\mathbf{p} = \frac{DL}{D\mathbf{x}} = \frac{m\mathbf{x}'}{\sqrt{1 - \frac{|\mathbf{x}'|^2}{c^2}}} + e\mathbf{A},$$

with Hamiltonian

$$H = c\sqrt{m^2c^2 + (\mathbf{p} - e\mathbf{A})^2} + e\phi.$$

16.3 Hamiltonian Monte Carlo

Now we are going to switch things up and demonstrate how Hamiltonian systems which arise in Physics give rise to one of the most important random sampling algorithms of the past hundred years. Although to be fair, there aren't many random sampling algorithms that originated more than 100 years ago. In fact, there weren't a lot of computers 100+ years ago so there weren't a lot of algorithms then either...

First, we must necessarily delve into some particulars about sampling algorithms (details and notation will eventually be made completely consistent with Volume III of this series but for now just bear with us for a bit). We will first state the Metropolis-Hastings construction as it is given in the version at the time of writing. We will then describe a more general framework for the Metropolis-Hastings algorithm, and then demonstrate how Hamiltonian dynamics can be used to construct a transition kernel (and hence Markov chain) that has many desirable properties for sampling (none of which we prove here).

16.3.1 Metropolis-Hastings for continuous random variables

To begin, recall Definition 8.2.1 (at least as it is right now)TODO: check this later from Volume III:

Definition 16.3.1 (Metropolis–Hastings construction). *Assume that $\boldsymbol{\pi}$ is a distribution (the* target distribution*) on a set S. Write π_s for the probability of s (or the density of s, if $\boldsymbol{\pi}$ is continuous) under this distribution. Assume that $\pi_s > 0$ for every $s \in S$ (if not, remove all s from S that have $\pi_s = 0$). Let X be a Markov chain with state space S, having transition probabilities $q_{s's} = P(X_{t+1} = s' \,|\, X_t = s)$. We call X the* proposal Markov Chain*. Define a new Markov chain Y on S as follows.*

(*i*) *For $Y_t = s$ choose a* proposal *$s' \in S$ using the Markov chain X; that is, set $X_t = s$ and draw from $X_{t+1} \,|\, X_t = s$ to get a proposal s' (with probability $q_{s's}$).*

(*ii*) *If s' is proposed, then randomly, with probability*

$$a_{s's} = \min\left(\frac{\pi_{s'} q_{ss'}}{\pi_s q_{s's}}, 1\right), \tag{16.2}$$

set $Y_{t+1} = s'$ (accept the proposal), and otherwise set $Y_{t+1} = s$ (reject the proposal). In other words, choose $B \sim \text{Bernoulli}(a_{s's})$ and accept the proposal whenever $B = 1$.

The resulting Markov chain Y is called the Metropolis–Hastings *construction.*

Remark 16.3.2. The following derivation can probably be skipped as a version of it takes place in Volume III...most readers (and students) are probably content with skipping ahead to the subsection labeled 'Metropolis Hastings' below, just prior to Section 16.3.2.

*Deriving the Metropolis-Hastings update

Note that you can likely skip this derivation/notation, and move straight to the description of the algorithm below...this is done here just in case it hasn't been adequately covered elsewhere.

In our setting we will be interested in continuous random variables, and so we are interested in a transition kernel for the Markov Chain constructed via Metropolis-Hastings, i.e. we will use the following notation:

Definition 16.3.3. *A* Markov chain *is a discrete time stochastic process consisting of random variables $\{X_j\}_{j=1}^{\infty}$ such that the probability that $X_{j+1} = x_{j+1}$ conditioned on $X_1 = x_1, \ldots X_j = x_j$ is equal to the same probability conditioned only on $X_j = x_j$. Symbolically this is written as:*

$$\pi_{X_{j+1}}(x_{j+1} \mid x_1, \ldots, x_j) = \pi_{X_{j+1}}(x_{j+1} \mid x_j).$$

Another way to think of Markov chains is that the have only short term (one step) memory.

A *time-homogenous Markov chain* with probability transition kernel P has the additional property that

$$\pi_{X_{j+1}}(x_{j+1} \mid x_j) = P(x_j, x_{j+1})$$

for all j. In other words, the probability transition kernel P does not change with the index j. This means that

$$\pi_{X_{j+1}}(B_{j+1}) = \pi_{X_j} P(B_{j+1}) = \int_{\mathbb{R}^n} P(x_j, B_{j+1}) \pi_{X_j}(x_j)\, dx_j, \tag{16.3}$$

where B_{j+1} is a measurable set referring to the region that the Markov Chain may move to after being at state X_j.

The probability measure π is an *invariant measure* of the transition kernel P if $\pi P = \pi$, meaning that the distribution which defines the random variable X_j at time step j is the same as the distribution that defines the random variable X_{j+1} at time step $j+1$, i.e. the distribution of the points does is not changed following the action of the transition kernel.

Remark 16.3.4. Existence of an invariant measure does not indicate that the transition kernel P doesn't do anything. It only means that the probability distribution doesn't change. For example if the Markov chain had only two states the kernel which moves one state to the next will have an invariant distribution where each state is equally likely.

We say that the transition kernel P is *irreducible* if given any starting point x, the Markov chain generated by P will visit any set of positive measure with nonzero probability. This means that the chain will eventually visit every possible location in the space.

