PHYS 111: Theoretical Mechanics

Connor Neely, Fall 2024

* Adapted from FA24 lectures.

1 Lagrangian Mechanics

1.1 The Principle of Least Action

Newtonian mechanics provides a nice, intuitive interpretation of how forces cause motion: a force pushes on an object at a given place and at a given time, nudging it along its trajectory. The issue with this "local" approach is that, if we only have information about some snapshots of time, some more general properties or symmetries of the system can get obscured. It's also quite unpleasant to write Newton's laws in different coordinate systems, even relatively simple ones like polar coordinates!

For our new formulation of mechanics, let's start with 1D motion under the influence of conservative forces. Suppose we toss a ball into the air, and we want to find a function $x(t)$ on $t \in [t_i, t_f]$ that represents its height over time. In our new approach, every path the ball could take has an associated number S called the action:

$$S \equiv \int_{t_i}^{t_f} L(x(t), \dot{x}(t), t)\, dt,$$

where L is called the Lagrangian, which is in turn defined by

$$L \equiv T - U.$$

Here, T and U are the kinetic and potential energies, respectively. ($S[x(t)]$ is called a functional. as it takes a path as an input and spits out a number.) The principle of least action states that the path the ball actually takes is the one with the smallest S—to actually apply this, we'll need the calculus of variations.

Minimizing the Action

Recall that the Taylor expansion about x_0 of a single-variable function is

$$f(x) = f(x_0 + \Delta x) = f(x_0) + \frac{df}{dx}(x_0)\Delta x + \frac{1}{2}\frac{d^2 f}{dx^2}(x_0)\Delta x^2 + \cdots .$$

If x_0 has a minimum for f then to first order in Δx we have $f(x_0) \approx f(x_0 + \Delta x)$. This gives an alternative interpretation for minima which will prove useful when we go to minimize our action functional. Note that this we can generalize this to multivariable functions—for two variables we have, to first order,

$$\frac{\partial f}{\partial x}(x_0, y_0)\Delta x + \frac{\partial f}{\partial y}(x_0, y_0)\Delta y = 0$$

at a minimum. Since this equation holds for any small Δx and Δy, the only solution is for both partial derivatives to be zero, like we'd expect.

Now let's apply all this in our action minimization problem. Suppose there is a path $x_0(t)$ that minimizes S, meaning the action is unchanged under tiny variations and $S[x_0(t) + \Delta x(t)] = S[x_0(t)]$. We could show, using a first-order Taylor expansion, that

$$S[x_0(t) + \Delta x(t)] - S[x_0(t)] = \int_{t_1}^{t_2} dt \left(\frac{\partial L(x, \dot{x}, t)}{\partial x}\Delta x + \frac{\partial L(x, \dot{x}, t)}{\partial \dot{x}}\Delta \dot{x} \right).$$

So to minimize the action we solve the equation

$$\int_{t_1}^{t_2} \left(\frac{\partial L}{\partial x}\Delta x + \frac{\partial L}{\partial \dot{x}}\frac{d}{dt}(\Delta x) \right) dt = 0.$$

We'd like to make this look a little simpler, though. Ignoring the first term for now, we'll do integration by parts on the second term:

$$\int_{t_i}^{t_f} \frac{\partial L}{\partial \dot{x}}\frac{d}{dt}(\Delta x)\, dt = \frac{\partial L}{\partial \dot{x}}\Delta x \Bigg|_{t_i}^{t_f} - \int_{t_i}^{t_f} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) \Delta x\, dt.$$

In order for this variational approach to make sense, we should keep $\Delta x(t_i) = \Delta x(t_f) = 0$ so that our "varied" function still starts and ends at the same point. So the first term disappears, and our equation becomes

$$\int_{t_1}^{t_2} \left[\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right)\right] \Delta x \, dt = 0.$$

Because this must hold for any choice of $\Delta x(t)$, the bracketed factor must be zero. This gives us the Euler-Lagrange equation

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = \frac{\partial L}{\partial x}.$$

This variational principle can be applied to a variety of scenarios—the general method is to write down the "action" integral we'd like to minimize, extract a "Lagrangian", and write down the corresponding Euler-Lagrange equation.

Generalized Coordinates

Now, perhaps most important about what we've done here is that it is completely equivalent to Newton's framework! For a particle moving along the x-axis we have the Lagrangian $L = m\dot{x}^2/2 - U(x)$ and the Euler-Lagrange equation

$$m\ddot{x} = -\frac{dU}{dx},$$

which we recognize this as Newton's second law. But in a sense the Lagrangian formalism is more powerful than the Newtonian one because for any choice of coordinate q we can write

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

but we cannot in general write $m\ddot{q} = -\partial U/\partial q$. (A modification of this equation does work, of course, but things get messy when q is non-rectangular.) We call q a generalized coordinate because it really can be anything—linear, angular, or something more creative and exotic. In practice we take whatever best respects the constraints and symmetries of the problem at hand.

Even better, Lagrangian mechanics allows us to more easily work in accelerating reference frames! Because the Euler-Lagrange equation is valid for any generalized coordinate, we can simply calculate the Lagrangian in an inertial reference frame (using an accelerating coordinate) and the Euler-Lagrange equation will spit out the equation of motion we're looking for.

Lastly, as expected, everything we've just done also applies to systems that are described using several generalized coordinates. In two dimensions the action is

$$S = \int_{t_i}^{t_f} L(q_1, \dot{q}_1, q_2, \dot{q}_2, t)\, dt$$

which, when minimized, yields the Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_1} = \frac{\partial L}{\partial q_1}, \qquad \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_2} = \frac{\partial L}{\partial q_2}.$$

This generalizes to systems with several coordinates. (We define a system's number of degrees of freedom to be the minimum number of generalized coordinates needed to characterize its motion.)

1.2 Noether's Theorem and the Hamiltonian

We've already shown that the Euler-Lagrange equations are equivalent to Newton's second law $dp/dt = F$, but we might also notice that the two sets of equations take very similar forms. So we'll often refer to the quantities

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \quad \text{and} \quad F_i = \frac{\partial L}{\partial q_i}$$

as the generalized momentum and generalized force, respectively, in the q_i direction.

Now, a coordinate q_i is called cyclic if $\partial L/\partial q_i = 0$ or, equivalently, if q_i does not appear anywhere in L. The Euler-Lagrange equations reveal that cyclic coordinates have $dp_i/dt = 0$, meaning their component of the generalized momentum is conserved!

We might formalize this a little more in terms of symmetries, coordinate transformations which leave the action unchanged—for example, if q_i does not appear in the Lagrangian then we can translate q_i in any way we'd like (say, by adding a Δq_i) without changing the action. In particular, such a transformation encodes a continuous symmetry because the "step size" Δq_i can vary continuously while still being a symmetry. This particular kind of symmetry is what leads to conservation of p_i.

These observations bring us to Noether's theorem: every continuous symmetry of the action has a corresponding conserved quantity. These symmetries tend to have striking physical interpretations—translational symmetry, for example, just means that the evolution of a system is completely independent of where it is in space, or perhaps how it's oriented. This fact alone is enough to conclude that the system's linear momentum or angular momentum is conserved.

Translational symmetry is easy enough to intuit when spatial coordinates are being translated. When it's time being translated, we're really just looking at whether a system's evolution is dependent on when we set $t = 0$. Still simple in concept, but it isn't at all obvious what the associated conserved quantity should be.

Consider a Lagrangian $L(x, \dot{x}, t)$ for which $\partial L/\partial t = 0$. Certainly L is not conserved, but maybe computing dL/dt will be illuminating anyway:

$$\begin{aligned}
\frac{dL}{dt} &= \frac{\partial L}{\partial x}\frac{dx}{dt} + \frac{\partial L}{\partial \dot{x}}\frac{d\dot{x}}{dt} + \frac{\partial L}{\partial t} \\
&= \frac{\partial L}{\partial x}\dot{x} + \frac{\partial L}{\partial \dot{x}}\frac{d}{dt}(\dot{x}) \\
&= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right)\dot{x} + \frac{\partial L}{\partial \dot{x}}\frac{d}{dt}(\dot{x}) \\
&= \frac{d}{dt}\left(\dot{x}\frac{\partial L}{\partial \dot{x}}\right).
\end{aligned}$$

Noticing derivatives on either side of the equation, we can force a conserved quantity by simply rearranging:

$$0 = \frac{d}{dt}\left(\dot{x}\frac{\partial L}{\partial \dot{x}} - L\right).$$

We call the conserved parenthetical the Hamiltonian,

$$H = \dot{x}\frac{\partial L}{\partial \dot{x}} - L.$$

In several dimensions this is

$$H = \left(\sum_i p_i \dot{q}_i\right) - L = \left(\sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i}\right) - L.$$

This is a pretty abstract quantity, but it might not be too hard to guess that it's closely connected to the total mechanical energy in a system. In particular, we have $H = T + U$ whenever the transformation from Cartesian coordinates to whatever generalized coordinates we're using is time-independent. We'll spend the rest of the section proving this fact in three dimensions.

