Chapter 2

Lagrange's and Hamilton's Equations

In this chapter, we consider two reformulations of Newtonian mechanics, the Lagrangian and the Hamiltonian formalism. The first is naturally associated with configuration space, extended by time, while the latter is the natural description for working in phase space.

Lagrange developed his approach in 1764 in a study of the libration of the moon, but it is best thought of as a general method of treating dynamics in terms of generalized coordinates for configuration space. It so transcends its origin that the Lagrangian is considered the fundamental object which describes a quantum field theory.

Hamilton's approach arose in 1835 in his unification of the language of optics and mechanics. It too had a usefulness far beyond its origin, and the Hamiltonian is now most familiar as the operator in quantum mechanics which determines the evolution in time of the wave function.

We begin by deriving Lagrange's equation as a simple change of coordinates in an unconstrained system, one which is evolving according to Newton's laws with force laws given by some potential. Lagrangian mechanics is also and especially useful in the presence of constraints, so we will then extend the formalism to this more general situation.

2.1 Lagrangian for unconstrained systems

For a collection of particles with conservative forces described by a potential, we have in inertial cartesian coordinates

$$m\ddot{x}_i = F_i.$$

The left hand side of this equation is determined by the kinetic energy function as the time derivative of the momentum $p_i = \partial T/\partial \dot{x}_i$, while the right hand side is a derivative of the potential energy, $-\partial U/\partial x_i$. As T is independent of x_i and U is independent of $\dot{x}_i$ in these coordinates, we can write both sides in terms of the **Lagrangian** $L = T - U$, which is then a function of both the coordinates and their velocities. Thus we have established

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}_i} - \frac{\partial L}{\partial x_i} = 0,$$

which, once we generalize it to arbitrary coordinates, will be known as Lagrange's equation. Note that we are treating L as a function of the $2N$ independent variables x_i and $\dot{x}_i$, so that $\partial L/\partial \dot{x}_i$ means vary one $\dot{x}_i$ holding all the other $\dot{x}_j$ and **all** the x_k fixed. Making this particular combination of $T(\dot{\vec{r}})$ with $U(\vec{r})$ to get the more complicated $L(\vec{r}, \dot{\vec{r}})$ seems an artificial construction for the inertial cartesian coordinates, but it has the advantage of preserving the form of Lagrange's equations for any set of generalized coordinates.

As we did in section 1.3.3, we assume we have a set of generalized coordinates $\{q_j\}$ which parameterize all of coordinate space, so that each point may be described by the $\{q_j\}$ or by the $\{x_i\}$, $i, j \in [1, N]$, and thus each set may be thought of as a function of the other, and time:

$$q_j = q_j(x_1, ...x_N, t) \qquad x_i = x_i(q_1, ...q_N, t). \tag{2.1}$$

We may consider L as a function[1] of the generalized coordinates q_j and $\dot{q}_j$,

[1] Of course we are not saying that $L(x, \dot{x}, t)$ is the same function of its coordinates as $L(q, \dot{q}, t)$, but rather that these are two functions which agree at the corresponding physical points. More precisely, we are defining a new function $\tilde{L}(q, \dot{q}, t) = L(x(q, t), \dot{x}(q, \dot{q}, t), t)$, but we are being physicists and neglecting the tilde. We are treating the Lagrangian here as a *scalar* under coordinate transformations, in the sense used in general relativity, that its value at a given physical point is unchanged by changing the coordinate system used to define that point.

and ask whether the same expression in these coordinates

$$\frac{d}{dt}\frac{\partial L}{\partial \dot q_j} - \frac{\partial L}{\partial q_j}$$

also vanishes. The chain rule tells us

$$\frac{\partial L}{\partial \dot x_j} = \sum_k \frac{\partial L}{\partial q_k}\frac{\partial q_k}{\partial \dot x_j} + \sum_k \frac{\partial L}{\partial \dot q_k}\frac{\partial \dot q_k}{\partial \dot x_j}. \tag{2.2}$$

The first term vanishes because q_k depends only on the coordinates x_k and t, but not on the $\dot x_k$. From the inverse relation to (1.10),

$$\dot q_j = \sum_i \frac{\partial q_j}{\partial x_i}\dot x_i + \frac{\partial q_j}{\partial t}, \tag{2.3}$$

we have

$$\frac{\partial \dot q_j}{\partial \dot x_i} = \frac{\partial q_j}{\partial x_i}.$$

Using this in (2.2),

$$\frac{\partial L}{\partial \dot x_i} = \sum_j \frac{\partial L}{\partial \dot q_j}\frac{\partial q_j}{\partial x_i}. \tag{2.4}$$

Lagrange's equation involves the time derivative of this. Here what is meant is not a partial derivative $\partial/\partial t$, holding the point in configuration space fixed, but rather the derivative along the path which the system takes as it moves through configuration space. It is called the **stream derivative**, a name which comes from fluid mechanics, where it gives the rate at which some property defined throughout the fluid, $f(\vec r, t)$, changes for a fixed element of fluid as the fluid as a whole flows. We write it as a *total* derivative to indicate that we are following the motion rather than evaluating the rate of change at a fixed point in space, as the partial derivative does.

For any function $f(x,t)$ of extended configuration space, this total time derivative is

$$\frac{df}{dt} = \sum_j \frac{\partial f}{\partial x_j}\dot x_j + \frac{\partial f}{\partial t}. \tag{2.5}$$

Using Leibnitz' rule on (2.4) and using (2.5) in the second term, we find

$$\frac{d}{dt}\frac{\partial L}{\partial \dot x_i} = \sum_j \left(\frac{d}{dt}\frac{\partial L}{\partial \dot q_j}\right)\frac{\partial q_j}{\partial x_i} + \sum_j \frac{\partial L}{\partial \dot q_j}\left(\sum_k \frac{\partial^2 q_j}{\partial x_i \partial x_k}\dot x_k + \frac{\partial^2 q_j}{\partial x_i \partial t}\right). \tag{2.6}$$

On the other hand, the chain rule also tells us

$$\frac{\partial L}{\partial x_i} = \sum_j \frac{\partial L}{\partial q_j}\frac{\partial q_j}{\partial x_i} + \sum_j \frac{\partial L}{\partial \dot q_j}\frac{\partial \dot q_j}{\partial x_i},$$

where the last term does not necessarily vanish, as $\dot q_j$ in general depends on both the coordinates and velocities. In fact, from 2.3,

$$\frac{\partial \dot q_j}{\partial x_i} = \sum_k \frac{\partial^2 q_j}{\partial x_i \partial x_k}\dot x_k + \frac{\partial^2 q_j}{\partial x_i \partial t},$$

so

$$\frac{\partial L}{\partial x_i} = \sum_j \frac{\partial L}{\partial q_j}\frac{\partial q_j}{\partial x_i} + \sum_j \frac{\partial L}{\partial \dot q_j}\left(\sum_k \frac{\partial^2 q_j}{\partial x_i \partial x_k}\dot x_k + \frac{\partial^2 q_j}{\partial x_i \partial t}\right). \tag{2.7}$$

Lagrange's equation in cartesian coordinates says (2.6) and (2.7) are equal, and in subtracting them the second terms cancel[2], so

$$0 \;=\; \sum_j \left(\frac{d}{dt}\frac{\partial L}{\partial \dot q_j} - \frac{\partial L}{\partial q_j}\right)\frac{\partial q_j}{\partial x_i}.$$

The matrix $\partial q_j/\partial x_i$ is nonsingular, as it has $\partial x_i/\partial q_j$ as its inverse, so we have derived Lagrange's Equation in generalized coordinates:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot q_j} - \frac{\partial L}{\partial q_j} = 0.$$

Thus we see that Lagrange's equations are form invariant under changes of the generalized coordinates used to describe the configuration of the system. It is primarily for this reason that this particular and peculiar combination of kinetic and potential energy is useful. Note that we implicity assume the Lagrangian itself transformed like a scalar, in that its value at a given physical point of configuration space is independent of the choice of generalized coordinates that describe the point. The change of coordinates itself (2.1) is called a **point transformation**.