An irreducible transition kernel P is also *periodic* if the Markov chain generated by P stays in a periodic loop forever. Likewise, an irreducible transition kernel P is *aperiodic* if it is not periodic. Finally, we denote $P^{(k)}$ as the transition kernel that propagates the stochastic process k time steps forward, so that

$$P^{(k)}(x_j, B_{j+k}) = \pi_{X_{j+k}}(B_{j+k} \mid x_j).$$

With these definitions, we are now able to state this important result.

Proposition: Let π be a probability measure in $\mathbb{R}^n$ and $\{X_j\}$ a time-homogenous Markov chain with transition kernel P. Suppose that π is an invariant measure of the transition kernel P and that P is irreducible and aperiodic. Then, for all $x \in \mathbb{R}^n$,

$$\lim_{N\to\infty} P^{(N)}(x, B) = \pi(B) \qquad \text{for all } B \in \mathbb{B}.$$

Remark 16.3.5. I'm sure this is discussed ad nauseum in Volume III somewhere, but I'm having a hard time identifying exactly where this part of the discussion occurs. Basically this is probably a lot of repetition, but hopefully the notation will be standardized eventually (after all these books are supposed to be synced).

Thus, we can approximate the probability distribution π by drawing a sequence of points using the correct invariant, irreducible, aperiodic transition kernel P. The question now remains how we can aquire the correct transition kernel. The Metropolis Hastings Algorithm accomplishes this by *constructing* the correct transition kernel at each timestep.

We begin with an arbitrary transition kernel P, and note that we can separate this kernel as [30]

$$P(x, B) = r(x)\chi_B(x) + \int_B \mathscr{K}(x, y)\, dy \tag{16.4}$$

where r is the probability of x remaining where it is, and $\mathscr{K}$ is the probability of moving from x to y. So, if x does not move to B, then $P(x, B) = r(x)$, otherwise we have the intuitive result that $P(x, B) = \int_B \mathscr{K}(x, y)\, dy$. Since the probability of propagating somewhere into $\mathbb{R}^n$ is one, i.e. $P(x, \mathbb{R}^n) = 1$, we can rewrite

$$P(x, \mathbb{R}^n) = r(x)\chi_{\mathbb{R}^n}(x) + \int_{\mathbb{R}^n} \mathscr{K}(x, y)\, dy$$

as

$$1 - r(x) = \int_{\mathbb{R}^n} \mathscr{K}(x, y)\, dy. \tag{16.5}$$

Now, as P was chosen arbitrarily, we wish to modify P to satisfy the conditions of the Proposition stated above. We need the transition kernel P to be invariant, which means we must have

$$\begin{aligned}
\pi P(B) &= \pi \\
&= \int_{\mathbb{R}^n} P(x, B)\pi(x)\, dx && \text{by Equation (16.3)} \\
&= \int_{\mathbb{R}^n} \left(r(x)\chi_B(x) + \int_B \mathscr{K}(x, y)\, dy \right) \pi(x)\, dx && \text{by Equation(16.4)} \\
&= \int_B \left(r(y)\pi(y) + \int_{\mathbb{R}^n} \pi(x)\mathscr{K}(x, y)\, dx \right) dy \\
&= \int_B \pi(y)\, dy
\end{aligned}$$

[30] TODO: I don't like the notation here for $\mathscr{K}$ but it is at least different from the kinetic energy K

for all $B \in \mathbb{B}$. From the last two equalities, we see that

$$\int_B \pi(y)\ dy = \int_B \left(r(y)\pi(y) + \int_{\mathbb{R}^n} \pi(x)\mathscr{K}(x,y)\ dx \right)\ dy$$

which implies that

$$\pi(y)(1 - r(y)) = \pi(y) - r(y)\pi(y) = \int_{\mathbb{R}^n} \pi(x)\mathscr{K}(x,y)\ dx$$

almost everywhere. We can substitute the right hand side of Equation (16.5) (swapping the x and y variables) to obtain

$$\int_{\mathbb{R}^n} \pi(y)\mathscr{K}(y,x)\ dx = \int_{\mathbb{R}^n} \pi(x)\mathscr{K}(x,y)\ dx$$

which is known as the *balance equation*. This condition holds if we satisfy the *detailed balance equation*, given by

$$\pi(y)\mathscr{K}(y,x) = \pi(x)\mathscr{K}(x,y)$$

for all $x, y \in \mathbb{R}^n$. This means that as long as $\mathscr{K}$ satisfies the detailed balance equation, P is invariant. The Metropolis Hastings Algorithm updates $\mathscr{K}$ to ensure the detailed balance equation holds.

To show how this is done, let $\mathscr{K}$ be a given function such that $\int \mathscr{K}(x,y)\ dy = 1$. Then, $\mathscr{K}$ defines a probability transition kernel

$$Q(x,B) = \int_B \mathscr{K}(x,y)\ dy.$$

We call $\mathscr{K}$ the *proposal distribution* or *candidate-generating kernel*. We will correct $\mathscr{K}$ by a factor of $\alpha(x,y)$ so that $\mathscr{K}$ satisfies the detailed balance equation. To do this, we need to compute the appropriate α.