Let r_i and q_i denote Cartesian and generalized coordinates, respectively, for $i = 1, 2, 3$, and let the transformation between these coordinates be encoded by $r_i = f_i(q_1, q_2, q_3, t)$. If these r_i are time-independent then the kinetic energy is given by

$$T = \frac{1}{2}m\sum_i \dot{r}_i^2 = \frac{1}{2}m\sum_i\left(\sum_j \frac{\partial f_i}{\partial q_j}\dot{q}_j\right)^2 = \frac{1}{2}m\sum_{i,j,k}\frac{\partial f_i}{\partial q_j}\dot{q}_j\frac{\partial f_i}{\partial q_k}\dot{q}_k.$$

(Note that all sum indices are implied to range from 1 to 3 here.) We aim to show that the first term in the Hamiltonian is equal to $2T$. To this end we note that, for a velocity-independent potential, $\partial L/\partial \dot{q}_m$ is equivalent to $\partial T/\partial \dot{q}_m$. Also note that the f_i are velocity-independent, meaning

$$\begin{aligned}\frac{\partial T}{\partial \dot{q}_m} &= \frac{1}{2}m\sum_{i,j,k}\frac{\partial f_i}{\partial q_j}\frac{\partial f_i}{\partial q_k}\frac{\partial}{\partial \dot{q}_m}(\dot{q}_j\dot{q}_k) \\ &= \frac{1}{2}m\sum_{i,j,k}\frac{\partial f_i}{\partial q_j}\frac{\partial f_i}{\partial q_k}\left(\dot{q}_j\delta_{km} + \dot{q}_k\delta_{jm}\right),\end{aligned}$$

where δ_{ab} is a Kronecker delta. Using this to collapse some sums gives

$$\begin{aligned}&= \frac{1}{2}m\sum_{i,j}\frac{\partial f_i}{\partial q_j}\frac{\partial f_i}{\partial q_m}\dot{q}_j + \frac{1}{2}m\sum_{i,k}\frac{\partial f_i}{\partial q_m}\frac{\partial f_i}{\partial q_k}\dot{q}_k \\ &= m\sum_{i,k}\frac{\partial f_i}{\partial q_m}\frac{\partial f_i}{\partial q_k}\dot{q}_k.\end{aligned}$$

So the first bit of the Hamiltonian evaluates to

$$\sum_m \dot{q}_m\frac{\partial L}{\partial \dot{q}_m} = \sum_m \dot{q}_m\frac{\partial T}{\partial \dot{q}_m} = m\sum_{m,i,k}\frac{\partial f_i}{\partial q_m}\frac{\partial f_i}{\partial q_k}\dot{q}_k\dot{q}_m.$$

Comparing with the T we found earlier, this sum is actually $2T$. So $H = 2T - L = T + U$ and the Hamiltonian is the system's total mechanical energy. (If we didn't assume a time-independent transformation then each $\dot{r}_i$ would have an extra $\partial f_i/\partial t$ in it and all of this would fall apart.)

1.3 Lagrange Multipliers

Up to this point we have implemented any constraints on a system's motion via a choice of coordinates that reflects those constraints. But there are times where this isn't enough, either because the constraints are too complex or because we want to use them to learn something else about the system.

Here we'll focus on holonomic constraints, that is, those of the form $f(q_1, \ldots, q_N, t) = 0$. These constraints are implemented via the method of Lagrange multipliers—we take whatever Lagrangian we write down and tack on a $\lambda f(q_1, \ldots, q_N, t)$ at the end. We then compute the Euler-Lagrange equations for the N generalized coordinates *and* for the Lagrange multiplier λ. (This last equation is what encodes the constraint.)

From here we can recover all the results we're familiar with. If this is all we care about then we've potentially wasted a lot of time, but we can do something more! It turns out that the expression $\lambda(\partial f/\partial q_i)$ encodes the ith component of the force associated with the constraint function f. In the case of a simple pendulum, for example, we'd end up with

$$\lambda\frac{\partial f}{\partial r} = -mg\cos\theta - m\ell\dot{\theta}^2,$$

which we recognize as tension and some centripetal force. (Notice that the signs indicate the direction of the force, too, with negatives denoting the $-\hat{r}$ direction.) We call this the generalized constraint force.

2 Two-Body Dynamics

2.1 Two-Body Dynamics

Now that we've developed the Lagrangian formalism in reasonable depth, we'll take a look at what we can actually do with it. Broadly speaking, interactions between a pair of objects fall into two different categories: bound states and scattering states. In a bound state the objects "meet" infinitely times and stay relatively close together, while in a scattering state objects meet once and never again.

Either way, the two-body problem has the Lagrangian

$$L = \frac{1}{2}(m_1 + m_2)|\dot{\mathbf{R}}|^2 + \frac{1}{2}\frac{m_1 m_2}{m_1 + m_2}|\dot{\mathbf{r}}|^2 - U(r),$$

where $\mathbf{R}$ is the center of mass position and $\mathbf{r}$ is the displacement from $\mathbf{r}_1$ to $\mathbf{r}_2$. We can see that the COM momentum is conserved, so we can safely look at the COM frame to get

$$\begin{aligned} L &= \frac{1}{2}\frac{m_1 m_2}{m_1 + m_2}|\dot{\mathbf{r}}|^2 - U(r) \\ &= \frac{1}{2}\mu\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2\right) - U(r). \end{aligned}$$

So the two-body system is equivalent to a single-particle system with reduced mass μ and potential $U(r)$. We can further simplify this Lagrangian by recognizing that all of the interactions between the two bodies are in the same plane, meaning we can restrict our reference frame to $\theta = \pi/2$ and $\dot{\theta} = 0$ with no problems. So now we have

$$L = \frac{1}{2}\mu\left(\dot{r}^2 + r^2\dot{\phi}^2\right) - U(r).$$

Finally, we can note that ϕ is cyclic and define the conserved $p_\phi = \ell = \mu r^2\dot{\phi}$ to write the Lagrangian purely in terms of r and $\dot{r}$. This gives us the kind of spherical symmetry we'd expect from a central potential!

More interesting to us for now, though, is the Hamiltonian (which is also conserved):

$$H = T + U = \frac{1}{2}\mu\dot{r}^2 + \frac{\ell^2}{2\mu r^2} + U(r).$$

Since the latter two terms depend only on r we'll bring them together into one "effective potential" $U_{\text{eff}}(r)$, turning this into

$$H = \frac{1}{2}\mu\dot{r}^2 + U_{\text{eff}}(r).$$

When plotted, $U_{\text{eff}}(r)$ looks like the well-known attraction potential curve—asymptotes at $r = 0$ and $U_{\text{eff}}(r) = 0$ with a little dip below the horizontal axis. Oscillations about the bottom of the "well" with $U_{\text{eff}}(r) < 0$ correspond to elliptical trajectories, while ones with $U_{\text{eff}}(r) > 0$ are hyperbolic. The first term in $U_{\text{eff}}(r)$ serves as a "centrifugal barrier"—if the masses get too close, conservation of angular momentum slingshots them away from one another.

With constant H we can interpret the above equation as a separable differential equation with solution

$$t(r) - t(r_0) = \pm\sqrt{\frac{\mu}{2}}\int_{r_0}^{r}\frac{dr'}{\sqrt{H - U_{\text{eff}}(r')}}.$$

The $\pm$ here accounts for the two possible orientations of the orbit. But $t(r)$ isn't quite what we want—employing a change of variables $dt = (\mu r^2/\ell)d\phi$ gives

$$\phi(r) - \phi(r_0) = \pm\frac{\ell}{\sqrt{2\mu}}\int_{r_0}^{r}\frac{dr'}{r'^2}\frac{1}{\sqrt{H - U_{\text{eff}}(r')}}.$$

This equation will be central to our study of two-body dynamics.

2.2 Kepler Orbits

Now we'll look at the case of $U(r) = -\alpha/r$, which might correspond to a gravitational or electrostatic potential. Let r_0 denote the minimum of $U_{\text{eff}}(r)$, so in this case we'd find that $r_0 = \ell^2/\mu\alpha$. The integral from the previous section turns out to be

$$\phi(r) - \phi(r_0) = \pm \arcsin\left(\frac{\alpha r - \ell^2/\mu}{r\sqrt{\alpha^2 + 2E\ell^2/\mu}}\right).$$

By convention we define $(r_0) = \pi/2$. It follows that

$$\cos\phi = \pm\frac{\alpha r - \ell^2/\mu}{r\sqrt{\alpha^2 + 2E\ell^2/\mu}},$$

and solving for r gives

$$r(\phi) = \frac{\ell^2}{\mu\alpha}\frac{1}{1 \mp \varepsilon\cos\phi}, \qquad \varepsilon \equiv \sqrt{1 + \frac{2E\ell^2}{\mu\alpha^2}},$$

where by convention we take $\mp = +$ so that $r(0) = r_{\min}$. The quantity ε is called the eccentricity, and it denotes the path's deviation from circularity (which is at $\varepsilon = 0$). Bound orbits have $\varepsilon < 1$.

Notice that if we substitute $r = \sqrt{x^2 + y^2}$ and $\cos\phi = x/r$ we get

$$\frac{(x+d)^2}{a^2} + \frac{y^2}{b^2} = 1, \qquad a = \frac{l^2/\mu\alpha}{1-\varepsilon^2}, \quad b = a\sqrt{1-\varepsilon^2}, \quad d = a\varepsilon.$$

This proves Kepler's first law, that gravitational orbits are elliptical. The three parameters are the semimajor axis, the semiminor axis, and the center-focus distance, respectively.

Kepler's second law states that orbits sweep out equal areas in equal times—approximating each dA as a triangle gives

$$dA = \frac{1}{2}r\,(r\,d\phi) \implies \frac{dA}{dt} = \frac{1}{2}r^2\dot{\phi}^2 = \frac{\ell}{2\mu}.$$

As for Kepler's third law we have

$$(\text{period})^2 = \left(\frac{\text{area}}{dA/dt}\right)^2 = \frac{4\pi^2 a^3}{\ell^2}\cdot\mu^2 a(1-\varepsilon^2) = \frac{4\pi^2 a^3}{\ell^2}\frac{\ell^2}{\mu\alpha},$$

meaning $(\text{period})^2 \propto a^3$.

Finally, let's bring this back to our original two-body problem in the center-of-mass frame. Here we have

$$\mathbf{R}_{\text{CM}} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2} = 0;$$

together with $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ we get a system of equations whose solution is

$$\mathbf{r}_1 = \frac{m_2}{m_1+m_2}\mathbf{r}, \qquad \mathbf{r}_2 = -\frac{m_1}{m_1+m_2}\mathbf{r}.$$

So both bodies have the same orbital shape (ellipses with the same orientation and eccentricity), but are opposite in position along those orbits.

As a side note, it turns out (by Bertrand's theorem) that all this theory of closed orbits is only possible for central potentials that look like

$$U(r) = -\frac{\alpha}{r}, \quad U(r) = kr^2.$$

Small deviations from these potentials give *approximately* elliptical orbits. The actual behavior of such a system is quite complex! (Think about phenomena like orbital precession.)