[2]This is why we chose the particular combination we did for the Lagrangian, rather than $L = T - \alpha U$ for some $\alpha \neq 1$. Had we done so, Lagrange's equation in cartesian coordinates would have been $\alpha\ d(\partial L/\partial \dot x_j)/dt - \partial L/\partial x_j = 0$, and in the subtraction of (2.7) from $\alpha\times$(2.6), the terms proportional to $\partial L/\partial \dot q_i$ (without a time derivative) would not have cancelled.

2.2 Lagrangian for Constrained Systems

We now wish to generalize our discussion to include contraints. At the same time we will also consider possibly nonconservative forces. As we mentioned in section 1.3.2, we often have a system with internal forces whose effect is better understood than the forces themselves, with which we may not be concerned. We will assume the constraints are holonomic, expressible as k real functions $\Phi_\alpha(\vec{r}_1, ..., \vec{r}_n, t) = 0$, which are somehow enforced by constraint forces $\vec{F}_i^C$ on the particles $\{i\}$. There may also be other forces, which we will call F_i^D and will treat as having a dynamical effect. These are given by known functions of the configuration and time, possibly but not necessarily in terms of a potential.

This distinction will seem artificial without examples, so it would be well to keep these two in mind. In each of these cases the full configuration space is $\mathbb{R}^3$, but the constraints restrict the motion to an allowed subspace of extended configuration space.

1. In section 1.3.2 we discussed a mass on a light rigid rod, the other end of which is fixed at the origin. Thus the mass is constrained to have $|\vec{r}| = L$, and the allowed subspace of configuration space is the surface of a sphere, independent of time. The rod exerts the constraint force to avoid compression or expansion. The natural assumption to make is that the force is in the radial direction, and therefore has no component in the direction of allowed motions, the tangential directions. That is, for all allowed displacements, $\delta\vec{r}$, we have $\vec{F}^C \cdot \delta\vec{r} = 0$, and the constraint force does no work.

2. Consider a bead free to slide without friction on the spoke of a rotating bicycle wheel[3], rotating about a fixed axis at fixed angular velocity ω. That is, for the polar angle θ of inertial coordinates, $\Phi := \theta - \omega t = 0$ is a constraint[4], but the r coordinate is unconstrained. Here the allowed subspace is not time independent, but is a helical sort of structure in extended configuration space. We expect the force exerted by the spoke on the bead to be in the $\hat{e}_\theta$ direction. This is again perpendicular to any **virtual displacement**, by which we mean an allowed change in configuration *at a fixed time*. It is important to distinguish this virtual displacement from a small segment of the trajectory of the particle. In this case a virtual displacement is a change in r without a change in θ, and is perpendicular to $\hat{e}_\theta$. So again, we have the "net virtual work" of the constraint forces is zero. It is important to note that this does not mean that the net real work is zero. In a small time interval, the displacement $\Delta\vec{r}$ includes a component $r\omega\Delta t$ in the tangential direction, and the force of constraint does do work!

[3] Unlike a real bicycle wheel, we are assuming here that the spoke is directly along a radius of the circle, pointing directly to the axle.

[4] There is also a constraint $z = 0$.

We will assume that the constraint forces in general satisfy this restriction that no net *virtual* work is done by the forces of constraint for any possible virtual displacement. Newton's law tells us that $\dot{\vec{p}}_i = F_i = F_i^C + F_i^D$. We can multiply by an arbitrary virtual displacement

$$\sum_i \left(\vec{F}_i^D - \dot{\vec{p}}_i\right) \cdot \delta\vec{r}_i = -\sum_i \vec{F}_i^C \cdot \delta\vec{r}_i = 0,$$

where the first equality would be true even if $\delta\vec{r}_i$ did not satisfy the constraints, but the second requires $\delta\vec{r}_i$ to be an allowed virtual displacement. Thus

$$\sum_i \left(\vec{F}_i^D - \dot{\vec{p}}_i\right) \cdot \delta\vec{r}_i = 0, \tag{2.8}$$

which is known as **D'Alembert's Principle**. This gives an equation which determines the motion on the constrained subspace and does not involve the unspecified forces of constraint F^C. We drop the superscript ^{D} from now on.

Suppose we know generalized coordinates $q_1, \ldots, q_N$ which parameterize the constrained subspace, which means $\vec{r}_i = \vec{r}_i(q_1, \ldots, q_N, t)$, for $i = 1, \ldots, n$, are known functions and the N q's are independent. There are $N = 3n - k$ of these independent coordinates, where k is the number of holonomic constraints. Then $\partial\vec{r}_i/\partial q_j$ is no longer an invertable, or even square, matrix, but we still have

$$\Delta\vec{r}_i = \sum_j \frac{\partial\vec{r}_i}{\partial q_j}\Delta q_j + \frac{\partial\vec{r}_i}{\partial t}\Delta t.$$

For the velocity of the particle, divide this by Δt, giving

$$\vec{v}_i = \sum_j \frac{\partial\vec{r}_i}{\partial q_j}\dot{q}_j + \frac{\partial\vec{r}_i}{\partial t}, \tag{2.9}$$

but for a *virtual* displacement $\Delta t = 0$ we have

$$\delta\vec{r}_i = \sum_j \frac{\partial\vec{r}_i}{\partial q_j}\delta q_j.$$

Differentiating (2.9) we note that,

$$\frac{\partial \vec{v}_i}{\partial \dot{q}_j} = \frac{\partial \vec{r}_i}{\partial q_j}, \tag{2.10}$$

and also

$$\frac{\partial \vec{v}_i}{\partial q_j} = \sum_k \frac{\partial^2 \vec{r}_i}{\partial q_j \partial q_k}\dot{q}_k + \frac{\partial^2 \vec{r}_i}{\partial q_j \partial t} = \frac{d}{dt}\frac{\partial \vec{r}_i}{\partial q_j}, \tag{2.11}$$

where the last equality comes from applying (2.5), with coordinates q_j rather than x_j, to $f = \partial\vec{r}_i/\partial q_j$. The first term in the equation (2.8) stating D'Alembert's principle is

$$\sum_i \vec{F}_i \cdot \delta\vec{r}_i = \sum_j \sum_i \vec{F}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j}\delta q_j = \sum_j Q_j \cdot \delta q_j.$$

The generalized force Q_j has the same form as in the unconstrained case, as given by (1.9), but there are only as many of them as there are *unconstrained* degrees of freedom.