If the detailed balance equation does not hold, then we must have an inequality. Suppose

$$\pi(y)\mathscr{K}(y,x) < \pi(x)\mathscr{K}(x,y)$$

for some $x, y \in \mathbb{R}^n$. We need to compute α so that

$$\pi(y)\alpha(y,x)\mathscr{K}(y,x) = \pi(x)\alpha(x,y)\mathscr{K}(x,y). \tag{16.6}$$

This equality holds if we set

$$\begin{aligned} a(x,y) &= \frac{\pi(y)\mathscr{K}(y,x)}{\pi(x)\mathscr{K}(x,y)} < 1 \\ a(y,x) &= 1. \end{aligned}$$

Similarly, if

$$\pi(y)\mathscr{K}(y,x) > \pi(x)\mathscr{K}(x,y),$$

then we obtain

$$\begin{aligned} a(x,y) &= 1 \\ a(y,x) &= \frac{\pi(x)\mathscr{K}(x,y)}{\pi(y)\mathscr{K}(y,x)} < 1. \end{aligned}$$

Using these two results, we see that the Equation (16.6) is always satisfied if we define

$$\alpha(x,y) = \min\left(1, \frac{\pi(y)\mathscr{K}(y,x)}{\pi(x)\mathscr{K}(x,y)}\right),$$

which is the transition kernel used in the Metropolis-Hastings Algorithm. This is the term with which we correct our proposal kernel $\mathscr{K}$ at each step.

Metropolis Hastings

The Metropolis-Hastings Algorithm proceeds as follows for a specified proposal kernel $\mathscr{K}(x, y)$: starting with an initial sample x, draw a candidate sample y from the proposal kernel $\mathscr{K}$, and accept it as the next step in the stochastic process with probability $\alpha(x, y) = \min\left(1, \frac{\pi(y)\mathscr{K}(y,x)}{\pi(x)\mathscr{K}(x,y)}\right)$, otherwise set the next step equal to the current step.

By the proposition stated above, this stochastic process will converge to the desired probability distribution as the number of samples processed approaches infinity. The Metropolis-Hastings Algorithm permits the adoption of any transition kernel desired, so long as it integrates to 1. One very popular transition kernel is the Random Walk kernel, whose proposal kernel is a symmetric Gaussian centered at the current sample point x, i.e. $\mathscr{K}(x, y) \sim e^{(x-y)^T \Gamma (x-y)}$, where Γ is a (typically diagonal) covariance matrix whose values are hyperparameters of the algorithm. In Random Walk Metropolis-Hastings, the probability transition kernel is symmetric, meaning that $\mathscr{K}(y, x) = \mathscr{K}(x, y)$, so the acceptance probability becomes

$$\alpha_{RW}(x, y) = \min\left(1, \frac{\pi(y)}{\pi(x)}\right).$$

One advantage of Random Walk Metropolis-Hastings is that the acceptance probability is easier to compute, but the obvious drawback to randomly selecting candidate samples (centered around the current sample) is that the rate of acceptance is usually quite poor.[31] A natural question, then, is how to construct a transition kernel that is both easy to compute and draws better candidate samples that are more likely to be accepted.

16.3.2 Hamiltonian Monte Carlo

The big idea behind Hamiltonian Monte Carlo (HMC) is to translate the sampling problem into a physical system and then use Hamiltonian dynamics with a special transition kernel to optimally draw samples from the desired distribution. While the computational cost of drawing each sample is larger than for Random Walk Metropolis-Hastings, the rate of acceptance for candidate samples is much higher with HMC. We proceed as outlined by Neal. TODO: get correct citation

As seen previously in this Chapter, the Hamiltonian $H(q, p)$ is commonly defined as

$$H(q, p) = U(q) + K(p),$$

where $U(q)$ is the *potential energy* (q is the position or state) and $K(p)$ is the *kinetic energy* (p is the momentum variable), typically given by

$$K(p) = p^T M^{-1} p/2$$

where M is a symmetric, positive-definite, "mass matrix." In the Hamiltonian, q is an n-dimensional position vector, and p is an n-dimensional momentum vector. Over time, q and p change in accordance with Hamilton's Equations, given by

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} = [M^{-1}p]_i$$
$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} = -\frac{\partial U}{\partial q_i}.$$

[31] Indeed, the 'standard' acceptance ratio for random walk Metropolis-Hastings is only 23.4% of all samples meaning that over 3/4 of the samples drawn are not retained in the sampling process.

Notice that this system approaches the minimum potential energy U with respect to q (via the negative partial derivative of U). Thus, if we set U to be the negative log probability density of the target distribution of q, then minimizing U with respect to q is equivalent to maximizing the probability that the target density accurately represents q. Thus, we set

$$U(q) = -\log\left[\pi(q)\right]$$

where $\pi(q)$ is the target density.

In statistical mechanics, the canonical distribution given some energy function $E(x)$ has a density function given by

$$\bar{\pi}(x) = \frac{1}{Z}\exp\left(\frac{-E(x)}{T}\right)$$

where Z is a constant for normalization and T is the temperature of the system. In our case, we have as an energy function given by the Hamiltonian $H(q,p)$, and we wish to find the probability density for q and p, so we adopt the canonical distribution with temperature $T = 1$ (other versions of HMC can be done with different temperatures) so that

$$\begin{aligned}\bar{\pi}(q,p) &= \frac{1}{Z}\exp\left(-H(q,p)\right)\\ &= \frac{1}{Z}\exp\left(-U(q) - K(p)\right)\\ &= \frac{1}{Z}\exp\left(-U(q)\right)\exp\left(-K(p)\right),\end{aligned}$$

which defines a joint density on the variables q that we care about, joint with the momentum p. Since we defined $U(q)$ to be the negative log probability density of the posterior, we see that