2.3 Particle Scattering

Now we'll move into a discussion about scattering states. First we will show that, very conveniently, all trajectories in a central potential are symmetric about a line between the origin and the point of closest approach. Let B denote this point and let A and C be points on opposite sides of B satisfying $r_A = r_C$. We'll show that the $\Delta\phi$ from B satisfy $\Delta\phi_{BA} = \Delta\phi(CB)$:

$$\Delta\phi_{BA} = -\frac{\ell}{\sqrt{2\mu}}\int_{r_A}^{r_B}\frac{dr'}{r'^2}\frac{1}{\sqrt{E-U_{\text{eff}}(r')}} = -\frac{\ell}{\sqrt{2\mu}}\int_{r_A}^{r_C}(\cdots) = \frac{\ell}{\sqrt{2\mu}}\int_{r_C}^{r_A}(\cdots) = \Delta\phi_{CB}.$$

Combined with the fact that both branches of the trajectory have the same shape, we conclude that the trajectory is symmetric.

Another important fact is that scattering in a central potential is elastic, so kinetic energy is conserved in the infinite "before" and "after" of the collision. This has to do with the fact that $U_{\text{eff}}(\infty)$ must be finite to allow for unbound trajectories; we'll take $U_{\text{eff}}(\infty) = 0$ for the remainder of our discussion.

Each scattering trajectory is determined entirely by the energy E and angular momentum ℓ relative to the center of the potential. We could also choose other pairs of constants, like the initial speed v_∞ and the impact parameter b, the particle's center-offset as it approaches from infinity. (These two pairs of parameters are related by $E = \mu v_\infty^2/2$ and $\ell = \mu v_\infty b$.) The potential then determines the scattering angle θ as a function of these constants.

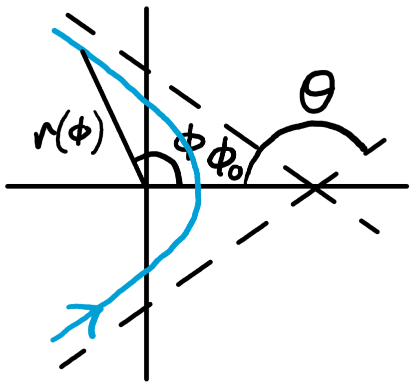
There's a couple of choices of $U(r)$ that are particularly relevant to us. The first is the case of hard sphere scattering, for which $U(r)$ is infinite within a radius a and zero outside. If $b > R$ then the problem is easy—the incident particle never "hits" the potential and $\theta = 0$. Otherwise, if α is the angle the initial trajectory makes with the line of symmetry, then we have $\sin\alpha = b/R$ and

$$\sin\left(\frac{\pi}{2}-\frac{\theta}{2}\right) = \frac{b}{R} \implies \theta = 2\arccos\left(\frac{b}{R}\right).$$

So for a hard sphere the scattering angle is

$$\theta = \begin{cases} 2\arccos(b/R) & b \le R, \\ 0 & b > R. \end{cases}$$

Now let's look at Rutherford scattering, for which $U(r) = -\alpha/r$. We know, from previous discussion, that our particle will follow a hyperbolic trajectory; such a trajectory is drawn at right. We know



$$r(\phi) = \frac{\ell^2}{\mu\alpha}\frac{1}{1+\varepsilon\cos\phi}, \quad \varepsilon \ge 1,$$

and if $r \to \infty$ we get $\cos(\pi - \phi_0) = -1/\varepsilon$ and $\cos\phi_0 = 1/\varepsilon$. We can relate this to our parameters using

$$\varepsilon = \sqrt{1+\frac{2E\ell^2}{\mu\alpha^2}} = \sqrt{1+\frac{\mu^2 b^2 v_\infty^4}{\alpha^2}}.$$

Also, because $1/\varepsilon = \cos\phi_0 = \cos(\pi/2 - \theta/2) = \sin(\theta/2)$,

$$\frac{1}{\sin^2(\theta/2)} = 1 + \cot^2\frac{\theta}{2} = \varepsilon^2 = 1 + \frac{\mu^2 b^2 v_\infty^2}{\alpha^2}.$$

So we end up with the scattering angle

$$\theta = 2\arctan\left(\frac{|\alpha|}{\mu b v_\infty^2}\right);$$

we would've gotten the same result with a repulsive potential rather than an attractive one.

2.4 Beam Scattering

All this theory is great, but it has some practical limitations. In particular, we'd normally employ lots of different particles in any given scattering experiment we'd do, and we can't aim them precisely. So we don't exactly know b, v_∞, and θ for each particle!

To characterize the behavior of a beam of particles, we define the current I and current density $\mathbf{J}$ so that

$$dI = \mathbf{J} \cdot \hat{n}\, d\sigma = J\, d\sigma,$$

where $d\sigma$ is an infinitesimal cross section of the beam. The number of particles incident on an infinitesimal target in a time dt is $N_{\text{inc}} = J\, d\sigma dt$; each $d\sigma$ scatters at a range of angles $[0,\, \theta + d\theta]$.

Suppose we have a detector with area $dA = r^2 d\Omega$ that's just big enough to catch all of these scattered particles when orthogonal to the beam. Here r is the distance from the detector to the potential's center at $r = 0$, and $d\Omega = \sin\theta\, d\theta d\phi$ is the solid angle subtended by the detector when viewed from $r = 0$. There are $N_{\text{det}} = J_{\text{det}}\, dA dt$ particles detected in a time dt, and since $N_{\text{inc}} = N_{\text{det}}$ we have

$$\begin{aligned}
J_{\text{det}} dA &= J_{\text{inc}} d\sigma \\
J_{\text{det}} r^2 d\Omega &= J_{\text{inc}} d\sigma \\
r^2 J_{\text{det}} &= J_{\text{inc}} \cdot \frac{d\sigma}{d\Omega}
\end{aligned}$$

Thus illustrates how the scattered current J_{det} gets thinner with distance, as the particles all scatter at slightly different angles. The $d\sigma/d\Omega$ is called the differential cross section and has units of area—it is a measure of how many particles passing through $d\sigma$ get scattered into $d\Omega$. A larger $d\sigma/d\Omega$ to more scattering!

To compute the differential cross section, suppose a beam is incident upon a potential with a given $b(\theta)$, and consider a width-db ring of particles with impact parameter b. We have $d\sigma = 2\pi b\, db$ and $d\Omega = 2\pi \sin\theta\, d\theta$, meaning

$$\frac{d\sigma}{d\Omega} = \frac{2\pi b\, db}{2\pi \sin\theta\, d\theta} = \frac{b}{\sin\theta} \left| \frac{db}{d\theta} \right|.$$

Let us now define the total cross section

$$\begin{aligned}
\sigma &= \int \left| \frac{d\sigma}{d\Omega} \right| d\Omega \\
&= 2\pi \int_0^\pi \sin\theta \frac{b}{\sin\theta} \left| \frac{db}{d\theta} \right| d\theta \\
&= 2\pi \int_{b_{\min}}^{b_{\max}} b\, db \\
&= \pi \left(b_{\max}^2 - b_{\min}^2 \right).
\end{aligned}$$

For a hard sphere we get $d\sigma/d\Omega = R^2/4$ and $= \pi R^2$. (This illustrates that σ is the "effective area" of the beam that gets scattered at all!) In the case of Rutherford scattering,

$$b(\theta) = \frac{|\alpha|}{\mu v_\infty^2} \cot\frac{\theta}{2}$$

and $\sigma = \infty$ because, strictly speaking, the range of the Coulomb potential is infinite. A more physically interesting quantity is the differential cross section, which turns out to be

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4\mu^2 v_\infty^4} \frac{1}{\sin^4(\theta/2)}.$$

2.5 Visualizing Dynamics

Now we'll take a step back and describe some graphical methods for studying more general dynamical systems. Starting with just one degree of freedom, the state of whatever system we're interested in can be described entirely by the behavior of q and $\dot{q}$. A parametric plot $(q(t), \dot{q}(t))$ of these two quantities is called the system's phase space. In the case of a simple harmonic oscillator, for example, we have the conserved Hamiltonian

$$H = \frac{1}{2}M\dot{\delta}^2 + \frac{1}{2}K\delta^2,$$

which takes the form of an ellipse centered at the origin. (We call the origin here an attractor since the system seems to be oscillating around it.) Note that conservation of the Hamiltonian ensures that there is a one-to-one mapping between H and system trajectories, so given a general time-independent potential $U(x)$ we can think carefully about how the relationship between E and $U(x)$ can provide us information about the system's evolution over time.

It can also be interesting to study how the behavior of a system changes as we vary its parameters. Consider a mass-m bead on a radius-R hoop spinning with angular speed ω. We could show that the equilibrium points of this system are at

$$\theta_{\text{eq}} = 0, \pi, \qquad \cos\theta_{\text{eq}} = \left(\frac{g}{R\omega^2}\right) \equiv \gamma.$$

Our analysis of this system is split into two cases.

- If $\gamma < 1$ then we only have two equilibrium points. The one at the bottom of the hoop is stable, while the one at the top is unstable.
- If $\gamma > 1$ then we have four equilibrium points. The ones at the top and bottom of the hoop are unstable, and the two on the sides are stable.

So as we vary our system's parameters we may create equilibrium points, destroy them, or change their stability. We call such changes bifurcations.

Now let's take a step up and consider a system with N degrees of freedom with state vector

$$\mathbf{x} = (q_1, \ldots, q_N, \dot{q}_1, \ldots, q_N).$$

Taking $n = 2N$ for the remainder of our discussion, the state space of this system is n-dimensional—we get two dimensions for each degree of freedom.

In general this state space is very difficult to conceptualize, even with as few as two degrees of freedom. To visualize things we'll often take an $(n-1)$-dimensional hyperplane S of this space, called a surface of section, and plot all of the points $\mathbf{x}_n$ at which $\mathbf{x}$ intercepts S. These points are related by a Poincaré map P via

$$\mathbf{x}_{k+1}P(\mathbf{x}_k).$$

If $P(\mathbf{x}_\star) = \mathbf{x}_\star$ then $\mathbf{x}_\star$ is called a fixed point of P; the existence of such a point indicates the existence of a closed, periodic trajectory in state space.

Even better, each conserved quantity in our system removes a dimension from state space. For example, the state of a pair of uncoupled pendulums is described entirely by $\mathbf{x} = (\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2)$, but because we can write $\dot{\theta}_2$ in terms of H we can equivalently write $\mathbf{x} = (\theta_1, \theta_2, \dot{\theta}_1, H)$. To capture the periodicity of θ_1, θ_2 in this system we can picture our state space as a tours—θ_1 might circle around the main tube of the torus, while θ_2 circles about its central axis. $\dot{\theta}_1$ is on the minor radius of the torus.