The second term of (2.8) involves

$$\begin{aligned}
\sum_i \dot{\vec{p}}_i \cdot \delta\vec{r}_i &= \sum_i \frac{dp_i}{dt}\frac{\partial \vec{r}_i}{\partial q_j}\delta q_j \\
&= \sum_j \frac{d}{dt}\left(\sum_i \vec{p}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j}\right)\delta q_j - \sum_{ij} p_i \cdot \left(\frac{d}{dt}\frac{\partial \vec{r}_i}{\partial q_j}\right)\delta q_j \\
&= \sum_j \frac{d}{dt}\left(\sum_i \vec{p}_i \cdot \frac{\partial \vec{v}_i}{\partial \dot{q}_j}\right)\delta q_j - \sum_{ij} p_i \cdot \frac{\partial \vec{v}_i}{\partial q_j}\delta q_j \\
&= \sum_j \left[\frac{d}{dt}\sum_i m_i \vec{v}_i \cdot \frac{\partial \vec{v}_i}{\partial \dot{q}_j} - \sum_i m_i v_i \cdot \frac{\partial \vec{v}_i}{\partial q_j}\right]\delta q_j \\
&= \sum_j \left[\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j}\right]\delta q_j,
\end{aligned}$$

where we used (2.10) and (2.11) to get the third line. Plugging in the expressions we have found for the two terms in D'Alembert's Principle,

$$\sum_j \left[\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} - Q_j\right]\delta q_j = 0.$$

We assumed we had a holonomic system and the q's were all independent, so this equation holds for arbitrary virtual displacements δq_j, and therefore

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} - Q_j = 0. \tag{2.12}$$

Now let us restrict ourselves to forces given by a potential, with $\vec{F}_i = -\vec{\nabla}_i U(\{\vec{r}\}, t)$, or

$$Q_j = -\sum_i \frac{\partial \vec{r}_i}{\partial q_j} \cdot \vec{\nabla}_i U = -\left.\frac{\partial \tilde{U}(\{q\}, t)}{\partial q_j}\right|_t.$$

Notice that Q_j depends only on the value of U on the constrained surface. Also, U is independent of the $\dot{q}_i$'s, so

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} + \frac{\partial U}{\partial q_j} = 0 = \frac{d}{dt}\frac{\partial (T-U)}{\partial \dot{q}_j} - \frac{\partial (T-U)}{\partial q_j},$$

or

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0. \tag{2.13}$$

This is Lagrange's equation, which we have now derived in the more general context of constrained systems.

2.2.1 Some examples of the use of Lagrangians

Atwood's machine

Atwood's machine consists of two blocks of mass m_1 and m_2 attached by an inextensible cord which suspends them from a pulley of moment of inertia I with frictionless bearings. The kinetic energy is

$$\begin{aligned}
T &= \frac{1}{2}m_1\dot{x}^2 + \frac{1}{2}m_2\dot{x}^2 + \frac{1}{2}I\omega^2 \\
U &= m_1 g x + m_2 g(K - x) = (m_1 - m_2)gx + \text{const}
\end{aligned}$$

where we have used the fixed length of the cord to conclude that the sum of the heights of the masses is a constant K. We assume the cord does not slip on the pulley, so the angular velocity of the pulley is $\omega = \dot{x}/r$, and

$$L = \frac{1}{2}(m_1 + m_2 + I/r^2)\dot{x}^2 + (m_2 - m_1)gx,$$

and Lagrange's equation gives

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0 = (m_1 + m_2 + I/r^2)\ddot{x} - (m_2 - m_1)g.$$

Notice that we set up our system in terms of only one degree of freedom, the height of the first mass. This one degree of freedom parameterizes the line which is the allowed subspace of the unconstrained configuration space, a three dimensional space which also has directions corresponding to the angle of the pulley and the height of the second mass. The constraints restrict these three variables because the string has a fixed length and does not slip on the pulley. Note that this formalism has permitted us to solve the problem without solving for the forces of constraint, which in this case are the tensions in the cord on either side of the pulley.

Bead on spoke of wheel

As a second example, reconsider the bead on the spoke of a rotating bicycle wheel. In section (1.3.4) we saw that the kinetic energy is $T = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\omega^2$. If there are no forces other than the constraint forces, $U(r,\theta) \equiv 0$, and the Lagrangian is

$$L = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\omega^2.$$

The equation of motion for the one degree of freedom is easy enough:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{r}} = m\ddot{r} = \frac{\partial L}{\partial r} = mr\omega^2,$$

which looks like a harmonic oscillator with a negative spring constant, so the solution is a real exponential instead of oscillating,

$$r(t) = Ae^{-\omega t} + Be^{\omega t}.$$

The velocity-independent term in T acts just like a potential would, and can in fact be considered the potential for the centrifugal force. But we see that the total energy T is not conserved but blows up as $t \to \infty$, $T \sim mB^2\omega^2 e^{2\omega t}$. This is because the force of constraint, while it does no *virtual* work, does do real work.

Mass on end of gimballed rod

Finally, let us consider the mass on the end of the gimballed rod. The allowed subspace is the surface of a sphere, which can be parameterized by an azimuthal angle ϕ and the polar angle with the upwards direction, θ, in terms of which

$$z = \ell\cos\theta, \quad x = \ell\sin\theta\cos\phi, \quad y = \ell\sin\theta\sin\phi,$$

and $T = \frac{1}{2}m\ell^2(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2)$. With an arbitrary potential $U(\theta,\phi)$, the Lagrangian becomes

$$L = \frac{1}{2}m\ell^2(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2) - U(\theta,\phi).$$

From the two independent variables θ, ϕ there are two Lagrange equations of motion,

$$m\ell^2\ddot{\theta} = -\frac{\partial U}{\partial \theta} + \frac{1}{2}\sin(2\theta)\dot{\phi}^2, \qquad (2.14)$$

$$\frac{d}{dt}\left(m\ell^2\sin^2\theta\dot{\phi}\right) = -\frac{\partial U}{\partial \phi}. \qquad (2.15)$$

Notice that this is a dynamical system with two coordinates, similar to ordinary mechanics in two dimensions, except that the mass matrix, while diagonal, is coordinate dependent, and the space on which motion occurs is not an infinite flat plane, but a curved two dimensional surface, that of a sphere. These two distinctions are connected—the coordinates enter the mass matrix because it is impossible to describe a curved space with unconstrained cartesian coordinates.

Often the potential $U(\theta,\phi)$ will not actually depend on ϕ, in which case Eq. 2.15 tells us $m\ell^2\sin^2\theta\dot{\phi}$ is constant in time. We will discuss this further in Section 2.4.1.

2.3 Hamilton's Principle

The configuration of a system at any moment is specified by the value of the generalized coordinates $q_j(t)$, and the space coordinatized by these $q_1, \ldots, q_N$ is the **configuration space**. The time evolution of the system is given by the trajectory, or motion of the point in configuration space as a function of time, which can be specified by the *functions* $q_i(t)$.

One can imagine the system taking many paths, whether they obey Newton's Laws or not. We consider only paths for which the $q_i(t)$ are differentiable. Along any such path, we define the **action** as

$$S = \int_{t_1}^{t_2} L(q(t), \dot{q}(t), t)dt. \tag{2.16}$$

The action depends on the starting and ending points $q(t_1)$ and $q(t_2)$, but beyond that, the value of the action depends on the path, unlike the work done by a conservative force on a point moving in ordinary space. In fact, it is exactly this dependence on the path which makes this concept useful — Hamilton's principle states that the actual motion of the particle from $q(t_1) = q_i$ to $q(t_2) = q_f$ is along a path $q(t)$ for which the action is stationary. That means that for any small deviation of the path from the actual one, keeping the initial and final configurations fixed, the variation of the action vanishes to first order in the deviation.

To find out where a differentiable function of one variable has a stationary point, we differentiate and solve the equation found by setting the derivative to zero. If we have a differentiable function f of several variables x_i, the first-order variation of the function is $\Delta f = \sum_i (x_i - x_{0i})\, \partial f/\partial x_i|_{x_0}$, so unless $\partial f/\partial x_i|_{x_0} = 0$ for all i, there is some variation of the $\{x_i\}$ which causes a first order variation of f, and then x_0 is not a stationary point.

But our action is a **functional**, a function of functions, which represent an infinite number of variables, even for a path in only one dimension. Intuitively, at each time $q(t)$ is a separate variable, though varying q at only one point makes $\dot{q}$ hard to interpret. A rigorous mathematician might want to describe the path $q(t)$ on $t \in [0, 1]$ in terms of Fourier series, for which $q(t) = q_0 + q_1 t + \sum_{n=1} a_n \sin(n\pi t)$. Then the functional $S(f)$ given by

$$S = \int f(q(t), \dot{q}(t), t)dt$$

becomes a function of the infinitely many variables $q_0, q_1, a_1, \ldots$. The endpoints fix q_0 and q_1, but the stationary condition gives an infinite number of equations $\partial S/\partial a_n = 0$.