$$\exp\left(-U(q)\right) = \pi(q),$$

so we finally have

$$\bar{\pi}(q,p) = \frac{1}{Z}\ \pi(q)\ e^{-K(p)}.$$

If the Hamiltonian system is solved exactly then the Hamiltonian will remain constant for all time so that

$$\frac{1}{Z}\exp\left(-H(q,p)\right) = \frac{1}{Z}\exp\left(-H(q^*,p^*)\right)$$

for different values of q, q^* and p, p^* that are evolved from initial values (q,p), that is (q^*,p^*) are solutions of the Hamiltonian system for some time $t > 0$ with initial condition (q,p). This implies that

$$\frac{1}{Z}\ \pi(q)\ e^{-K(p)} = \frac{1}{Z}\ \pi(q^*)\ e^{-K(p^*)}.$$

If we let $\frac{1}{Z}e^{-K(p)}$ act as our probability transition kernel, then this equation satisfies the detailed balance equation. Unfortunately, in practice, the Hamiltonian doesn't remain perfectly constant (because we are using an imperfect numerical discretization to propagate the solution forward in time), but as before, we can add a correcting term into this equation to make the equality hold. Adding $\alpha(q,q^*)$ into the equation so that

$$\pi(q)\ \alpha(q,q^*)\ \frac{1}{Z}e^{-K(p)} = \pi(q^*)\ \alpha(q^*,q)\ \frac{1}{Z}e^{-K(p^*)},$$

we find that this equation holds if we set

$$\alpha(q,q^*) = \min\left(1,\ \frac{\pi(q^*)\ e^{-K(p^*)}}{\pi(q)\ e^{-K(p)}}\right)$$

which becomes our new acceptance probability. In short, HMC is simply the Metropolis-Hastings Algorithm with a special transition kernel.

There is nice intuition behind this transition kernel. Rather than determining the probability of moving given start and end points, the HMC kernel can be thought of as determining the probability of moving given a momentum vector p. Notice that since $K(p) = p^T M^{-1} p/2$, we see that in fact, $p \sim \mathcal{N}(0, M)$, i.e. the momentum follows a normal distribution with covariance given by the mass matrix.

- At each step of the HMC algorithm, we draw an initial sample p from $\mathcal{N}(0, M)$. This indicates that the momentum will most likely be small, but with smaller probability the momentum may be large.
- The momentum and position terms p and q are updated using Hamilton's Equations to reach the final proposed position q^*.
- Large momentum increases the kinetic energy which will push the state closer to the minimum potential energy, i.e. the maximum probability, but the excessive kinetic energy also tends to lead to more significant errors when integrating the Hamiltonian system which produces a lower acceptance rate.
- The main benefit of HMC is that occasional selection of large momentum values will allow the Hamiltonian system to jump to significantly different values and avoid local maxima of the target distribution $\pi(q)$ and yet HMC will still focus its sampling on regions of higher probability.

Time integrators for Hamiltonian systems

TODO: Explain the Leapfrog Method To get the best acceptance rate possible, we need to have an accurate integrator for the Hamiltonian system. This requires a numerical method that matches the physics of the system in question. One can verify (maybe this is a good HW exercise?) that Hamiltonian systems will have all eigenvalues along the imaginary axis requiring a very unique type of numerical integrator. It turns out the 'best' integrator for this setting is the so-called Leapfrog method which is horribly unstable for most problems but works remarkable well for skew-Hermitian systems such as these Hamiltonian systems considered here.

Rather than work through the fully general setting for Leapfrog, we will consider the special case where the mass matrix M is diagonal with entries m_i. In that case the Leapfrog method for a single step is explicitly described as:

$$\begin{aligned} p_i(t + \Delta t/2) &= p_i(t) - \frac{\Delta t}{2} \frac{\partial U}{\partial q_i}(q(t)), \\ q_i(t + \Delta t) &= q_i(t) + (\Delta t) \frac{p_i(t + \Delta t/2)}{m_i}, \\ p_i(t + \Delta t) &= p_i(t + \Delta t/2) - \frac{\Delta t}{2} \frac{\partial U}{q_i}(q(t + \Delta t)). \end{aligned}$$

Essentially we update the momentum half a time step, and then use that to update the state to the full next step, and then update the momentum again. Also, just to make certain the notation above is clear, the first and third equation above indicate that the potential energy (logarithm of the target distribution) is evaluated at q at time t and q at time $t + \Delta t$ respectively.

Summary of Hamiltonian Monte Carlo

The HMC Algorithm proceeds as follows: begin with an initial candidate sample q, and iterate the following:

(i) Draw the sample momentum p from $\mathcal{N}(0, M)$

(ii) Use the Leapfrog method L times (integrating from time 0 to time $T = L\Delta t$ where Δt is the time step size of the Leapfrog integrator, to approximate q^*, p^* using Hamilton's equations

$$\frac{dq_i}{dt} = [M^{-1}p]_i$$
$$\frac{dp_i}{dt} = -\frac{\partial U}{\partial q_i}$$

(iii) Accept q^* as the new sample with probability

$$\alpha(q, q^*) = \min\left(1,\ \exp(U(q) - U(q^*) + K(p) - K(p^*))\right),$$

otherwise, set q as the new sample (note that this acceptance ratio is actually the same as that described above, just reworked for the current format).

One of the drawbacks of HMC is that M, T, Δt, and L must all be selected carefully, as each of these hyperparameters can greatly affect the outcome.