Consider a surface of section S at $\theta_2 = 0$. If the ratio between the frequencies of the pendulums is rational we'll see some regular behavior—only a finite number of points will ever appear on S. But for irrational frequency ratios the points become dense on a circle with radius $\dot{\theta}_1$ (and the trajectories will "fill in" in the torus). This irrational behavior is an example of quasiperiodic motion—the motion is periodic in both θ_1 and θ_2, but not in both.

Note that we call a system like this one separable, since we can look at the behavior of $\theta_1, \dot{\theta}_1$ and $\theta_2, \dot{\theta}_2$ (or θ_1, H_1 and θ_2, H_2) completely separate from one another with no problems.

3 Oscillators

3.1 Oscillations About Equilibria

Now we'll spend some time developing the theory of small oscillations about equilibria. It turns out to be incredibly powerful and far-reaching, as far as physics goes!

The general form of a one-dimensional Lagrangian is

$$L = \frac{1}{2}M(q)\dot{q}^2 - U(q).$$

We know, by now, that every equilibrium point q_0 looks like a harmonic oscillator when we look closely enough; when we define the displacement $\delta \equiv q - q_0$ from equilibrium and do a second-order (to preserve the kinetic energy) Taylor expansion our Lagrangian becomes

$$\begin{aligned} L &= \frac{1}{2}M(q_0+\delta)\dot{\delta}^2 - U(q_0+\delta) \\ &\approx \frac{1}{2}M(q_0)\dot{\delta}^2 - \frac{1}{2}U''(q_0)\delta^2, \end{aligned}$$

plus a constant that vanishes upon finding the Euler-Lagrange equation

$$M(q_0)\ddot{\delta} = -U''(q_0)\delta \implies \ddot{\delta} = \frac{U''(q_0)}{M(q_0)}\delta.$$

For $U''(q_0) < 0$ we get exponential solutions and our implicit $|\delta| \ll q_0$ assumption quickly goes invalid. But $U''(q_0) > 0$ gives oscillations with

$$\omega^2 = \frac{U''(q_0)}{M(q_0)}.$$

This idea of determining the behavior of the system will reappear in our discussion of more complex oscillators.

3.2 Two Coupled Oscillators

Consider two masses m in a space between two walls. Each mass is connected to its respective wall with a spring k, and the masses are connected to one another with a spring k'. The masses' rightward displacements from equilibrium are given by x_1, x_2. The Lagrangian of this setup is

$$L = \frac{1}{2}m\left(\dot{x}_1^2 + \dot{x}_2^2\right) - \frac{1}{2}k\left(x_1^2 + x_2^2\right) - \frac{1}{2}k'\left(x_1 - x_2\right)^2,$$

and the Euler-Lagrange equations are

$$\begin{aligned} m\ddot{x}_1 &= -kx_1 - k'x_1 + k'x_2, \\ m\ddot{x}_2 &= -kx_2 - k'x_2 + k'x_1. \end{aligned}$$

If we define $x_+ \equiv x_1 + x_2$ and $x_- \equiv x_1 - x_2$ then we can decouple these equations:

$$m\ddot{x}_+ = -kx_+, \qquad m\ddot{x}_- = -(k+2k')x_-.$$

So our solutions are

$$\begin{aligned} x_+(t) &= A_+\cos(\omega_+ t + \delta_+), \quad \omega_+^2 \equiv k/m, \\ x_-(t) &= A_-\cos(\omega_- t + \delta_-), \quad \omega_-^2 \equiv (k+2k')/m. \end{aligned}$$

We can write this in a way that's a bit more suggestive of what this is hinting at. If we define a vector $\mathbf{X}(t)$ whose components are $x_1(t)$ and $x_2(t)$ then we have

$$\mathbf{X}(t) = \begin{bmatrix}(x_+ + x_-)/2\\(x_+ - x_-)/2\end{bmatrix} = \frac{A_+}{2}\begin{bmatrix}1\\1\end{bmatrix}\cos(\omega_+ t + \delta_+) + \frac{A_-}{2}\begin{bmatrix}1\\-1\end{bmatrix}\cos(\omega_- t + \delta_-).$$

Each term here corresponds to an eigenmode of our system: the first represents the "together" motion of the two masses, while the second corresponds to the "apart" motion.

It will soon become useful for us to have this solution in complex form. We can write

$$\mathbf{X}(t) = \mathrm{Re}\left(\frac{A_+}{2}\begin{bmatrix}1\\1\end{bmatrix}e^{i\omega_+ t}e^{i\delta_+} + \frac{A_-}{2}\begin{bmatrix}1\\-1\end{bmatrix}e^{i\omega_- t}e^{i\delta_-}\right),$$

and if we define $C_+ = (A_+/2)e^{i\delta_+}$ with the first normal mode $\mathbf{Z}_+$ (along with analogous quantities $C_-, \mathbf{Z}_-$),

$$= \mathrm{Re}\left(C_+\mathbf{Z}_+e^{i\omega_+ t} + C_-\mathbf{Z}_-e^{i\omega_- t}\right).$$

We've gone from four free parameters to just two! (This is just taking advantage of the fact that one complex number encodes the same amount of information as two real numbers.)

An Eigenvalue Problem

Now, we could've come to all the same conclusions in a much more streamlined way using some linear algebra. Notice that we could've written our original (coupled) Euler-Lagrange equations as

$$\ddot{\mathbf{X}} = -\frac{1}{m}\mathbf{K}\mathbf{X}, \qquad \mathbf{K} = \begin{bmatrix}k+k' & -k'\\-k' & k+k'\end{bmatrix}.$$

Rather than solve this problem directly, we'll say that $\mathbf{X}(t) = \mathrm{Re}(\mathbf{Z}(\mathbf{t}))$ for some complex-valued function $\mathbf{Z}$ and focus on $\ddot{\mathbf{Z}} = -(1/m)\mathbf{K}\mathbf{Z}$. Substituting the ansatz $\mathbf{Z}(t) = \mathbf{Z}_0e^{i\omega t}$ for some constant vector $\mathbf{Z}_0$ gives

$$\begin{aligned}(i\omega)^2\mathbf{Z}_0e^{i\omega t} &= -\frac{1}{m}\mathbf{K}\mathbf{Z}_0e^{i\omega t},\\ m\omega^2\mathbf{Z}_0 &= \mathbf{K}\mathbf{Z}_0.\end{aligned}$$

This is a classic eigenvalue problem! Going through the motions tells us that we have eigenvalues when

$$\omega^2 = \frac{k}{m} \text{ or } \omega^2 = \frac{k+2k'}{m},$$

just like we've already found. The corresponding eigenvectors are the $\mathbf{Z}_+$ and $\mathbf{Z}_-$ from before.

Weak Coupling

Now let's quickly look at the weak coupling limit $k' \ll k$. Here we have

$$\omega_2 = \sqrt{\frac{k}{m}}\sqrt{1+\frac{2k'}{k}} \approx \sqrt{\frac{k}{m}}\left(1+\frac{1}{2}\frac{2k'}{k}\right) = \omega_1 + \frac{k'}{\sqrt{mk}}.$$

If $\omega_0 = (\omega_+ + \omega_-)/2$ and $\Delta\omega = \omega_- - \omega_+$ then our solution from before becomes

$$\mathbf{Z}(t) = e^{i\omega_0 t}\left(C_1\begin{bmatrix}1\\1\end{bmatrix}e^{-i\Delta\omega\, t} + C_2\begin{bmatrix}1\\-1\end{bmatrix}e^{i\Delta\omega\, t}\right).$$

If mass 2 initially has zero displacement then substituting $x_2(0) = 0$ gives $C_1 = C_2$ and

$$\mathbf{Z}(t) = C_1e^{i\omega_0 t}\begin{bmatrix}e^{-i\Delta\omega\, t} + e^{i\Delta\omega\, t}\\e^{-i\Delta\omega\, t} - e^{i\Delta\omega\, t}\end{bmatrix} = 2C_1e^{i\omega_0 t}\begin{bmatrix}\cos\Delta\omega\, t\\-i\sin\Delta\omega\, t\end{bmatrix}.$$

Defining $C_1 = D_1e^{i\delta_1}$ for real D_1, δ_1 gives

$$\mathbf{X}(t) = \mathrm{Re}(\mathbf{Z}(t)) = D_1\begin{bmatrix}\cos(\omega_0 t + \delta_1)\cos(\Delta\omega\, t)\\\sin(\omega_0 t + \delta_1)\sin(\Delta\omega\, t)\end{bmatrix}.$$

So $x_1(t)$ and $x_2(t)$ both look like beats with out-of-phase envelopes!