It is not really necessary to be so rigorous, however. Under a change $q(t) \to q(t) + \delta q(t)$, the derivative will vary by $\delta\dot{q} = d\,\delta q(t)/dt$, and the functional S will vary by

$$\begin{aligned} \delta S &= \int \left(\frac{\partial f}{\partial q}\delta q + \frac{\partial f}{\partial \dot{q}}\delta\dot{q} \right) dt \\ &= \left. \frac{\partial f}{\partial \dot{q}}\delta q \right|_i^f + \int \left[\frac{\partial f}{\partial q} - \frac{d}{dt}\frac{\partial f}{\partial \dot{q}} \right] \delta q dt, \end{aligned}$$

where we integrated the second term by parts. The boundary terms each have a factor of δq at the initial or final point, which vanish because Hamilton tells us to hold the q_i and q_f fixed, and therefore the functional is stationary if and only if

$$\frac{\partial f}{\partial q} - \frac{d}{dt}\frac{\partial f}{\partial \dot{q}} = 0 \quad \text{for } t \in (t_i, t_f) \tag{2.17}$$

We see that if f is the Lagrangian, we get exactly Lagrange's equation. The above derivation is essentially unaltered if we have many degrees of freedom q_i instead of just one.

2.3.1 Examples of functional variation

In this section we will work through some examples of functional variations both in the context of the action and for other examples not directly related to mechanics.

The falling particle

As a first example of functional variation, consider a particle thrown up in a uniform gravitional field at $t = 0$, which lands at the same spot at $t = T$. The Lagrangian is $L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - mgz$, and the boundary conditions are $x(t) = y(t) = z(t) = 0$ at $t = 0$ and $t = T$. Elementary mechanics tells us the solution to this problem is $x(t) = y(t) \equiv 0$, $z(t) = v_0 t - \frac{1}{2}gt^2$ with $v_0 = \frac{1}{2}gT$. Let us evaluate the action for any other path, writing $z(t)$ in terms of its deviation from the suspected solution,

$$z(t) = \Delta z(t) + \frac{1}{2}gTt - \frac{1}{2}gt^2.$$

We make no assumptions about this path other than that it is differentiable and meets the boundary conditions $x = y = \Delta z = 0$ at $t = 0$ and at $t = T$.

The action is

$$S = \int_0^T \left\{ \frac{1}{2}m \left[\dot{x}^2 + \dot{y}^2 + \left(\frac{d\Delta z}{dt}\right)^2 + g(T-2t)\frac{d\Delta z}{dt} + \frac{1}{4}g^2(T-2t)^2 \right] \right.$$
$$\left. -mg\Delta z - \frac{1}{2}mg^2 t(T-t) \right\} dt.$$

The fourth term can be integrated by parts,

$$\int_0^T \frac{1}{2}mg(T-2t)\frac{d\Delta z}{dt}\,dt = \frac{1}{2}mg(T-2t)\Delta z\Big|_0^T + \int_0^T mg\Delta z(t)\,dt.$$

The boundary term vanishes because $\Delta z = 0$ where it is evaluated, and the other term cancels the sixth term in S, so

$$S = \int_0^T \frac{1}{2}mg^2 \left[\frac{1}{4}(T-2t)^2 - t(T-t)\right] dt$$
$$+ \int_0^T \frac{1}{2}m \left[\dot{x}^2 + \dot{y}^2 + \left(\frac{d\Delta z}{dt}\right)^2 \right].$$

The first integral is independent of the path, so the minimum action requires the second integral to be as small as possible. But it is an integral of a non-negative quantity, so its minimum is zero, requiring $\dot{x} = \dot{y} = d\Delta z/dt = 0$. As $x = y = \Delta z = 0$ at $t = 0$, this tells us $x = y = \Delta z = 0$ at all times, and the path which minimizes the action is the one we expect from elementary mechanics.

Is the shortest path a straight line?

The calculus of variations occurs in other contexts, some of which are more intuitive. The classic example is to find the shortest path between two points in the plane. The length ℓ of a path $y(x)$ from (x_1, y_1) to (x_2, y_2) is given[5] by

$$\ell = \int_{x_1}^{x_2} ds = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\,dx.$$

[5]Here we are assuming the path is monotone in x, without moving somewhere to the left and somewhere to the right. To prove that the straight line is shorter than other paths which might not obey this restriction, do Exercise 2.2.

We see that length ℓ is playing the role of the action, and x is playing the role of t. Using $\dot{y}$ to represent dy/dx, we have the integrand $f(y, \dot{y}, x) = \sqrt{1 + \dot{y}^2}$, and $\partial f/\partial y = 0$, so Eq. 2.17 gives

$$\frac{d}{dx}\frac{\partial f}{\partial \dot{y}} = \frac{d}{dx}\frac{\dot{y}}{\sqrt{1+\dot{y}^2}} = 0, \quad \text{so } \dot{y} = \text{const.}$$

and the path is a straight line.

2.4 Conserved Quantities

2.4.1 Ignorable Coordinates

If the Lagrangian does not depend on one coordinate, say q_k, then we say it is an **ignorable coordinate**. Of course, we still want to solve for it, as its derivative may still enter the Lagrangian and effect the evolution of other coordinates. By Lagrange's equation

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} = \frac{\partial L}{\partial q_k} = 0,$$

so if in general we define

$$P_k := \frac{\partial L}{\partial \dot{q}_k},$$

as the **generalized momentum**, then in the case that L is independent of q_k, P_k is conserved, $dP_k/dt = 0$.

Linear Momentum

As a very elementary example, consider a particle under a force given by a potential which depends only on y and z, but not x. Then

$$L = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) - U(y, z)$$

is independent of x, x is an ignorable coordinate and

$$P_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x}$$

is conserved. This is no surprize, of course, because the force is $F = -\nabla U$ and $F_x = -\partial U/\partial x = 0$.

Note that, using the definition of the generalized momenta

$$P_k = \frac{\partial L}{\partial \dot{q}_k},$$

Lagrange's equation can be written as

$$\frac{d}{dt} P_k = \frac{\partial L}{\partial q_k} = \frac{\partial T}{\partial q_k} - \frac{\partial U}{\partial q_k}.$$

Only the last term enters the definition of the generalized force, so if the kinetic energy depends on the coordinates, as will often be the case, it is not true that $dP_k/dt = Q_k$. In that sense we might say that the generalized momentum and the generalized force *have not been defined consistently.*

Angular Momentum

As a second example of a system with an ignorable coordinate, consider an axially symmetric system described with inertial polar coordinates (r, θ, z), with z along the symmetry axis. Extending the form of the kinetic energy we found in sec (1.3.4) to include the z coordinate, we have $T = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 + \frac{1}{2}m\dot{z}^2$. The potential is independent of θ, because otherwise the system would not be symmetric about the z-axis, so the Lagrangian

$$L = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 + \frac{1}{2}m\dot{z}^2 - U(r, z)$$

does not depend on θ, which is therefore an ignorable coordinate, and

$$P_\theta := \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta} = \text{constant}.$$

We see that the conserved momentum P_θ is in fact the z-component of the angular momentum, and is conserved because the axially symmetric potential can exert no torque in the z-direction:

$$\tau_z = -\left(\vec{r} \times \vec{\nabla} U\right)_z = -r\left(\vec{\nabla} U\right)_\theta = -r^2 \frac{\partial U}{\partial \theta} = 0.$$

Finally, consider a particle in a spherically symmetric potential in spherical coordinates. In section (3.1.2) we will show that the kinetic energy in spherical coordinates is $T = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 + \frac{1}{2}mr^2\sin^2\theta\dot{\phi}^2$, so the Lagrangian with a spherically symmetric potential is

$$L = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 + \frac{1}{2}mr^2\sin^2\theta\dot{\phi}^2 - U(r).$$

Again, ϕ is an ignorable coordinate and the conjugate momentum P_ϕ is conserved. Note, however, that even though the potential is independent of θ as well, θ does appear undifferentiated in the Lagrangian, and it is *not* an ignorable coordinate, nor is P_θ conserved[6].