Remark 16.3.6. There are of course other drawbacks to HMC other than the difficulties in tuning the hyper-parameters. The evolution of the underlying Hamiltonian system requires evaluation of the gradient of the target distribution whose entries are $\frac{\partial U}{\partial q_i}$. For simple settings where the target distribution has an explicit formula, this is relatively straightforward (particularly with symbolic programming). For more complicated settings, particularly those motivated by Bayes' Theorem (as is often the case for these sampling problems) this can be a significant difficulty. Luckily recent advances in automatic differentiation and the promulgation of packages that incorporate said automatic differentiation (PyTorch, TensorFlow, JAX), have made this an accessible approach for many complicated sampling problems.

This is yet another reason why computationally intensive Bayesian approaches to inverse problems have only recently been feasible. Massively parallel computational architectures have only recently surfaced to the extent that sampling procedures are feasible, and algorithmically, easy access to automatic differentiation packages is a new advance that allows for more sophisticated sampling algorithms such as HMC.

16.4 Noether's theorem

Several of the examples in the last two sections included the derivation of certain conserved quantities. This conservation is very closely related to underlying symmetries of the problem. In example 16.2.8 the preservation of angular momentum is a direct result of the inherent radial symmetry of both the force acting on the particle (potential energy) and the type of constraint that was enforced. Noether's Theorem quantifies this.

16.4.1 Preliminaries

Consider the generic functional

$$J[\mathbf{y}] = \int_a^b L(x; \mathbf{y}, \mathbf{y}')dt, \tag{16.7}$$

for $\mathbf{y} \in C^2[a, b]$. We will consider the set of transformations

$$x^* = \Phi(x, \mathbf{y}, \mathbf{y}'), \qquad \mathbf{y}^* = \Psi(x, \mathbf{y}, \mathbf{y}'), \tag{16.8}$$

where Ψ is a vector valued set of functions that maps each y_i to a new y_i^*, each defined on the new interval $[a^*, b^*]$.

Definition 16.4.1. *The standard cost functional* (16.7) *is said to be invariant under the transformation* (16.8) *if* $J[\mathbf{y}] = J['\mathbf{y}^*]$, *i.e.*

$$\int_{a^*}^{b^*} L\left(x^*, \mathbf{y}^*, \frac{d\mathbf{y}^*}{dx^*}\right) dx^* = \int_a^b L\left(x, \mathbf{y}, \frac{d\mathbf{y}}{dx}\right) dx.$$

Remark 16.4.2. This remark is actually meaningful and has distinct purpose. It is important to note that the transformation above is given by $y(x) \to y^*(x)$, but not necessarily that $y(x) \to y^*(x^*)$, i.e. this transformation is independent of x^* but not independent of x.

Example 16.4.3. The functional

$$J[y] = \int_a^b (y')^2 dx,$$

is invariant under the transformation $x^* = x + \varepsilon$ with $y^* = y$, where ε is an arbitrary constant. The shifted curve for this transformation is given by $y^* = y(x^* - \varepsilon = y^*(x^*)$ for $x^* \in [a + \varepsilon, b + \varepsilon]$. Hence

$$\begin{aligned} J[y^*] &= \int_{a^*}^{b^*} \left(\frac{dy^*(x)}{dx^*}\right)^2 dx^* \\ &= \int_{a+\varepsilon}^{b+\varepsilon} \left(\frac{dy(x^* - \varepsilon)}{dx^*}\right)^2 dx^* \\ &= \int_a^b \left(\frac{dy(x)}{dx}\right)^2 dx = J[y], \end{aligned}$$

where the final step above is simply due to the change of variables instituted by $x^* = x + \varepsilon$, implying that $dx^* = dx$.

Remark 16.4.4. We used ε in the previous example to parameterize the transformation under consideration. This was not accidental. Typically we think of these transformations as small variations in the system, although 'small' isn't necessary. In the following we will continue to use ε in like manner, and it will become more apparent how this plays a role in the proof of Noether's Theorem itself.

Unexample 16.4.5. The functional

$$J[y] = \int_a^b x(y')^2 dx$$

is *NOT* invariant under the same transformation as considered in the previous example. In fact following the same calculation as in Example 16.4.3 we see that in this case

$$\begin{aligned} J[y^*] &= \int_{a^*}^{b^*} x^* \left(\frac{dy^*}{dx^*}\right)^2 dx^* \\ &= \int_{a+\varepsilon}^{b+\varepsilon} x^* \left(\frac{dy(x^*-\varepsilon)}{dx^*}\right)^2 dx^* \\ &= \int_a^b (x+\varepsilon)\left(\frac{dy(x)}{dx}\right)^2 dx \\ &= J[y] + \varepsilon \int_a^b \left(\frac{dy(x)}{dx}\right)^2 dx \neq J[y]. \end{aligned}$$

Following these examples, we will consider in the following only those transformations that are parameterized by a single parameter ε, i.e.

$$x^* = \Phi(x, \mathbf{y}, \mathbf{y}'; \varepsilon), \qquad \mathbf{y}^* = \Psi(x, \mathbf{y}, \mathbf{y}'; \varepsilon),$$

where in addition we constrain that $\varepsilon = 0$ is the identify transformation, i.e.

$$\Phi(x, \mathbf{y}, \mathbf{y}'; \varepsilon = 0) = x, \qquad \Psi(x, \mathbf{y}, \mathbf{y}'; \varepsilon = 0) = \mathbf{y}.$$