3.3 Several Coupled Oscillators

Now let's look at a system of N coupled oscillators near equilibrium (at $q_1 = \cdots = 0$). To second order,

$$U(q) \approx U(0) + \frac{1}{2}\sum_{i,j} q_i q_j \frac{\partial^2 U(0)}{\partial q_i \partial q_j},$$

where the linear terms have gone to zero because $U(0)$ is a minimum. We'll define a symmetric matrix $\mathbf{K}$ with entries

$$\mathbf{K}_{ij} = \frac{\partial^2 U(0)}{\partial q_i \partial q_j}$$

so we can write

$$U(q) = \frac{1}{2}\sum_{i,j} q_I \mathbf{K}_{ij} q_j = \frac{1}{2}\mathbf{X}^T\mathbf{K}\mathbf{X},$$

where $\mathbf{X}$ is a coordinate vector with components $q_1, \ldots, q_N$. We can do a similar thing with the kinetic energy which, including only the second-order terms, looks like

$$\begin{aligned}
T &\approx \frac{1}{2}\mathbf{M}(q)_{11}\dot{q}_1^2 + \frac{1}{2}\mathbf{M}(q)_{12}\dot{q}_1\dot{q}_2 + \frac{1}{2}\mathbf{M}(q)_{13}\dot{q}_1\dot{q}_3 + \cdots \\
&= \frac{1}{2}\sum_{i,j}\dot{q}_i \mathbf{M}_{ij}(0)\dot{q}_j \\
&= \frac{1}{2}\dot{\mathbf{X}}^T\mathbf{M}(0)\dot{\mathbf{X}}, \qquad \mathbf{M}_{ij} = \frac{\partial T}{\partial \dot{q}_i \partial \dot{q}_j}.
\end{aligned}$$

So the Lagrangian is

$$L = \frac{1}{2}\sum_{i,j}\dot{q}_i \mathbf{M}_{ij}\dot{q}_j - \frac{1}{2}\sum_{i,j} q_i \mathbf{K}_{ij} q_j.$$

Let's now derive the Euler-Lagrange equation for q_k. We have

$$\begin{aligned}
\frac{\partial L}{\partial q_k} &= -\frac{1}{2}\sum_{i,j}\mathbf{K}_{ij}\frac{\partial}{\partial q_k}(q_i q_j) \\
&= -\frac{1}{2}\sum_{i,j} K_{ij}\left(q_j \delta_{ik} + q_i \delta_{jk}\right) \\
&= -\frac{1}{2}\sum_{j} K_{kj} q_j - \frac{1}{2}\sum_{i} K_{ik} q_i \\
&= -(\mathbf{K}\mathbf{X})_k
\end{aligned}$$

We could similarly show that $\partial L/\partial \dot{q}_k = (\mathbf{M}\dot{\mathbf{X}})_k$, meaning $(\mathbf{M}\ddot{\mathbf{X}})_k = -(\mathbf{K}\mathbf{X})_k$ and, in all, the Euler-Lagrange equations are

$$\mathbf{M}\ddot{\mathbf{X}} = -\mathbf{K}\mathbf{X}.$$

When we go to solve these, we once again define a complex-valued $\mathbf{Z}$ satisfying $\mathbf{X}(t) = \mathrm{Re}(\mathbf{Z}(t))$ to write $\mathbf{M}\ddot{\mathbf{Z}} = -\mathbf{K}\mathbf{Z}$. Substituting the ansatz $\mathbf{Z}(t) = \mathbf{Z}_0 e^{i\omega t}$ gives

$$\omega^2\mathbf{M}\mathbf{Z}_0 = \mathbf{K}\mathbf{Z}_0.$$

This is what we call a generalized eigenvalue problem! We'd proceed by solving $\det(\mathbf{K} - \omega^2\mathbf{M}) = 0$ and them finding the null space of $\mathbf{K} - \omega^2\mathbf{M}$ for each of the resulting eigenvalues. Note that each of the eigenvalues ω^2 will turn out to be real and positive, and that all of the eigenvectors $\mathbf{Z}_i$ satisfy $\mathbf{Z}_i^T\mathbf{M}\mathbf{Z}_j = 0$. (This is what we'll mean by orthogonality in this context.)

3.4 Linear Resonance

Consider, now, an isolated damped oscillator with mass m, spring constant k, damping coefficient b, and driving force F_d. By Newton's second law we have

$$
\begin{aligned}
m\ddot{x} + b\dot{x} + kx &= F_d, \\
\ddot{x} + 2\beta\dot{x} + \omega_0^2 x &= \frac{F_d}{m},
\end{aligned}
$$

where $2\beta \equiv b/m$ and $\omega_0^2 \equiv k/m$. We could show that the homogeneous solutions to this differential equation decay exponentially in time, so we'll ignore them in this discussion. As for the particular solution, take $F_d(t) = \cos(\omega t)$ and define $x(t) = \mathrm{Re}(z(t))$, turning the equation into

$$\ddot{z} + 2\beta\dot{z} + \omega_0^2 z = f e^{i\omega t}.$$

Substituting the ansatz $z = Ae^{i\omega t}$ yields

$$A = \frac{f}{\omega_0^2 - \omega^2 + 2i\beta\omega} = \frac{f(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + 4\beta^2\omega^2} - i\frac{2\beta\omega f}{(\omega_0^2 - \omega^2)^2 + 4\beta^2\omega^2},$$

where in the last expression we'e multiplied by the conjugate of the first denominator. So if $A = a + bi$ then we could show that our solution looks like

$$
\begin{aligned}
x_p(t) &= \mathrm{Re}\left[(a + bi)(\cos\omega t + i\sin\omega t)\right] \\
&= C\cos(\omega t + \delta),
\end{aligned}
$$

where

$$C = \frac{f}{\left[(\omega_0^2 - \omega^2)^2 + 4\beta^2\omega^2\right]^{1/2}}, \qquad \tan\delta = \frac{2\beta\omega}{\omega_0^2 - \omega^2}.$$

Note that C is maximized at $\omega = \omega_0$, and that the maximum value is $C_{\max} = f/2\beta\omega_0$. But ω_0 is not the only frequency that gets amplified—to communicate the "characteristic width" of amplified frequencies, we'll say we want $(\omega_0^2 - \omega^2)^2 \sim 4\beta^2\omega^2$ and so

$$(\omega_0 + \omega)(\omega_0 - \omega) \sim 2\beta\omega \implies (\omega_0 - \omega) \sim \beta$$

if we take $\omega_0 + \omega \sim 2\omega$. So 2β, in a sense, communicates the width of frequencies that get amplified! The output phase changes with different input frequencies, too—in certain limiting cases we have

$$\delta = \begin{cases} 0 & \omega \ll \omega_0, \\ \pi & \omega \gg \omega_0, \\ \pi/2 & \omega = \omega_0. \end{cases}$$

4 Rigid Bodies

4.1 Rotational Energy and Momentum

The theory of oscillations, coupled or otherwise, proves to be very useful in simplifying the study of macroscopic systems. Another way we might study these is by assuming the particles comprise a rigid body, in which case the particles are said to be fixed in place. (In principle this isn't really true, of course, but we're not going to write down a Lagrangian with 10^{23} degrees of freedom to accommodate this technicality.) This leaves us with six degrees of freedom: three for the body's position and three for its orientation.

Take our object and break it into chunks indexed by α; the object's kinetic energy is

$$T = \frac{1}{2}\sum_\alpha \Delta m_\alpha |\mathbf{r}_\alpha|^2.$$

We'd like to break this up into translational and rotational components. To this end, we'll introduce a "tagged point" that specifies the position of the object. Let $\mathbf{R}$ be the position of the tagged point and let $\mathbf{r}'_\alpha$ be the position of α relative to the tagged point; with these our kinetic energy becomes

$$= \frac{1}{2}\sum_\alpha \Delta m_\alpha |\dot{\mathbf{R}} + \dot{\mathbf{r}}'_\alpha|^2.$$

Now, $\mathbf{r}'_\alpha$ has constant magnitude for a rigid body, meaning $\dot{\mathbf{r}}'_\alpha$ is pure rotation. In particular, if the object has angular velocity $\boldsymbol{\omega}$ at an angle θ to $\mathbf{r}'_\alpha$ then $|\dot{\mathbf{r}}'_\alpha| = (r'_\alpha \sin\theta)\,\omega = |\boldsymbol{\omega} \times \mathbf{r}'_\alpha|$ and, by the right-hand rule, $\dot{\mathbf{r}}'_\alpha = \boldsymbol{\omega} \times \mathbf{r}'_\alpha$. So the kinetic energy becomes

$$= \frac{1}{2}\sum_\alpha \Delta m_\alpha |\dot{\mathbf{R}} + (\boldsymbol{\omega} \times \mathbf{r}'_\alpha)|^2.$$

Expanding and manipulating turns this into

$$T = \frac{1}{2}M|\dot{\mathbf{R}}|^2 + \dot{\mathbf{R}} \cdot \left[\boldsymbol{\omega} \times \left(\sum_\alpha \Delta m_\alpha \mathbf{r}'_\alpha\right)\right] + \frac{1}{2}\sum_\alpha \Delta m_\alpha |\boldsymbol{\omega} \times \mathbf{r}'_\alpha|^2,$$

where $M = \sum_\alpha \Delta m_\alpha$ is the mass of the body. Notice that the first and third terms are purely translation and purely rotational, respectively, and that the middle term goes to zero if $\dot{\mathbf{R}} = 0$ or $\mathbf{R} = \mathbf{R}_{\mathrm{CM}}$. One of these will always be true for us, so we can ignore this term and write

$$T = \frac{1}{2}M|\dot{\mathbf{R}}|^2 + \frac{1}{2}\sum_\alpha \Delta m_\alpha (\boldsymbol{\omega} \times \mathbf{r}'_\alpha) \cdot (\boldsymbol{\omega} \times \mathbf{r}'_\alpha).$$

Let's now focus our attention on the rotational component, which we can write as

$$T_{\mathrm{rot}} = \frac{1}{2}\sum_\alpha \Delta m_\alpha \big[\omega^2 r_\alpha^2 - (\boldsymbol{\omega} \cdot \mathbf{r}'_\alpha)^2\big]$$

or, in the continuum limit,

$$= \frac{1}{2}\int \rho(\mathbf{r}')\, d^3\mathbf{r}' \big[\omega^2 r'^2 - (\boldsymbol{\omega} \cdot \mathbf{r}')^2\big].$$

We'll write our dot products in index notation via

$$\omega^2 = \sum_{i,j} \omega_i \delta_{ij} \omega_j, \qquad (\boldsymbol{\omega} \cdot \mathbf{r}')^2 = \sum_{i,j} (\omega_i r'_i)(\omega_j r'_j),$$

and substitute them into $T_{\rm rot}$ to get

$$\begin{aligned} T_{\rm rot} &= \frac{1}{2}\sum_{i,j}\int \rho(\mathbf{r}')\,d^3\mathbf{r}'\left[\omega_i\omega_j r'^2\delta_{ij} - \omega_i\omega_j r_i' r_j'\right] \\ &= \frac{1}{2}\sum_{i,j}\omega_i\omega_j\left[\int \rho(\mathbf{r}')\,d^3\mathbf{r}'\left(r'^2\delta_{ij} - r_i' r_j'\right)\right]. \end{aligned}$$

The bracketed expression is the (i,j) component of the moment of inertia tensor $\mathbf{I}$. So the rotational kinetic energy becomes

$$T_{\rm rot} = \frac{1}{2}\sum_{i,j}\mathbf{I}_{ij}\omega_i\omega_j.$$

Notice how this reduces to the familiar $I\omega^2/2$ for rotation about a single fixed axis. Now, the inertia tensor $\mathbf{I}$ is a (time-dependent) symmetric matrix, and its diagonal components look like

$$\mathbf{I}_{xx} = \int \rho(\mathbf{r}')\,d^3\mathbf{r}'\left[r'^2\cdot 1 - x^2\right] = \int \rho(\mathbf{r}')\,d^3\mathbf{r}'(y^2+z^2).$$

Physically, this is a projection of $\mathbf{r}'$ into the yz-plane a measure of how resistant our object is to rotations about $\hat{x}$. The off-diagonal terms, however, have no clear physical meaning other than the fact that nonzero values tend to couple rotations about two axes—if we try to generate a rotation about $\hat{x}$ then we'd also get one about $\hat{y}$. (We'll soon find that this happens when the object at hand lacks some kind of symmetry.)