If q_j is an ignorable coordinate, not appearing undifferentiated in the Lagrangian, any possible motion $q_j(t)$ is related to a different trajectory $q'_j(t) = q_j(t) + c$, in the sense that they have the same action, and if one is an extremal path, so will the other be. Thus there is a symmetry of the system under $q_j \to q_j + c$, a continuous symmetry in the sense that c can take on any value. As we shall see in Section 8.3, such symmetries generally lead to conserved quantities. The symmetries can be less transparent than an ignorable coordinate, however, as in the case just considered, of angular momentum for a spherically symmetric potential, in which the conservation of L_z follows from an ignorable coordinate ϕ, but the conservation of L_x and L_y follow from symmetry under rotation about the x and y axes respectively, and these are less apparent in the form of the Lagrangian.

2.4.2 Energy Conservation

We may ask what happens to the Lagrangian along the path of the motion.

$$\frac{dL}{dt} = \sum_i \frac{\partial L}{\partial q_i}\frac{dq_i}{dt} + \sum_i \frac{\partial L}{\partial \dot{q}_i}\frac{d\dot{q}_i}{dt} + \frac{\partial L}{\partial t}$$

In the first term the first factor is

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}$$

[6] It seems curious that we are finding straightforwardly one of the components of the conserved momentum, but not the other two, L_y and L_x, which are also conserved. The fact that not all of these emerge as conjugates to ignorable coordinates is related to the fact that the components of the angular momentum do not commute in quantum mechanics. This will be discussed further in section (6.6.1).

by the equations of motion, so

$$\frac{dL}{dt} = \frac{d}{dt}\left(\sum_i \frac{\partial L}{\partial \dot q_i}\dot q_i\right) + \frac{\partial L}{\partial t}.$$

We expect energy conservation when the potential is time invariant and there is not time dependence in the constraints, *i.e.* when $\partial L/\partial t = 0$, so we rewrite this in terms of

$$H(q,\dot q,t) = \sum_i \dot q_i \frac{\partial L}{\partial \dot q_i} - L = \sum_i \dot q_i P_i - L$$

Then for the actual motion of the system,

$$\frac{dH}{dt} = -\frac{\partial L}{\partial t}.$$

If $\partial L/\partial t = 0$, H is conserved.

H is essentially the Hamiltonian, although strictly speaking that name is reserved for the function $H(q,p,t)$ on extended phase space rather than the function with arguments $(q,\dot q,t)$. What is H physically? In the case of Newtonian mechanics with a potential function, L is an inhomogeneous quadratic function of the velocities $\dot q_i$. If we write the Lagrangian $L = L_2 + L_1 + L_0$ as a sum of pieces purely quadratic, purely linear, and independent of the velocities respectively, then

$$\sum_i \dot q_i \frac{\partial}{\partial \dot q_i}$$

is an operator which multiplies each term by its order in velocities,

$$\sum_i \dot q_i \frac{\partial L_n}{\partial \dot q_i} = nL_n, \qquad \sum_i \dot q_i \frac{\partial L}{\partial \dot q_i} = 2L_2 + L_1,$$

and

$$H = L_2 - L_0.$$

For a system of particles described by their cartesian coordinates, L_2 is just the kinetic energy T, while L_0 is the negative of the potential energy $L_0 = -U$, so $H = T + U$ is the ordinary energy. There are, however, constrained systems, such as the bead on a spoke of Section 2.2.1, for which the Hamiltonian is conserved but is not the ordinary energy.

2.5 Hamilton's Equations

We have written the Lagrangian as a function of q_i, $\dot q_i$, and t, so it is a function of $N + N + 1$ variables. For a free particle we can write the kinetic energy either as $\frac{1}{2}m\dot x^2$ or as $p^2/2m$. More generally, we can[7] reexpress the dynamics in terms of the $2N + 1$ variables q_k, P_k, and t.

The motion of the system sweeps out a path in the space $(q,\dot q,t)$ or a path in (q,P,t). Along this line, the variation of L is

$$\begin{aligned} dL &= \sum_k \left(\frac{\partial L}{\partial \dot q_k} d\dot q_k + \frac{\partial L}{\partial q_k} dq_k\right) + \frac{\partial L}{\partial t} dt \\ &= \sum_k \left(P_k d\dot q_k + \dot P_k dq_k\right) + \frac{\partial L}{\partial t} dt \end{aligned}$$

where for the first term we used the definition of the generalized momentum and in the second we have used the equations of motion $\dot P_k = \partial L/\partial q_k$. Then examining the change in the Hamiltonian $H = \sum_k P_k \dot q_k - L$ along this actual motion,

$$\begin{aligned} dH &= \sum_k (P_k d\dot q_k + \dot q_k dP_k) - dL \\ &= \sum_k \left(\dot q_k dP_k - \dot P_k dq_k\right) - \frac{\partial L}{\partial t} dt. \end{aligned}$$

If we think of $\dot q_k$ and H as functions of q and P, and think of H as a function of q, P, and t, we see that the physical motion obeys

$$\dot q_k = \left.\frac{\partial H}{\partial P_k}\right|_{q,t}, \qquad \dot P_k = -\left.\frac{\partial H}{\partial q_k}\right|_{P,t}, \qquad \left.\frac{\partial H}{\partial t}\right|_{q,P} = -\left.\frac{\partial L}{\partial t}\right|_{q,\dot q}$$

The first two constitute **Hamilton's equations of motion**, which are first order equations for the motion of the point representing the system in phase space.

Let's work out a simple example, the one dimensional harmonic oscillator. Here the kinetic energy is $T = \frac{1}{2}m\dot x^2$, the potential energy is $U = \frac{1}{2}kx^2$, so

[7] In field theory there arise situations in which the set of functions $P_k(q_i,\dot q_i)$ cannot be inverted to give functions $\dot q_i = \dot q_i(q_j, P_j)$. This gives rise to local gauge invariance, and will be discussed in Chapter 8, but until then we will assume that the phase space (q,p), or cotangent bundle, is equivalent to the tangent bundle, *i.e.* the space of $(q,\dot q)$.

$L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2$, the only generalized momentum is $P = \partial L/\partial\dot{x} = m\dot{x}$, and the Hamiltonian is $H = P\dot{x} - L = P^2/m - (P^2/2m - \frac{1}{2}kx^2) = P^2/2m + \frac{1}{2}kx^2$. Note this is just the *sum* of the kinetic and potential energies, or the total energy.

Hamilton's equations give

$$\dot{x} = \left.\frac{\partial H}{\partial P}\right|_x = \frac{P}{m}, \qquad \dot{P} = -\left.\frac{\partial H}{\partial x}\right|_P = -kx = F.$$

These two equations verify the usual connection of the momentum and velocity and give Newton's second law.