Using this notation we have the following famed result:

Theorem 16.4.6 (Noether's Theorem). *If the functional*

$$J[\mathbf{y}] = \int_a^b L(x, \mathbf{y}, \mathbf{y}')dx$$

is invariant under the one parameter family of transformations as described above for arbitrary a and b then

$$\sum_{i=1}^n \frac{\partial L}{\partial y_i'}\psi_i + \left(L - \sum_{i=1}^n y_i' \frac{\partial L}{\partial y_i'}\right)\phi = \text{constant}, \tag{16.9}$$

along each extremal of $J[\mathbf{y}]$, where

$$\begin{aligned} \phi(x, \mathbf{y}, \mathbf{y}') &= \left.\frac{\partial \Phi(x, \mathbf{y}, \mathbf{y}'; \varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0}, \\ \psi(x, \mathbf{y}, \mathbf{y}') &= \left.\frac{\partial \Psi(x, \mathbf{y}, \mathbf{y}'; \varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0}. \end{aligned}$$

Remark 16.4.7. Before proceeding to the proof of Noether's Theorem, we note that (16.9) is referred to as the Noether charge of the one-parameter family Φ, Ψ, that is the Noether charge is the conserved quantity for this specific system under this specific symmetry.

Proof. First we suppose that ε is sufficiently small so that the family of transformations is well approximated by

$$\begin{aligned} x^* &= \Phi(x, \mathbf{y}, \mathbf{y}'; 0) + \varepsilon\phi(x, \mathbf{y}, \mathbf{y}') + o(\varepsilon), \\ \mathbf{y}^* &= \Psi_i(x, \mathbf{y}, \mathbf{y}'; 0) + \varepsilon\psi(x, \mathbf{y}, \mathbf{y}') + o(\varepsilon), \end{aligned}$$

for each $i = 1, \dots n$. Now, before proceeding it is important that we clarify what this means. Note that in this sense, $x^* = x + \varepsilon\phi$ and $\mathbf{y}^* = \mathbf{y} + \varepsilon\boldsymbol{\psi}$ indicates that the coordinates x and $\mathbf{y}$ are modified along the curve $\mathbf{y}(x)$ that represents the extremal of $J[\mathbf{y}]$, that is $\boldsymbol{\psi}$ should not be thought of as a function of x added to $\mathbf{y}(x)$ (as was the case for variations $\mathbf{h}(x)$ we have considered before). Indeed, if we wanted to consider functional variations $\mathbf{h}(x)$ added to the extrema $\mathbf{y}(x)$ (for the same parameter ε), then these would be related to $\boldsymbol{\psi}$ via

$$\mathbf{y}(x + \varepsilon\phi) + \varepsilon\mathbf{h}(x + \varepsilon\phi) = \mathbf{y}(x) + \varepsilon\boldsymbol{\psi} + o(\varepsilon), \tag{16.10}$$

i.e. the function $\mathbf{h}(x)$ is capturing the relationship between the changes in x (represented by ϕ) and the changes in the 'vertical' coordinate $\mathbf{y}$ (represented by $\boldsymbol{\psi}$). Such a transformation allows us to consider the Gateaux differential with respect to this transformation, so long as we recall that (16.10) adequately captures this transformation everywhere except at the edge of the domain $x = a, b$. In other words the invariance of $J[\mathbf{y}]$ to this transformation implies that

$$J[\mathbf{y}] = \int_{a+\varepsilon\phi}^{b+\varepsilon\phi} L(x^*, \mathbf{y} + \varepsilon\boldsymbol{\psi}, \mathbf{y}' + \boldsymbol{\psi}')dx^* = \int_{a+\varepsilon\phi}^{b+\varepsilon\phi} L(x, \mathbf{y} + \varepsilon\mathbf{h}, \mathbf{y}' + \varepsilon\mathbf{h}')dx.$$

Now, because $J[\mathbf{y}]$ is invariant under this family of transformations, it follows that $\frac{\partial J}{\partial \varepsilon} = 0$ for all values of ε including $\varepsilon = 0$, that is following the standard steps as we have done multiple times by now:

$$\begin{aligned}
0 = \delta J[\mathbf{y}; \mathbf{h}] &= \frac{\partial}{\partial \varepsilon}\left|\int_{a+\varepsilon\phi}^{b+\varepsilon\phi} L(x, \mathbf{y} + \varepsilon\mathbf{h}, \mathbf{y}' + \varepsilon\mathbf{h}')dx\right|_{\varepsilon=0} \\
&= \int_a^b \sum_{i=1}^n \left[L_{y_i} h_i + L_{y_i'} h_i'\right] dx + \left. L\phi\right|_{x=a}^{x=b} \\
&= \int_a^b \sum_{i=1}^n \left[L_{y_i} - \frac{\partial}{\partial x_i} L_{y_i'}\right] dx + \left[\sum_{i=1}^n L_{y_i'} h + L\phi\right]_{x=a}^{x=b} \\
&= \int_a^b \sum_{i=1}^n \left[L_{y_i} - \frac{\partial}{\partial x_i} L_{y_i'}\right] dx + \left[\sum_{i=1}^n L_{y_i'} \psi_i + \left(L - \sum_{i=1}^n y_i' L_{y_i'}\right)\phi\right]_{x=a}^{x=b}.
\end{aligned}$$

where the last line is taken by expanding (16.10) (much like we did for the transversality boundary condition) to come to the relationship

$$\mathbf{y}(x) + \varepsilon\phi\mathbf{y}'(x) + \varepsilon\mathbf{h}(x) = \mathbf{y}(x) + \varepsilon\boldsymbol{\psi} + o(\varepsilon) \Rightarrow \mathbf{h}(x) \approx \boldsymbol{\psi} - \phi\mathbf{y}'(x).$$