For angular momentum we get a similar result by taking

$$\boldsymbol{\ell} = \sum_\alpha \Delta m_\alpha(\mathbf{r}_\alpha\times\dot{\mathbf{r}}_\alpha) = \sum_\alpha \Delta m_\alpha(\mathbf{R}+\mathbf{r}_\alpha')\times(\dot{\mathbf{R}}+\boldsymbol{\omega}\times\mathbf{r}_\alpha'),$$

from which one could deduce that

$$\ell_i = M(\mathbf{R}\times\dot{\mathbf{R}}) + \sum_j \mathbf{I}_{ij}\omega_j.$$

Note that this means conservation of angular momentum doesn't imply conservation of $\boldsymbol{\omega}$.

4.2 Principal Axes

The fact that the inertia tensor has off-diagonal entries suggests that the Cartesian basis isn't the most "natural" one for this problem. We'd like to have one in which $\boldsymbol{\ell} = \lambda\boldsymbol{\omega}$ if $\boldsymbol{\omega}$ points along one of the axes; since $\boldsymbol{\ell} = \mathbf{I}\boldsymbol{\omega}$ for non-translational frames we can write $\mathbf{I}\boldsymbol{\omega} = \lambda\boldsymbol{\omega}$. This is an eigenvalue problem, and because $\mathbf{I}$ is symmetric we get real eigenvalues and orthogonal eigenvectors. The eigenvalues λ_i are the principal moments of our object, and the eigenvectors $\hat{e}_i$ are its principal axes. Define $\hat{x}' = \hat{e}_1$, $\hat{y}' = \hat{e}_2$, and $\hat{z}' = \hat{e}_3$.

We can see that if $\boldsymbol{\omega}$ points along a principal axis then $\boldsymbol{\ell}\propto\boldsymbol{\omega}$ and $\boldsymbol{\ell}$-conservation implies $\boldsymbol{\omega}$-conservation. More generally,

$$\boldsymbol{\ell} = \mathbf{I}\boldsymbol{\omega} = \lambda_{x'}\omega_{x'}\hat{x}' + \lambda_{y'}\omega_{y'}\hat{y}' + \lambda_{z'}\omega_{z'}\hat{z}'.$$

This indicates that, in the principal axis basis, the inertia tensor $\mathbf{I}$ is diagonal with entries $\lambda_{x'}, \lambda_{y'}, \lambda_{z'}$. This gives us

$$T = \frac{1}{2}\sum_{i',j'}\omega_{i'}\mathbf{I}_{i'j'}\omega_{j'} = \frac{1}{2}\left(\lambda_{x'}\omega_{x'}^2 + \lambda_{y'}\omega_{y'}^2 + \lambda_{z'}\omega_{z'}^2\right),$$

which looks a lot more like what we'd expect! So principal axes serve to reduce the number of components of $\mathbf{I}$ we need to deal with, which serves to make all the math a bit more physically intuitive. Unfortunately, an object's principal axes are intrinsic to its shape and so vary with time as the body spins. (This doesn't affect our calculation for T, though, since it's coordinate-independent and we did it in an inertial frame.)

Symmetry makes it very easy to identify some or all of an object's principal axes. It isn't difficult to show that the existence of a reflection symmetry, say $x\to -x$, implies that $\hat{x}$ is a principal axis; it follows that I_{xx} is a principal moment and $I_{xy} = I_{xz} = 0$. Similarly, if there is a rotational axis of symmetry then that axis is a principal axis, and any other pair of mutually orthogonal axes are also principal.

4.3 Rotating Frames of Reference

Motivated by the fact that a body's principal axes generally define a frame of reference that rotates relative to the body's surroundings, we'll spend some time talking about how to work in rotating frames in general.

Consider a reference frame that rotates with angular velocity $\boldsymbol{\omega}$ relative to an inertial frame, and suppose a vector $\mathbf{A}$ is stationary in this rotating frame. If $\mathbf{A}$ makes an angle θ with $\boldsymbol{\omega}$ then in a time dt we have $(dA)_{\text{in}} = (A\sin\theta)\omega\, dt$ and $(d\mathbf{A}/dt)_{\text{in}} = \boldsymbol{\omega}\times\mathbf{A}$. But if $\mathbf{A}$ is not stationary in the rotating frame then by the Galilean velocity transformation

$$\left(\frac{d\mathbf{A}}{dt}\right)_{\text{in}} = \left(\frac{d\mathbf{A}}{dt}\right)_{\text{rot}} + (\boldsymbol{\omega}\times\mathbf{A}).$$

Now we'll derive the equation of motion for a point mass in a rotating frame. This requires finding the Lagrangian, which should be written down in an inertial frame. The kinetic energy in such a frame is

$$\begin{aligned} T &= \frac{1}{2}m\big[\mathbf{v}_{\text{rot}} + (\boldsymbol{\omega}\times\mathbf{r})\big]^2 \\ &= \frac{1}{2}m\big[\mathbf{v}_{\text{rot}}\cdot\mathbf{v}_{\text{rot}} + 2\mathbf{v}_{\text{rot}}\cdot(\boldsymbol{\omega}\times\mathbf{r}) + \omega^2 r^2 - (\boldsymbol{\omega}\cdot\mathbf{r})^2\big] \\ &= \frac{1}{2}m\left[\sum_k v_{\text{rot},\,k}^2 + 2\sum_k v_{\text{rot},\,k}(\boldsymbol{\omega}\times\mathbf{r})_k + \omega^2\sum_k r_k^2 - \sum_{j,k}\omega_j\omega_k r_j r_k\right], \end{aligned}$$

and the Lagrangian is $L = T - U(\mathbf{r})$. Each component of the Euler-Lagrange equations looks like

$$\frac{d}{dt}\frac{\partial L}{\partial v_{\text{rot},\,i}} = \frac{\partial L}{\partial r_i},$$

and computing these would give the equation of motion

$$m\ddot{\mathbf{r}}_{\text{rot}} = -\nabla U - m\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}_{\text{rot}}) - 2m\boldsymbol{\omega}\times\mathbf{v}_{\text{rot}} - m(\dot{\boldsymbol{\omega}}\times\mathbf{r}_{\text{rot}}).$$

The $-\nabla U$ here is the inertial force, and it's all we'd see in an inertial frame. The second term, called the centrifugal force, tends to push away from the axis of rotation; we could demonstrate this using the right-hand rule. If we're at the equator of a planet with mass M and radius R, we could show that

$$\frac{|\mathbf{F}|_{\text{cent}}}{|\mathbf{F}|_{\text{grav}}} = \frac{\omega^2 r^3}{GM}.$$

For Earth this evaluates to about 0.3%. The third is the Coriolis force, and the fourth is the Euler term. (We generally ignore this term since $\dot{\boldsymbol{\omega}}$ is usually either zero or near zero.)

4.4 Euler's Equations

We're now in a position to determine the equations of motion for a rotating rigid body. If a torque $\boldsymbol{\Gamma}$ acts in the inertial frame then by the rotational vector transformation we have

$$\dot{\mathbf{L}}_{\text{rot}} + \boldsymbol{\omega}\times\mathbf{L}_{\text{rot}} = \boldsymbol{\Gamma}.$$

In the case of free rotation ($\boldsymbol{\Gamma} = 0$) we can expand both nonzero terms to get Euler's equations:

$$\begin{aligned} \lambda_1\dot{\omega}_{x'} - (\lambda_2-\lambda_3)\omega_{y'}\omega_{z'} &= 0, \\ \lambda_2\dot{\omega}_{y'} - (\lambda_3-\lambda_1)\omega_{x'}\omega_{z'} &= 0, \\ \lambda_3\dot{\omega}_{z'} - (\lambda_1-\lambda_2)\omega_{x'}\omega_{y'} &= 0. \end{aligned}$$

We can quickly leverage this set of equations to learn something about the stability of rotations about each principal axis. If we have rotation purely about $\hat{z}'$ and introduce a small perturbation so that $\omega_{z'}\approx 0$ is preserved, the first of Euler's equations gives

$$\lambda_1\ddot{\omega}_{x'} = (\lambda_2-\lambda_3)[\omega_{z'}\dot{\omega}_{y'} + \dot{\omega}_{z'}\omega_{y'}]$$

and after taking $\dot\omega_{z'} = 0$ and substituting for $\dot\omega_{y'}$ we get

$$\ddot\omega_{x'} = \frac{(\lambda_2 - \lambda_3)(\lambda_3 - \lambda_1)\omega_{z'}^2}{\lambda_1\lambda_2}\omega_{x'}.$$

Notice that, if $\lambda_1 < \lambda_3 < \lambda_2$, the above coefficient is positive and ω_x' grows exponentially! In this case rotation about $\hat z'$ is unstable. More generally, rotations about principal axes with "intermediate" moments are unstable; this is called the tennis racket theorem.

