Classical Mechanics

Nate Annau

Table of Contents

1. Physical Preliminaries

We review some fundamental concepts from Newtonian Mechanics.

1.1. Time and Space

Probably the two most fundamental ideas of physics are **time** and **space**. These are things we have an intuitive understanding of by observing the world around us—an exact definition is outside the realm of physics, and probably best left to philosophy. What is important to note, however, is that we can only measure time and space with respect to some reference point. I.e., we must define a **coordinate system**, with some reference point in space, or some reference time, before we can talk about distance or time taken. This is the idea behind a **reference frame**: some set of reference points and coordinate axes such that space and time can be defined precisely.

Define **position** to be a particular point in space within some reference frame. Now, supposing we have a coordinate system giving us a concrete notion of time and space, suppose we have a ball at position $\mathbf{P}$ at time t, and suppose at another instant, we measure that the ball has moved to position $\mathbf{P} + \Delta\mathbf{P}$ at time $t + \Delta t$. We may wonder how fast the ball is going, and in which direction. This is the idea of **velocity**. In particular, we might say that in the time in between our measurements, the ball moved $\Delta\mathbf{P}$ in time Δt, so its rate of change of position per unit time is

$$\frac{\Delta\mathbf{P}}{\Delta t}.$$

Yet, this is not particularly precise—it gives us an approximate idea of how fast the ball was moving at time t, but the ball could have sped up or slowed down within the time interval Δt—we have no way of knowing. To ensure a more accurate sample, we can try shrinking Δt, so there's less chance of fluctuations in the rate of change of position. We finally make this precise when Δt becomes infinitely small; this is how we formally define velocity:

$$\mathbf{v} = \lim_{\Delta t \to 0} \frac{\Delta P}{\Delta t} = \frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t}$$

This general process is called finding the **instantaneous rate of change**, known as a **derivative**. Finally, we define **acceleration** to be the derivative of velocity, i.e., it's the rate of change of the rate of change of position:

$$\mathbf{a} = \lim_{\Delta t \to 0} \frac{\Delta \mathbf{v}}{\Delta t} = \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t}.$$

1.2. Galilean Transformations

1.3. Newton's Laws

The foundation of classical mechanics is comprised of Newton's Three Laws, which we present out of order.

1. Through observation of the world, physicists found that in general, objects do not accelerate, a tendency we call **inertia**. By "in general", we specifically are operating under the assumption that objects are not interacting with other objects outside a given system. Newton's First Law states that there are special reference frames where this assumption holds, called **inertial reference frames (IRFs)**.

   An important point of emphasis: Newton's First Law only tells us that *these reference frames exist*, not that every possible frame of reference is inertial. So, the logical statement

$$\text{no interaction} \implies \text{no acceleration}$$

and the logically equivalent statement

$$\text{acceleration} \implies \text{interaction},$$

only hold true when we are in an inertial reference frame. Further, we say nothing about the converse; if there is an interaction in an inertial frame, there may or may not be acceleration.

3. Another observation physicists made is that in an isolated system consisting of two objects, the objects interact in a particular way: they accelerate towards each other. That is, suppose we are in an inertial reference frame consisting of two isolated objects. Newton's Third Law states that the objects will accelerate, subject to:
   1. Their accelerations will be opposite in direction (toward one another)
   2. The ratio of their accelerations will be constant

   Suppose we have Object 0 and Object 1 given as above. According to our law, the ratio of their accelerations is some constant

   $$\frac{\mathbf{a}_0}{\mathbf{a}_1} = c.$$

   Since the accelerations are opposite in direction, notice c will be negative. Now define the **mass** of Object 0 to be m_0, some arbitrary reference point.

   Then define the mass of Object 1 to be

   $$m_1 \equiv -cm_0 = -\frac{\mathbf{a}_0}{\mathbf{a}_1} m_0$$

   Notice that this definition implies mass is large when the other object is accelerating more quickly than the current object, and it's small when the current object accelerates more quickly. Thus, we can naturally interpret mass as a measure of resistance to acceleration, that is, inertia. Further, for every object whose mass we might want to measure, we can repeat this same process, as long as we can find a proper inertial reference frame. In other words, by doing these experiments with two isolated bodies, we can define a mass for every object we want.