Because we have assumed that $\mathbf{y}(x)$ is an extrema of $J[\mathbf{y}]$ then we note that the (EL) are necessarily satisfied, and hence the integral term above vanishes indicating that

$$\left[\sum_{i=1}^n L_{y_i'} \psi_i + \left(L - \sum_{i=1}^n y_i' L_{y_i'}\right)\phi\right]_{x=a}^{x=b} = 0,$$

which is true for arbitrary a and b (from our choice of defining the invariance above) which in turn implies the desired result. □

Example 16.4.8. Returning to Example 16.2.8 where we considered the motion of a particle on a cylinder with potential $U = \frac{1}{2}kr^2$ in cylindrical coordinates, first recall that the Lagrangian was given by:

$$L = \frac{1}{2}m(\rho^2\theta'^2 + z'^2) - \frac{1}{2}k(z^2 + \rho^2).$$

Note $\psi_\varepsilon(z, \theta) = (z, \theta + \varepsilon)$ is a 1-parameter group of transformations for which L is invariant (and hence so is the integral of L). From this one parameter group, we define the Noether charge as

$$\begin{aligned} I &= \frac{\partial L}{\partial z'}\frac{\psi_{\varepsilon,1}}{\partial \varepsilon} + \frac{\partial L}{\partial \theta'}\frac{\psi_{\varepsilon,2}}{\partial \varepsilon} \\ &= m\rho^2\theta' \end{aligned}$$

which is exactly the angular momentum.

Remark 16.4.9. Note in the previous example that we did not even derive the equations of motion for this problem without discovering a conserved quantity. Noether's Theorem is powerful because once we know the symmetries (invariant transformations) of the problem, we can determine the relevant conserved quantities without deriving the Euler-Lagrange equations let alone solving them.

Example 16.4.10. Suppose that we are now interested in a particle (or some more fascinating object such as Ant-man) that is governed (in cylindrical coordinates) by a potential (which then determines the potential energy) that can be described as $U(\rho, \theta, z) = V(\rho, a\theta + z)$ where a is a constant (with dimensions of length) and the object in question has mass m. Note that this θ here is the angle along the horizontal plane, and is not itself the symmetry of the system. The kinetic energy is computed exactly as we have done in the past (only in cylindrical coordinates now) so that the Lagrangian becomes

$$L = \frac{1}{2}m\left(\rho'^2 + \rho^2\theta'^2 + z'^2\right) - V(\rho, a\theta + z).$$

This system posses the one-parameter group of transformations defined by

$$\begin{aligned} \rho(\varepsilon) &= \rho \\ \theta(\varepsilon) &= \theta + \varepsilon \\ z(\varepsilon) &= z - \varepsilon a. \end{aligned}$$

The Noether charge for this symmetry is defined by:

$$\begin{aligned} I &= \frac{\partial L}{\partial \rho'}\left.\frac{\partial \rho(\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} + \frac{\partial L}{\partial \theta'}\left.\frac{\partial \theta(\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} + \frac{\partial L}{\partial z'}\left.\frac{\partial z(\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0} \\ &= m\rho^2\theta' - maz'. \end{aligned}$$

One can check (see the homework) that this quantity is indeed conserved for this system meaning that $\frac{dI}{dt} = 0$.

Remark 16.4.11. So far we have only discussed finding the conserved quantity if the symmetry is known. The reverse is possible, although the general derivation of a symmetry group from the known conserved quantity is beyond the scope of the current text. We suffice to say that such a calculation can be performed, and is fundamentally of interest in many situations.

Exercises

Note to the student: Each section of this chapter has several corresponding exercises, all collected here at the end of the chapter. The exercises between the first and second line are for Section 1, the exercises between the second and third lines are for Section 2, and so forth.

You should **work every exercise** (your instructor may choose to let you skip some of the advanced exercises marked with *). We have carefully selected them, and each is important for your ability to understand subsequent material. Many of the examples and results proved in the exercises are used again later in the text. Exercises marked with ⚠ are especially important and are likely to be used later in this book and beyond. Those marked with † are harder than average, but should still be done.

Although they are gathered together at the end of the chapter, we strongly recommend you do the exercises for each section as soon as you have completed the section, rather than saving them until you have finished the entire chapter.

16.1. Using Hamilton's principle, find the classical equation of motion for a mass attached to a spring, i.e. the potential energy of the spring is given by $U = \frac{1}{2}kx^2$ and the kinetic energy is given by $T = \frac{1}{2}m(x')^2$.

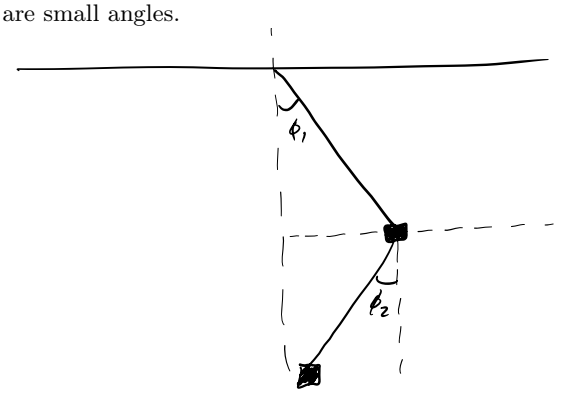
16.2. Solve the differential equation given above if the mass attached to the spring is released without any speed at $x = 1$ at time $t = 0$.