Now let's make a huge simplification and assume our body has cylindrical symmetry, so $\lambda_1 = \lambda_2 = \lambda$. Looking at the Euler equations, we can see that this makes $\omega_{z'}$ constant, so that nonlinear system of equations has become linear! Specifically, we get

$$\dot\omega_{x'} = \Omega_b\omega_{y'}, \quad \dot\omega_{y'} = -\Omega_b\omega_{x'}, \qquad \Omega_b \equiv \frac{(\lambda - \lambda_3)\omega_{z'}}{\lambda}.$$

We could combine these equations to get $\ddot\omega_{x'} = -\Omega_b^2\omega_{x'}$ which, with the initial conditions $\omega_{x'}(0) = \omega_0$ and $\omega_{y'}(0) = 0$, gives

$$\omega_{x'}(t) = \omega_0\cos\Omega_b t, \qquad \omega_{y'}(t) = -\omega_0\sin\Omega_b t.$$

So both $\boldsymbol\omega$ and $\boldsymbol\ell$ precess about $\hat z'$ at a rate Ω_b:

$$\begin{aligned}
\boldsymbol\omega &= \omega_0\cos\Omega_b t\,\hat x' - \omega_0\sin\Omega_b t\,\hat y' + \omega_{z'}\hat z',\\
\boldsymbol\ell &= \lambda\big[\omega_0\cos\Omega_b t\,\hat x' - \omega_0\sin\Omega_b t\,\hat y'\big] + \lambda_3\omega_{z'}\hat z'.
\end{aligned}$$

But remember, this is what the body sees in its own, rotating frame of reference. In the inertial "space frame" we note that $\boldsymbol\ell$ is fixed, that it makes a constant angle with $\boldsymbol\omega$, and that these two vectors lay in the same plane as $\hat z'$. Put together, it turns out that both $\hat z'$ and $\boldsymbol\omega$ precess about $\boldsymbol\ell$ at a rate $\Omega_s = \ell/\lambda$! (This will be easier to show via the Lagrangian approach that we'll develop soon.)

4.5 Euler Angles

The Euler equations do a good job at describing motion in the body frame, but we really seek equations of motion relative to the inertial frame. To make this happen we'll need to define a set of generalized coordinates that specify the orientation of our body relative to such a frame. We'll once again focus on objects with cylindrical symmetry.

The standard choice is to use the Euler angles, which are defined as follows. Start with the principal axes aligned with the inertial-frame axes.

- Rotate the body an angle ϕ about the z-axis.
- Tilt the body an angle θ from the z-axis (about $\hat x$').
- Spin the body an angle ψ about the z'-axis.

In practical terms, θ and ϕ are the polar and azimuthal angles of the z'-axis, while ψ is the amount that $\hat x'$ has rotated out of the xy-plane due to the spin of the object about $\hat z'$.

This gives rise to three kinds of motion: spin ($\dot\psi$), precession ($\dot\phi$), and nutation ($\dot\theta$). Each of these contributes to the full rotation of the body in different ways.

- ($\dot\psi$) This is purely rotation about $\hat z'$, so $\omega_{z'(\psi)} = \dot\psi$.
- ($\dot\phi$) This is rotation about $\hat z = \hat z'\cos\theta + \hat y'\sin\theta$, so $\omega_{z'(\phi)} = \dot\phi\cos\theta$ and $\omega_{y'(\phi)} = \dot\phi\sin\theta$.
- ($\dot\theta$) By cylindrical symmetry we can choose $\hat x'$ to lie in the xy-plane, in which case $\dot\theta$ is purely rotation about $\hat x'$ meaning $\omega_{x'(\theta)} = \dot\theta$.

With this we have the kinetic energy

$$\begin{aligned}
T &= \frac{1}{2}\left(\lambda\omega_{x'}^2 + \lambda\omega_{y'}^2 + \lambda_3\omega_{z'}^2\right)\\
&= \frac{1}{2}\lambda\left(\dot\theta^2 + \dot\phi^2\sin^2\theta\right) + \frac{1}{2}\lambda_3\left(\dot\psi + \dot\phi\,\cos\theta\right)^2
\end{aligned}$$

which, along with a potential, is just what we need to write down a Lagrangian! If we again focus our attention on the free rotation case (in which $L = T$) we can immediately see a couple of conserved quantities:

$$\begin{aligned} p_\psi &= \frac{\partial L}{\partial \dot\psi} = \lambda_3 \left(\dot\psi + \dot\phi \cos\theta\right), \\ p_\phi &= \frac{\partial L}{\partial \dot\phi} = \lambda \dot\phi \sin^2\theta + \lambda_3 \cos\theta \left(\dot\psi + \dot\phi\cos\theta\right). \end{aligned}$$

Also, if we define $\hat{z}$ such that $\boldsymbol{\ell} = \ell \hat{z}$, then by conservation of angular momentum we have

$$\begin{aligned} \ell\, \hat{z} &= \lambda\omega_{x'}\hat{x}' + \lambda\omega_{y'}\hat{y}' + \lambda\omega_{z'}\hat{z}', \\ \ell\cos\theta\,\hat{z}' + \ell\sin\theta\,\hat{y}' &= \lambda\dot\theta\,\hat{x}' + \lambda\dot\phi\sin\theta\,\hat{y}' + \lambda_3(\dot\psi + \dot\phi\cos\theta)\hat{z}'. \end{aligned}$$

From this it follows that $\dot\theta = 0$, so we have no nutation, and $\dot\phi = \ell/\lambda$, which is exactly the precession frequency we pointed out previously. We can also say something about spin—since $\ell\cos\theta = \lambda_3(\dot\psi + \dot\phi\cos\theta)$,

$$\dot\psi = \frac{1}{\lambda_3}\left(\ell\cos\theta - \lambda_3\dot\phi\cos\theta\right) = \frac{1}{\lambda_3}\left(\ell\cos\theta - \lambda_3\frac{\ell}{\lambda}\cos\theta\right) = \frac{\ell\cos\theta}{\lambda\lambda_3}(\lambda - \lambda_3),$$

and because $\ell\cos\theta = \lambda_3\omega_{z'}$, this becomes

$$\dot\psi = \frac{\omega_{z'}(\lambda - \lambda_3)}{\lambda},$$

which is exactly the frequency at which we found $\boldsymbol{\ell}$ to precess in the body frame.

We'll finish off with physical interpretations for the generalized momenta p_ψ, p_ϕ, along with a relationship between the two. In the body frame we have

$$\begin{aligned} p_\psi &= \lambda_3\left(\dot\psi + \dot\phi\cos\theta\right) = \lambda_{z'}, \\ p_\phi &= \lambda\dot\phi\sin^2\theta + \lambda_3\cos\theta\left(\dot\psi + \dot\phi\cos\theta\right) = \ell_{y'}\sin\theta + \ell_{z'}\cos\theta, \end{aligned}$$

and by moving into the inertial frame with $\boldsymbol{\ell} = \ell\hat{z}$ we get $p_\psi = \ell\cos\theta$ and

$$p_\phi = \ell\sin^2\theta + \ell\cos^2\theta = \ell.$$

Note that $p_\phi\cos\theta = p_\psi$ here. Lastly, let's take a look at the Euler-Lagrange equation for θ:

$$\begin{aligned} \lambda\ddot\theta &= \lambda\dot\phi^2\sin\theta\cos\theta + \lambda_3\left(\dot\psi + \dot\phi\cos\theta\right)\left(-\dot\phi\sin\theta\right) \\ &= \sin\theta\dot\phi\left(\lambda\dot\phi\cos\theta - p_\psi\right) \end{aligned}$$

Because $p_\phi = \lambda\sin^2\theta\,\dot\phi + p_\psi\cos\theta$, we can substitute for the innner $\dot\phi$ and, after some manipulation, get

$$= \frac{\dot\phi}{\sin\theta}\left(p_\phi\cos\theta - p_\psi\right).$$

So we can see that $\ddot\theta = 0$ if $p_\phi\cos\theta = p_\psi$, which occurs when $\boldsymbol{\ell}$ aligns with $\hat{z}$.

4.6 Gyroscope Precession

Now we'll consider a problem in which there is an external torque—a gyroscope under the influence of gravity. There's actually only a couple of modifications we'll need to make from the free rotation problem.

- Our tagged point will be the stationary point of the gyroscope in contact with the ground. This turns out to not affect our kinetic energy T at all if we go through the same motions as before.
- If R is the distance from the tagged point to the center of mass then $U = MgR\cos\theta$.

With these the Lagrangian is $L = T - U$ and the Euler-Lagrange equation in θ is

$$\lambda\ddot{\theta} = \lambda\dot{\phi}^2 \sin\theta\cos\theta - \lambda_3\left(\dot{\psi} + \dot{\phi}\cos\theta\right)\dot{\phi}\sin\theta + MgR\sin\theta.$$

In the case of no nutation ($\dot{\theta} = 0$) we get $\lambda\dot{\phi}^2\cos\theta - p_\psi\dot{\phi} + MgR = 0$. If p_ψ is large (i.e., if the top is spinning very quickly), solving the quadratic for $\dot{\phi}$ would give

$$\dot{\phi} = \frac{MgR}{p_\psi} \quad \text{or} \quad \dot{\phi} = \frac{p_\psi}{\lambda\cos\theta}$$

as the respective slow and fast solutions. This is called steady precession—all angles change at constant rates! Now, if there is nutation we'll need to use the fact that the Hamiltonian is conserved and equal to

$$H = \frac{1}{2}\lambda\left(\dot{\phi}^2\sin^2\theta + \dot{\theta}^2\right) + \frac{1}{2}\lambda_3\left(\dot{\psi} + \dot{\phi}\cos\theta\right)^2 + MgR\cos\theta.$$

Noting that p_ϕ and p_ψ are conserved, we can substitute $p_\psi = \lambda_3(\dot{\psi} + \dot{\phi}\cos\theta)$ and $\lambda\dot{\phi}\sin^2\theta = p_\phi - p_\psi\cos\theta$ to get

$$H = \frac{1}{2}\lambda\dot{\theta}^2 + \frac{(p_\phi - p_\psi\cos\theta)^2}{2\lambda\sin^2\theta} + \frac{p_\psi^2}{2\lambda_3} + MgR\cos\theta,$$

which allows us to define an effective potential as before. Solving as we did for the two-body problem would reveal that

$$\dot{\phi} = \frac{p_\phi - p_\psi\cos\theta}{\lambda\sin^2\theta}.$$

Notice that if $p_\phi > p_\psi$ then $\dot{\theta}$ is always positive, but if $p_\phi < p_\psi$ then the sign of $\dot{\phi}$ may change depending on the values of $p_\psi - p_\psi\cos\theta$ at the minimum and maximum θ.