The identification of H with the total energy is more general than our particular example. If T is purely quadratic in velocities, we can write $T = \frac{1}{2}\sum_{ij} M_{ij}\dot{q}_i\dot{q}_j$ in terms of a symmetric **mass matrix** M_{ij}. If in addition U is independent of velocities,

$$\begin{aligned} L &= \frac{1}{2}\sum_{ij} M_{ij}\dot{q}_i\dot{q}_j - U(q) \\ P_k &= \frac{\partial L}{\partial \dot{q}_k} = \sum_i M_{ki}\dot{q}_i \end{aligned}$$

which as a matrix equation in a n-dimensional space is $P = M\cdot\dot{q}$. Assuming M is invertible,[8] we also have $\dot{q} = M^{-1}\cdot P$, so

$$\begin{aligned} H &= P^T\cdot\dot{q} - L \\ &= P^T\cdot M^{-1}\cdot P - \left(\frac{1}{2}\dot{q}^T\cdot M\cdot\dot{q} - U(q)\right) \\ &= P^T\cdot M^{-1}\cdot P - \frac{1}{2}P^T\cdot M^{-1}\cdot M\cdot M^{-1}\cdot P + U(q) \\ &= \frac{1}{2}P^T\cdot M^{-1}\cdot P + U(q) = T + U \end{aligned}$$

so we see that the Hamiltonian is indeed the total energy under these circumstances.

[8] If M were not invertible, there would be a linear combination of velocities which does not affect the Lagrangian. The degree of freedom corresponding to this combination would have a Lagrange equation without time derivatives, so it would be a constraint equation rather than an equation of motion. But we are assuming that the q's are a set of independent generalized coordinates that have already been pruned of all constraints.

2.6 Don't plug Equations of Motion into the Lagrangian!

When we have a Lagrangian with an ignorable coordinate, say θ, and therefore a conjugate momentum P_θ which is conserved and can be considered a constant, we are able to reduce the problem to one involving one fewer degrees of freedom. That is, one can substitute into the other differential equations the value of $\dot{\theta}$ in terms of P_θ and other degrees of freedom, so that θ and its derivatives no longer appear in the equations of motion. For example, consider the two dimensional isotropic harmonic oscillator,

$$\begin{aligned} L &= \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) - \frac{1}{2}k\left(x^2 + y^2\right) \\ &= \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - \frac{1}{2}kr^2 \end{aligned}$$

in polar coordinates. The equations of motion are

$$\begin{aligned} \dot{P}_\theta &= 0, \quad \text{where } P_\theta = mr^2\dot{\theta}, \\ m\ddot{r} &= -kr + mr\dot{\theta}^2 \Longrightarrow m\ddot{r} = -kr + P_\theta^2\Big/ mr^3. \end{aligned}$$

The last equation is now a problem in the one degree of freedom r.

One might be tempted to substitute for $\dot{\theta}$ into the Lagrangian and then have a Lagrangian involving one fewer degrees of freedom. In our example, we would get

$$L = \frac{1}{2}m\dot{r}^2 + \frac{P_\theta^2}{2mr^2} - \frac{1}{2}kr^2,$$

which gives the equation of motion

$$m\ddot{r} = -\frac{P_\theta^2}{mr^3} - kr.$$

This is wrong

Notice that the last equation has the sign of the P_θ^2 term reversed from the correct equation. Why did we get the wrong answer? In deriving the Lagrange equation which comes from varying r, we need

$$\frac{d}{dt}\left.\frac{\partial L}{\partial \dot{r}}\right|_{r,\theta,\dot{\theta}} = \left.\frac{\partial L}{\partial r}\right|_{\dot{r},\theta,\dot{\theta}}.$$

But we treated P_θ as fixed, which means that when we vary r on the right hand side, we are not holding $\dot\theta$ fixed, as we should be. While we often write partial derivatives without specifying explicitly what is being held fixed, they are not defined without such a specification, which we are expected to understand implicitly. However, there are several examples in Physics, such as thermodynamics, where this implicit understanding can be unclear, and the results may not be what was intended.

2.7 Velocity-dependent forces

We have concentrated thus far on Newtonian mechanics with a potential given as a function of coordinates only. As the potential is a piece of the Lagrangian, which may depend on velocities as well, we should also entertain the possibility of velocity-dependent potentials. Only by considering such a potential can we possibly find velocity-dependent forces, and one of the most important force laws in physics is of that form. This is the Lorentz force[9] on a particle of charge q in the presence of electromagnetic fields $\vec{E}(\vec{r},t)$ and $\vec{B}(\vec{r},t)$,

$$\vec{F} = q\left(\vec{E} + \frac{\vec{v}}{c}\times\vec{B}\right). \tag{2.18}$$

If the motion of a charged particle is described by Lagrangian mechanics with a potential $U(\vec{r},\vec{v},t)$, Lagrange's equation says

$$0 = \frac{d}{dt}\frac{\partial L}{\partial v_i} - \frac{\partial L}{\partial r_i} = m\ddot{r}_i - \frac{d}{dt}\frac{\partial U}{\partial v_i} + \frac{\partial U}{\partial r_i}, \qquad \text{so } F_i = \frac{d}{dt}\frac{\partial U}{\partial v_i} - \frac{\partial U}{\partial r_i}.$$

We want a force linear in $\vec{v}$ and proportional to q, so let us try

$$U = q\left(\phi(\vec{r},t) + \vec{v}\cdot\vec{C}(\vec{r},t)\right).$$

Then we need to have

$$\vec{E} + \frac{\vec{v}}{c}\times\vec{B} = \frac{d}{dt}\vec{C} - \vec{\nabla}\phi - \sum_j v_j\vec{\nabla}C_j. \tag{2.19}$$

[9] We have used Gaussian units here, but those who prefer S. I. units (rationalized MKS) can simply set $c = 1$.

The first term is a stream derivative evaluated at the time-dependent position of the particle, so, as in Eq. (2.5),

$$\frac{d}{dt}\vec{C} = \frac{\partial\vec{C}}{\partial t} + \sum_j v_j\frac{\partial\vec{C}}{\partial x_j}.$$

The last term looks like the last term of (2.19), except that the indices on the derivative operator and on $\vec{C}$ have been reversed. This suggests that these two terms combine to form a cross product. Indeed, noting (A.17) that

$$\vec{v}\times\left(\vec{\nabla}\times\vec{C}\right) = \sum_j v_j\vec{\nabla}C_j - \sum v_j\frac{\partial\vec{C}}{\partial x_j},$$

we see that (2.19) becomes

$$\vec{E} + \frac{\vec{v}}{c}\times\vec{B} = \frac{\partial\vec{C}}{\partial t} - \vec{\nabla}\phi - \sum_j v_j\vec{\nabla}C_j + \sum_j v_j\frac{\partial\vec{C}}{\partial x_j} = \frac{\partial\vec{C}}{\partial t} - \vec{\nabla}\phi - \vec{v}\times\left(\vec{\nabla}\times\vec{C}\right).$$

We have successfully generated the term linear in $\vec{v}$ if we can show that there exists a vector field $\vec{C}(\vec{r},t)$ such that $\vec{B} = -c\vec{\nabla}\times\vec{C}$. A curl is always divergenceless, so this requires $\vec{\nabla}\cdot\vec{B} = 0$, but this is indeed one of Maxwell's equations, and it ensures[10] there exists a vector field $\vec{A}$, known as the **magnetic vector potential**, such that $\vec{B} = \vec{\nabla}\times\vec{A}$. Thus with $\vec{C} = -\vec{A}/c$, we need only to find a ϕ such that

$$\vec{E} = -\vec{\nabla}\phi - \frac{1}{c}\frac{\partial\vec{A}}{\partial t}.$$

Once again, one of Maxwell's laws,

$$\vec{\nabla}\times\vec{E} + \frac{1}{c}\frac{\partial\vec{B}}{\partial t} = 0,$$

guarantees the existence of ϕ, the **electrostatic potential**, because after inserting $\vec{B} = \vec{\nabla}\times\vec{A}$, this is a statement that $\vec{E} + (1/c)\partial\vec{A}/\partial t$ has no curl, and is the gradient of something.