   The next step is to be more precise about the idea of "interactions". Since we observed that in this interaction, Object 0 is accelerating, and we found that mass is some resistance to acceleration, it is natural to define a new concept, called **force**, to be the product of the two.

   $$\mathbf{F} \equiv m\mathbf{a}.$$

   We are also motivated by a desire to write the equation defining mass more simply; by defining $\mathbf{F}_0 = m_0\mathbf{a}_0$ and $\mathbf{F}_1 = m_1\mathbf{a}_1$, the equation becomes

   $$\mathbf{F}_1 = -\mathbf{F}_0.$$

   Notice the logical chain here: we started with empirical observations, wrote down a law, and then some definitions naturally followed to make it easier to describe the phenomena. In particular, it is essential to realize that force and mass are ways to concisely describe what Newton's laws are saying, and are not necessarily more fundamental concepts.

2. The Second Law states that if an object of mass m in an inertial frame experiences some forces $\{\mathbf{F}_i\}_{i\in I}$, where each $\mathbf{F}_i$ represents the force exerted in an isolated system with no other forces, then the acceleration $\mathbf{a}$ of the object satisfies

   $$\mathbf{F}_{\text{tot}} = \sum \mathbf{F}_i = m\mathbf{a};$$

   that is, we have a superposition law: the total force exerted on an object with many interactions is the same as if considered each pair of objects as an isolated system.

The most important thing to remember about these laws is that, while they may appear similar to mathematical axioms, the key difference is that they are based on empirical observation. It is at all not self-evident that inertial reference frames exist, for example; careful observation of real-world phenomena suggests that it is so. Unlike mathematics, we cannot do physics by writing whatever definitions wherever we please. They must come from experiment.

1.4. Energy

Suppose we wish to determine the motion of a generic particle due to the action of particular force. By Newton's Laws, we merely need to solve the differential equation

$$\mathbf{F} = m\mathbf{a}.$$

Path integrating both sides with respect to position gives

$$\begin{aligned}\int \mathbf{F} \cdot \mathrm{d}\mathbf{r} &= m \int \mathbf{a} \cdot \mathrm{d}\mathbf{r} \\ &= m \int \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} \cdot \left(\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\, \mathrm{d}t \right) \\ &= m \int \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} \cdot \left(\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t}\, \mathrm{d}t \right) \\ &= m \int \mathbf{v} \cdot \mathrm{d}\mathbf{v} \\ &= \frac{1}{2} m v^2\end{aligned}$$

We define this new quantity on the right to be **kinetic energy**, and call the quantity on the left **work**. The reason we specify this energy as "kinetic" is that it is fundamentally coming from the motion caused by the object's acceleration, as opposed to another type of energy, called **potential energy**, that comes from work. While we can loosey describe energy as some measure of motion or potential motion of a particle, it is perhaps best to think of energy as some mysterious quantity that proves to be extraordinarily useful in doing physics. Work then, can be thought of as a change in energy resulting from the action of a force over a distance.

Thus, we have just found a key result:

Theorem: **Work-energy theorem**

> The work W done by a force is equal to the change in an object's kinetic energy:
>
> $$W = \Delta K.$$

It initially seems like we are making these definitions for no reason: why do we care about these quantities, when we alreadd(y) understand forces? The reason that energy is useful is that for particular systems, it is **conserved**, i.e., invariant with respect to time. This means that at each point in a physical system, we can calculate the total energy, and make deductions about quantities based on that. This, of course, leaves an important question: why is it conserved? The fundamental reason is temporal symmetry, which will be explained later on in our coverage of **Noether's Theorem**.

1.5. Conservative Forces

It turns out that the key property of forces that make energy calculations easy is that the work done by the force is independent of the path taken. Forces that satisfy this are called **conservative**.

By the Fundamental Theorem of Calculus for Line Integrals, we then have

$$\int_{\mathbf{r}_0}^{\mathbf{r}_1} \mathbf{F} \cdot \mathrm{d}\mathbf{r} = -U(\mathbf{r}_1) + U(\mathbf{r}_0).$$

or equivalently,

$$\mathbf{F} = -\boldsymbol{\nabla}\mathbf{U}.$$

where U is some potential function. Combining this with the work-energy theorem gives

$$K_a + U_a = K_b + U_b,$$

for a system subject to a conservative force, a particularly simple statement of the conservation of energy. Note the reason for the sign convention on U is so that we have sums in the above equation.