16.3. Consider the plane pendulum problem with a bob of mass m attached to a string of length l. After the pendulum is set in motion the string is shortened by a constant rate $dl/dt = -\alpha$ (a constant). Formulate Hamilton's principle and determine the equation of motion.

16.4. Complete the steps necessary to reduce the Euler-Lagrange equations for the trapeze problem on the cruise ship to the given form in the problem description.

16.5. *For the hanging trapeze aboard the cruise ship, consider oscillations that are small about the equilibrium angle, i.e. $\theta = \theta_e + \phi$ where ϕ is small. Using this ansatz show that the system can be reduced to an oscillation equation, i.e. $\phi'' = -\omega^2\phi$. Find the value of the oscillatory frequency ω in this case.

16.6. A double planar pendulum is simply two pendula each with a bob of the same mass m attached to strings of equal length l, however the second pendulum is attached to the bob of the first. If both pendula are confined to move in the same plane, find the Euler Lagrange equations of motion for the system. Do not assume there are small angles.



16.7. *Show that for a Lagrangian $L = L(y, y')$ independent of the time variable t (the independent variable in this case), the total energy of the system is conserved following Hamilton's Principle where the kinetic energy $T = T(y')$ and the potential energy $U = U(y)$.

16.8. Suppose that the Hamiltonian of a system is given by $H = H(y_1, y_2, \ldots, y_n, p_1, p_2, \ldots p_n)$, that is there is no explicit dependence on time t (the system is autonomous). Show that the Hamiltonian in this setting is conserved.

16.9. Two blocks, each of equal mass m are connected by a uniform, non-elastic, and massless string of length l. One block is placed on a smooth, horizontal surface and the other block is suspended over the side, the string passing through a frictionless pulley. Describe the motion of the system assuming that the string has no mass.

16.10. Repeat the previous problem, only now suppose that the string is uniform in density with total mass m, meaning that if the fraction α of the string is hanging over the edge then this will allow αm of the mass to be influenced by gravity.

16.11. For the previous problem, determine the canonical momenta and corresponding evolution equations for them. Is this an easier approach than the standard (EL)?

16.12. Identify the potential and kinetic energy for a Hamiltonian system given by $H = q^2/2 + p^2/2$. Set up Hamilton's equations for this system, and compute the solution exactly for initial conditions $q(t = 0) = 0$ and $p(t = 0) = 1$.

16.13. Simulate the Hamiltonian system with Hamiltonian $H = q^2/2 + p^2/2$ using the Leapfrog method as described in the text starting with the initial state $q = 0$ and $p = 1$ with a time step size of $\Delta t = 0.3$ and $\Delta t = 1.2$. Run the simulation for 20 total steps. Note that for this Hamiltonian the mass matrix is the one-dimensional identity, i.e. $M = m = 1$. Compare the numerical solution with the exact solution from the previous problem. To visualize the solution, try plotting it in phase space, i.e. a 2D plot of $q(t)$ versus $p(t)$.

16.14. Simulate the same Hamiltonian system from the previous problem (with the same initial conditions and step size etc.), but now use the forward Euler method, i.e.

$$p(t+\Delta t) = p(t) - (\Delta t)\frac{\partial U}{\partial q}(q(t)),$$
$$q(t+\Delta t) = q(t) + (\Delta t)p(t).$$

Comparing these simulations with the previous problem, do you see why the Leapfrog method is better?

16.15. *Implement HMC to sample from the two welled distribution $\pi(x) = \frac{1}{2\sqrt{\pi}}\left(e^{-x^2} + e^{-(x-1)^2}\right)$ which is a multi-peaked distribution that is typically difficult to sample from.

16.16. Consider a Lagrangian $L(t, x, y, x', y') = L(t, y, x', y')$, i.e. independent of the coordinate variable x. Using Noether's Theorem find a conserved quantity and the corresponding family of transformations for this system.

16.17. Let $L = \frac{1}{2}m|\mathbf{x}'|^2 - U(|\mathbf{x}|)$ be a Lagrangian for planar curves $\mathbf{x}(t) : [0, t] \to \mathbb{R}^2$, and where the potential energy U is dependent on the magnitude of the vector $\mathbf{x} = (x, y)^T$ (where $|\cdot|$ refers to the vector l^2 norm). Show that L is invariant under the one-parameter symmetry group $\phi_\varepsilon : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$\phi_\varepsilon(x, y) = (x\cos(\varepsilon) - y\sin(\varepsilon), x\sin(\varepsilon) + y\cos(\varepsilon)).$$

16.18. Find the conserved quantity guaranteed by Noether's theorem for the system described in the previous problem.

16.19. Show that the 1-parameter family of transformations $\psi_\varepsilon(t, \mathbf{y}) = \mathbf{y}(t + \varepsilon a)$ for a specific $a \in \mathbb{R}$ is an invariant family of transformations for the cost functional

$$J[\mathbf{y}] = \int_\alpha^\beta L(\mathbf{y})dt.$$

16.20. Using Noether's Theorem, determine what the conserved quantity for the system considered in the previous problem is.

Notes