[10] This is but one of many consequences of the Poincaré lemma, discussed in section 6.5 (well, it should be). The particular forms we are using here state that if $\vec{\nabla}\cdot\vec{B} = 0$ and $\vec{\nabla}\times\vec{F} = 0$ in all of $\mathbb{R}^3$, then there exist a scalar function ϕ and a vector field $\vec{A}$ such that $\vec{B} = \vec{\nabla}\times\vec{A}$ and $\vec{F} = \vec{\nabla}\phi$.

Thus we see that the Lagrangian which describes the motion of a charged particle in an electromagnetic field is given by a velocity-dependent potential

$$U(\vec{r},\vec{v}) = q\left(\phi(r,t) - (\vec{v}/c)\cdot\vec{A}(\vec{r},t)\right).$$

Note, however, that this Lagrangian describes only the motion of the charged particle, and not the dynamics of the field itself.

Arbitrariness in the Lagrangian In this discussion of finding the Lagrangian to describe the Lorentz force, we used the lemma that guaranteed that the divergenceless magnetic field $\vec{B}$ can be written in terms of some magnetic vector potential $\vec{A}$, with $\vec{B} = \vec{\nabla}\times\vec{A}$. But $\vec{A}$ is not uniquely specified by $\vec{B}$; in fact, if a change is made, $\vec{A}\to\vec{A}+\vec{\nabla}\lambda(\vec{r},t)$, $\vec{B}$ is unchanged because the curl of a gradient vanishes. The electric field $\vec{E}$ will be changed by $-(1/c)\partial\vec{A}/\partial t$, however, unless we also make a change in the electrostatic potential, $\phi\to\phi-(1/c)\partial\lambda/\partial t$. If we do, we have completely unchanged electromagnetic fields, which is where the physics lies. This change in the potentials,

$$\vec{A}\to\vec{A}+\vec{\nabla}\lambda(\vec{r},t), \qquad \phi\to\phi-(1/c)\partial\lambda/\partial t, \tag{2.20}$$

is known as a **gauge transformation**, and the invariance of the physics under this change is known as **gauge invariance**. Under this change, the potential U and the Lagrangian are not unchanged,

$$L\to L - q\left(\delta\phi - \frac{\vec{v}}{c}\cdot\delta\vec{A}\right) = L + \frac{q}{c}\frac{\partial\lambda}{\partial t} + \frac{q}{c}\vec{v}\cdot\vec{\nabla}\lambda(\vec{r},t) = L + \frac{q}{c}\frac{d\lambda}{dt}.$$

We have here an example which points out that there is not a unique Lagrangian which describes a given physical problem, and the ambiguity is more that just the arbitrary constant we always knew was involved in the potential energy. This ambiguity is quite general, not depending on the gauge transformations of Maxwell fields. In general, if

$$L^{(2)}(q_j,\dot{q}_j,t) = L^{(1)}(q_j,\dot{q}_j,t) + \frac{d}{dt}f(q_j,t) \tag{2.21}$$

then $L^{(1)}$ and $L^{(2)}$ give the same equations of motion, and therefore the same physics, for $q_j(t)$. While this can be easily checked by evaluating the Lagrange equations, it is best understood in terms of the variation of the action. For

any path $q_j(t)$ between q_{jI} at $t=t_I$ to q_{jF} at $t=t_F$, the two actions are related by

$$\begin{aligned} S^{(2)} &= \int_{t_I}^{t_F}\left(L^{(1)}(q_j,\dot{q}_j,t) + \frac{d}{dt}f(q_j,t)\right)dt \\ &= S^{(1)} + f(q_{jF},t_F) - f(q_{jI},t_I). \end{aligned}$$

The variation of path that one makes to find the stationary action does not change the endpoints q_{jF} and q_{jI}, so the difference $S^{(2)}-S^{(1)}$ is a constant independent of the trajectory, and a stationary trajectory for $S^{(2)}$ is clearly stationary for $S^{(1)}$ as well.

The conjugate momenta are affected by the change in Lagrangian, however, because $L^{(2)} = L^{(1)} + \sum_j \dot{q}_j\partial f/\partial q_j + \partial f/\partial t$, so

$$p_j^{(2)} = \frac{\partial L^{(2)}}{\partial\dot{q}_j} = p_j^{(1)} + \frac{\partial f}{\partial q_j}.$$

This ambiguity is not usually mentioned in elementary mechanics, because if we restict our attention to Lagrangians consisting of canonical kinetic energy and potentials which are velocity-independent, a change (2.21) to a Lagrangian $L^{(1)}$ of this type will produce an $L^{(2)}$ which is not of this type, unless f is independent of position q and leaves the momenta unchanged. That is, the only f which leaves U velocity independent is an arbitrary constant.

Dissipation Another familiar force which is velocity dependent is friction. Even the "constant" sliding friction met with in elementary courses depends on the direction, if not the magnitude, of the velocity. Friction in a viscous medium is often taken to be a force proportional to the velocity, $\vec{F} = -\alpha\vec{v}$. We saw above that a potential linear in velocities produces a force perpendicular to $\vec{v}$, and a term higher order in velocities will contribute a force that depends on acceleration. This situation cannot handled by Lagrange's equations. More generally, a Lagrangian can produce a force $Q_i = R_{ij}\dot{q}_j$ with antisymmetric R_{ij}, but not for a symmetric matrix. An extension to the Lagrange formalism, involving Rayleigh's dissipation function, can handle such a case. These dissipative forces are discussed in Ref. [6].

Exercises

2.1 (Galelean relativity): Sally is sitting in a railroad car observing a system of particles, using a Cartesian coordinate system so that the particles are at positions $\vec{r}_i^{(S)}(t)$, and move under the influence of a potential $U^{(S)}(\{\vec{r}_i^{(S)}\})$. Thomas is in another railroad car, moving with constant velocity $\vec{u}$ with respect to Sally, and so he describes the position of each particle as $\vec{r}_i^{(T)}(t) = \vec{r}_i^{(S)}(t) - \vec{u}t$. Each takes the kinetic energy to be of the standard form in his system, *i.e.* $T^{(S)} = \frac{1}{2}\sum m_i \left(\dot{\vec{r}}_i^{(S)}\right)^2$ and $T^{(T)} = \frac{1}{2}\sum m_i \left(\dot{\vec{r}}_i^{(T)}\right)^2$.

(a) Show that if Thomas assumes the potential function $U^{(T)}(\vec{r}^{(T)})$ to be the same as Sally's at the same physical points,

$$U^{(T)}(\vec{r}^{(T)}) = U^{(S)}(\vec{r}^{(T)} + \vec{u}t), \tag{2.22}$$

then the equations of motion derived by Sally and Thomas describe the same physics. That is, if $r_i^{(S)}(t)$ is a solution of Sally's equations, $r_i^{(T)}(t) = r_i^{(S)}(t) - \vec{u}t$ is a solution of Thomas'.

(b) show that if $U^{(S)}(\{\vec{r}_i\})$ is a function only of the displacements of one particle from another, $\{\vec{r}_i - \vec{r}_j\}$, then $U^{(T)}$ is the same function of its arguments as $U^{(S)}$, $U^{(T)}(\{\vec{r}_i\}) = U^{(S)}(\{\vec{r}_i\})$. This is a different statement than Eq. 2.22, which states that they agree at the same physical configuration. Show it will not generally be true if $U^{(S)}$ is not restricted to depend only on the differences in positions.

(c) If it is true that $U^{(S)}(\vec{r}) = U^{(T)}(\vec{r})$, show that Sally and Thomas derive the same equations of motion, which we call "form invariance" of the equations.

(d) Show that nonetheless Sally and Thomas disagree on the energy of a particular physical motion, and relate the difference to the total momentum. Which of these quantities are conserved?