5 Hamiltonian Mechanics

5.1 Hamilton's Equations

Now we'll go back to studying the underlying framework of classical mechanics, this time based on the Hamiltonian and its dependencies on positions and momenta. This Hamiltonian formulation turns out to have a lot of advantages (and disadvantages) when compared with the Lagrangian approach:

- it leads to a system of first-order differential equations rather than second-order ones, which are numerically easier objects to work with;
- it allows for a cleaner treatment of conserved quantities since any conserved momenta actually appear in the Hamiltonian;
- it ends up being practically easier in fields like statistical and quantum mechanics; and
- it puts the positions and momenta on similar footing, allowing for much more flexibility in the coordinates we choose to work with.

We must first establish that the Hamiltonian actually depends only on positions and momenta; that is, that it's independent of any velocities. For a system with one degree of freedom,

$$\begin{aligned} H(p, q, \dot{q}, t) &= p\dot{q} - L(q, \dot{q}, t), \\ dH &= \dot{q}\,dp + p\,d\dot{q} - \left(\frac{\partial L}{\partial q}dq + \frac{\partial L}{\partial \dot{q}}d\dot{q} + \frac{\partial L}{\partial t}dt\right) \\ &= \dot{q}\,dp - \frac{\partial L}{\partial q}\,dq + \left(p - \frac{\partial L}{\partial \dot{q}}\right)d\dot{q} - \frac{\partial L}{\partial t}\,dt, \end{aligned}$$

and if we define the canonical momentum $p \equiv \partial L/\partial \dot{q}$, we can see that the explicit $\dot{q}$ dependence goes away! In practice, to transform from a Hamiltonian in $\dot{q}$ to one in p we simply compute $p = \partial L/\partial \dot{q}$, solve for $\dot{q}$, and substitute. (This generalizes naturally to systems with multiple degrees of freedom.)

Speaking in mathematical terms, what we've done here is take the Legendre transformation of the Lagrangian to get the Hamiltonian: $H(p, q, t) = p\dot{q} - L(q, \dot{q}, t)$. This transform can be inverted in an analogous way to recover the Lagrangian!

Now, by the Euler-Lagrange equation $\dot{p} = \partial L/\partial q$, we're left with

$$\begin{aligned} dH &= \dot{q}\,dp - \frac{\partial L}{\partial q}\,dq - \frac{\partial L}{\partial t}\,dt \\ &= \dot{q}\,dp - \dot{p}\,dq - \frac{\partial L}{\partial t}\,dt. \end{aligned}$$

Also, by the chain rule

$$dH = \frac{\partial H}{\partial p}\,dp + \frac{\partial H}{\partial q}\,dq + \frac{\partial H}{\partial t}\,dt,$$

and matching terms gives Hamilton's equations of motion:

$$\dot{q} = \frac{\partial H}{\partial p}, \qquad \dot{p} = -\frac{\partial H}{\partial q}.$$

For a system with n degrees of freedom we would end with a system of $2n$ equations, two for each coordinate. Also note that we could use the chain rule to show that $dH/dt = \partial H/\partial t$, so H is conserved if it has no explicit time dependence.

5.2 Liouville's Theorem

During our earlier discussion of Lagrangian mechanics we looked at state space as a tool for visualizing dynamical systems. We'll do something similar here, but rather than plot the evolution of the (q, q_i) we will track the (q_i, p_i) in what's called phase space.

Trajectories in phase space have some nice properties. One of them is that different trajectories cannot cross at the same time—and if H is time-independent then they also cannot cross at all. Another is Liouville's theorem, which states that volumes enclosed by surfaces in phase space are time-independent. More specifically, if we look at a region of phase space and imagine it being occupied by a bunch of infinitesimal "particles", then the amount of space those particles occupy is constant over time.

We'll prove this result for a system with one degree of freedom, but we could to larger systems quite easily. Consider a phase space characterized by a velocity vector $\mathbf{v} = \dot{q}\,\hat{q} + \dot{p}\,\hat{p}$ and a closed area Σ in that space, whose outward normal is $\mathbf{n}$. In a small time dt the area enclosed by $\partial\Sigma$ changes by an amount

$$dV = dt \oint_{\partial\Sigma} (\mathbf{v} \cdot \mathbf{n})\, d\ell = dt \iint_{\Sigma} (\nabla \cdot \mathbf{v})\, dA,$$

and so by Hamilton's equations

$$\frac{dV}{dt} = \iint_{\Sigma} \left(\frac{\partial \dot{q}}{\partial q} + \frac{\partial \dot{p}}{\partial p} \right) dA = \iint_{\Sigma} \left[\frac{\partial}{\partial q}\frac{\partial H}{\partial p} + \frac{\partial}{\partial p}\left(-\frac{\partial H}{\partial q} \right) \right] dA = 0.$$

We can also make a more local statement: if ρ denotes the "density" of these particles in phase space then $d\rho/dt = 0$. (This is because we can take an arbitrarily small region of particles and see how the number of particles and size of the region are both constant.) So the flow generated by the Hamiltonian here acts sort of like an incompressible fluid!

5.3 Poisson Brackets and Quantization

The Hamiltonian formulation of mechanics comes with a very elegant mathematical structure for time evolution. If we have some observable that depends on some combination of coordinates q_i, p_i (and, perhaps, time) then we can write

$$\begin{aligned} \frac{df}{dt} &= \frac{\partial f}{\partial t} + \sum_i \left(\frac{\partial f}{\partial q_i}\dot{q}_i + \frac{\partial f}{\partial p_i}\dot{p}_i \right) \\ &= \frac{\partial f}{\partial t} + \sum_i \left(\frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q_i} \right). \end{aligned}$$

We'll call the summation here the Poisson bracket $\{f, H\}$. Notice that f is conserved if and only if $\{f, H\} = 0$, in which case we might say that f "Poisson commutes" with the Hamiltonian. (Because of this role it plays in observables' time dependence, the Hamiltonian is often referred to as the "generator" of time evolution.)

More generally, for any observables A, B we have the Poisson bracket

$$\{A, B\} = \sum_i \left(\frac{\partial A}{\partial q_i}\frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial B}{\partial q_i} \right),$$

and so $\{A, B\} = -\{A, B\}$. Crucially, the position and momentum observables satisfy

$$\{q_i, q_j\} = \{p_i, p_j\} = 0, \qquad \{q_i, p_j\} = \delta_{ij}$$

for all valid indices i, j. We can also see that the Poisson bracket provides a nice way to rewrite Hamilton's equations:

$$\dot{q}_i = \{q_i, H\}, \qquad \dot{p}_i = \{p_i, H\}.$$

Now, one of the great advantages of the Hamiltonian formulation over the Lagrangian one is a greater flexibility in our choice of coordinates. In Lagrangian mechanics we could pick any set of coordinates q_i we wanted, and the $\dot{q}_i$ just fell out of that choice. But in Hamiltonian mechanics the q_i and p_i are much more removed from one another, so in principle we have twice as many choices to make for coordinates! This really increases our ability to simplify complex problems. But still, not every transformation is valid.

In particular, a coordinate transformation

$$(q_1, p_1, \ldots, q_N, p_N) \rightarrow (Q_1 P_1, \ldots, Q_N, P_N)$$

is allowed only if it preserves the structure of Hamilton's equations:

$$\dot{Q}_i = \frac{\partial H}{\partial P_i}, \qquad \dot{P}_i = -\frac{\partial H}{\partial Q_i}.$$

Alternatively, it must satisfy the Poisson "commutation relations"

$$\{Q_i, Q_j\} = \{P_i, P_j\} = 0, \qquad \{Q_i, P_j\} = \delta_{ij}.$$

Such a coordinate transformation is called a canonical transformation, and we call the Q_i, P_i canonical coordinates. The flexibility that comes with these transformations is reflected in how the transformation $Q = -p$, $P = q$ is perfectly valid!

Another surprising example of a valid canonical transformation is time evolution:

$$Q_i(t) = q_i(t + \Delta t), \qquad P_i(t) = p_i(t + \Delta t).$$

There's a couple of interesting lines of reasoning we could follow from this. On one hand, the Hamiltonian framework may encourage us to move all time dependence from the state of the system to the coordinates that describe the system. If we imagine a state as a collection of points on a grid, we might imagine the grid itself warping rather than the points moving around upon it.

But we can also take the idea of positions and momenta as transformations to the extreme and say that *all* observables are just transformations of the system's state. In this case we would need some new quantity on which these observables can act and transform—we might call it a wave function or, more suggestively, a quantum state.

In this way of thinking, all observables become operators rather than functions of time. These operators, of course, don't necessarily commute; to capture this we define the commutator

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} = -[\hat{B}, \hat{A}].$$

The commutation relation for a pair of quantum operators comes from the Poisson bracket:

$$[\hat{A}, \hat{B}] = i\hbar \{A, B\}.$$

This is where the canonical commutation relations come from, including the Heisenberg uncertainty principle:

$$[\hat{X}_i, \hat{P}_i] = i\hbar \{X_i, P_j\} = i\hbar\, \delta_{ij}.$$

Physically, operators that don't commute with one another have some fundamental uncertainty associated with their successive measurements. If we went to measure X_i in a quantum state, for example, we would not then be able to definitively measure P_i for that same state.

Given the relationship between commutators and Poisson brackets, it should be unsurprising that observable time evolution in the quantum picture is very similar to that in the classical one. We'll just slap on some expectation values:

$$\frac{d\left\langle \hat{f} \right\rangle}{dt} = \left\langle \frac{\partial \hat{f}}{\partial t} \right\rangle + \frac{1}{i\hbar} \left\langle [\hat{f}, \hat{H}] \right\rangle.$$

One final connection we can make between classical and quantum mechanics is that, in both pictures, we consider a variety of paths in determining how a system will move from point A to point B. In classical mechanics we look for a path with extremal action, while in quantum mechanics we look for a path with stationary phase. These two ideas are related by the fact that the wave function for a process going from A to B is given by

$$\psi(A \rightarrow B) = \sum_k \exp\left(\frac{iS_k}{\hbar}\right),$$

where k indexes the paths $A \rightarrow B$. From here we can see that the principle of extremal action follows from the principle of stationary phase.