2.2 In order to show that the shortest path in two dimensional Euclidean space is a straight line without making the assumption that Δx does not change sign along the path, we can consider using a parameter λ and describing the path by two functions $x(\lambda)$ and $y(\lambda)$, say with $\lambda \in [0,1]$. Then

$$\ell = \int_0^1 d\lambda \sqrt{\dot{x}^2(\lambda) + \dot{y}^2(\lambda)},$$

where $\dot{x}$ means $dx/d\lambda$. This is of the form of a variational integral with two variables. Show that the variational equations do *not* determine the functions $x(\lambda)$ and $y(\lambda)$, but do determine that the path is a straight line. Show that the pair of functions $(x(\lambda), y(\lambda))$ gives the same action as another pair $(\tilde{x}(\lambda), \tilde{y}(\lambda))$, where $\tilde{x}(\lambda) = x(t(\lambda))$ and $\tilde{y}(\lambda) = y(t(\lambda))$, where $t(\lambda)$ is any monotone function mapping $[0,1]$ onto itself. Explain why this equality of the lengths is obvious

in terms of alternate parameterizations of the path. [In field theory, this is an example of a local gauge invariance, and plays a major role in string theory.]

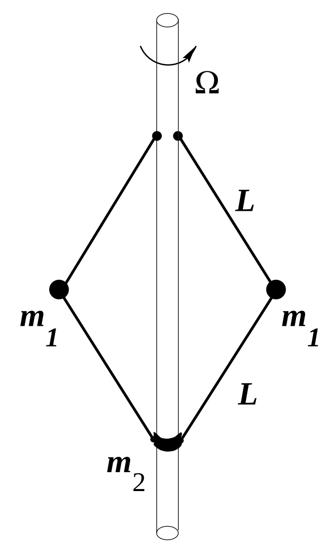
2.3 Consider a circular hoop of radius R rotating about a vertical diameter at a fixed angular velocity Ω. On the hoop there is a bead of mass m, which slides without friction on the hoop. The only external force is gravity. Derive the Lagrangian and the Lagrange equation using the polar angle θ as the unconstrained generalized coordinate. Find a conserved quantity, and find the equilibrium points, for which $\dot{\theta} = 0$. Find the condition on Ω such that there is an equilibrium point away from the axis.

2.4 Early steam engines had a feedback device, called a governor, to automatically control the speed. The engine rotated a vertical shaft with an angular velocity Ω proportional to its speed. On opposite sides of this shaft, two hinged rods each held a metal weight, which was attached to another such rod hinged to a sliding collar, as shown.

As the shaft rotates faster, the balls move outwards, the collar rises and uncovers a hole, releasing some steam. Assume all hinges are frictionless, the rods massless, and each ball has mass m_1 and the collar has mass m_2.

(a) Write the Lagrangian in terms of the generalized coordinate θ.

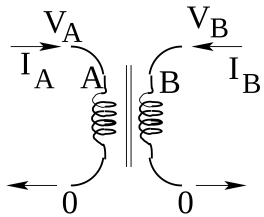
(b) Find the equilibrium angle θ as a function of the shaft angular velocity Ω. Tell whether the equilibrium is stable or not.



Governor for a steam engine.

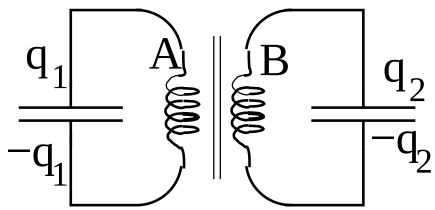
2.5 A transformer consists of two coils of conductor each of which has an inductance, but which also have a coupling, or mutual inductance.

If the current flowing into the upper posts of coils A and B are $I_A(t)$ and $I_B(t)$ respectively, the voltage difference or EMF across each coil is V_A and V_B respectively, where



$$\begin{aligned} V_A &= L_A \frac{dI_A}{dt} + M \frac{dI_B}{dt} \\ V_B &= L_B \frac{dI_B}{dt} + M \frac{dI_A}{dt} \end{aligned}$$

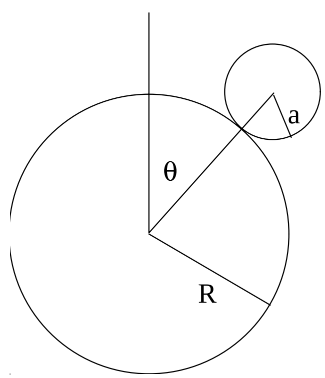
Consider the circuit shown, two capacitors coupled by a such a transformer, where the capacitances are C_A and C_B respectively, with the charges $q_1(t)$ and $q_2(t)$ serving as the generalized coordinates for this problem. Write down the two second order differential equations of "motion" for $q_1(t)$ and $q_2(t)$, and write a Lagrangian for this system.



2.6 A cylinder of radius R is held horizontally in a fixed position, and a smaller uniform cylindrical disk of radius a is placed on top of the first cylinder, and is released from rest. There is a coefficient of static friction μ_s and a coefficient of kinetic friction $\mu_k < \mu_s$ for the contact between the cylinders. As the equilibrium at the top is unstable, the top cylinder will begin to roll on the bottom cylinder.

(a) If μ_s is sufficiently large, the small disk will roll until it separates from the fixed cylinder. Find the angle θ at which the separation occurs, and find the minimum value of μ_s for which this situation holds.

(b) If μ_s is less than the minimum value found above, what happens differently, and at what angle θ does this different behavior begin?



A small cylinder rolling on a fixed larger cylinder.

2.7 (a) Show that if $\Phi(q_1, ..., q_n, t)$ is an arbitrary differentiable function on extended configuration space, and $L^{(1)}(\{q_i\}, \{\dot{q}_j\}, t)$ and $L^{(2)}(\{q_i\}, \{\dot{q}_j\}, t)$ are two

Lagrangians which differ by the total time derivative of Φ,

$$L^{(1)}(\{q_i\}, \{\dot{q}_j\}, t) = L^{(2)}(\{q_i\}, \{\dot{q}_j\}, t) + \frac{d}{dt}\Phi(q_1, ..., q_n, t),$$

show by explicit calculations that the equations of motion determined by $L^{(1)}$ are the same as the equations of motion determined by $L^{(2)}$.
(b) What is the relationship between the momenta $p_i^{(1)}$ and $p_i^{(2)}$ determined by these two Lagrangians respectively.

2.8 A particle of mass m_1 moves in two dimensions on a frictionless horizontal table with a tiny hole in it. An inextensible massless string attached to m_1 goes through the hole and is connected to another particle of mass m_2, which moves vertically only. Give a full set of generalized unconstrained coordinates and write the Lagrangian in terms of these. Assume the string remains taut at all times and that the motions in question never have either particle reaching the hole, and there is no friction of the string sliding at the hole.
Are there ignorable coordinates? Reduce the problem to a single second order differential equation. Show this is equivalent to single particle motion in one dimension with a potential $V(r)$, and find $V(r)$.

2.9 Consider a mass m on the end of a massless rigid rod of length ℓ, the other end of which is free to rotate about a fixed point. This is a spherical pendulum. Find the Lagrangian and the equations of motion.

2.10 (a) Find a differential equation for $\theta(\phi)$ for the shortest path on the surface of a sphere between two arbitrary points on that surface, by minimizing the length of the path, assuming it to be monotone in ϕ.
(b) By geometrical argument (that it must be a great circle) argue that the path should satisfy

$$\cos(\phi - \phi_0) = K \cot\theta,$$

and show that this is indeed the solution of the differential equation you derived.

2.11 Consider some intelligent bugs who live on a turntable which, according to inertial observers, is spinning at angular velocity ω about its center. At any one time, the inertial observer can describe the points on the turntable with polar coordinates r, ϕ. If the bugs measure distances between two objects at rest with respect to them, at infinitesimally close points, they will find