1.6. Momentum and Angular Momentum

1.7. Harmonic Oscillator

2. Mathematical Preliminaries

2.1. Coordinate Systems

We very often want to choose a coordinate system that makes a particular problem as easy as possible. While Cartesian coordinates $\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}$ are comfortable and familiar, they are not well-suited to every task. We thus discuss the other most commonly used coordinate systems in physics.

Definition: Polar Coordinates

Consider an arbitrary vector. We want to define unit vectors $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{\theta}}$ in such a way that

$$\mathbf{r} = x\hat{\boldsymbol{x}} + y\hat{\boldsymbol{y}} = r\hat{\boldsymbol{r}}.$$

In particular, from geometry we want

$$\mathbf{r} = \sqrt{x^2 + y^2}\hat{\boldsymbol{r}} = r\cos\theta\hat{\boldsymbol{x}} + r\sin\theta\hat{\boldsymbol{y}}.$$

with $\theta = \operatorname{atan2}\left(\frac{y}{x}\right)$. Thus in Cartesian coordinates, $\hat{\boldsymbol{r}}$ should be

$$\frac{1}{r}\begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$

and being orthonormal in a right-handed coordinate system, we have $\hat{\boldsymbol{\theta}} = (-\sin\theta, \cos\theta)$.

We now wish to determine the Euclidean line element in polar coordinates. Recall the Euclidean metric in Cartesian coordinates:

$$\mathrm{d}t^2 = \mathrm{d}x^2 + \mathrm{d}y^2\,.$$

From the definition of polar coordinates in terms of Cartesian coordinates, we see that

$$\mathrm{d}\mathbf{x} = \cos\theta\mathrm{d}\mathbf{r} - r\sin\theta\mathrm{d}\boldsymbol{\theta}; \qquad \mathrm{d}\mathbf{y} = \sin\theta\mathrm{d}\mathbf{r} + r\cos\theta\mathrm{d}\boldsymbol{\theta}$$

so that our line element becomes

$$\begin{aligned}\mathrm{d}s^2 &= (\cos\theta\mathrm{d}\mathbf{r} - r\sin\theta\mathrm{d}\boldsymbol{\theta})^2 + (\sin\theta\mathrm{d}\mathbf{r} + r\cos\theta\mathrm{d}\boldsymbol{\theta})^2 \\ &= \cos^2\theta\,\mathrm{d}r^2 + r^2\sin^2\theta\,\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}r^2 + r^2\cos^2\mathrm{d}\theta^2 \\ &= \mathrm{d}r^2 + r^2\,\mathrm{d}\theta^2\,.\end{aligned}$$

Finally, we now want to express a generic position vector and its derivatives in polar coordinates. We have

$$\begin{aligned}\mathbf{r} &= (x, y) = r(\cos\theta, \sin\theta) = r\hat{\boldsymbol{r}} \\ \dot{\mathbf{r}} &= (\dot{x}, \dot{y}) = \dot{r}(\cos\theta, \sin\theta) + r\dot{\theta}(-\sin\theta, \cos\theta) = \dot{r}\hat{\boldsymbol{r}} + r\dot{\theta}\hat{\boldsymbol{\theta}} \\ \ddot{\mathbf{r}} &= (\ddot{x}, \ddot{y}) = \ddot{r}(\cos\theta, \sin\theta) + 2\dot{r}\dot{\theta}(-\sin\theta, \cos\theta) + r\ddot{\theta}(-\sin\theta, \cos\theta) - r\dot{\theta}^2(\cos\theta, \sin\theta) \\ &= \left(\ddot{r} - r\dot{\theta}^2\right)\hat{\boldsymbol{r}} + \left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right)\hat{\boldsymbol{\theta}}.\end{aligned}$$

Note this equation for $\ddot{\mathbf{r}}$ will be expanded upon in the later section on non-IRFs.

2.1.1. Cylindrical Coordinates

2.1.2. Spherical Coordinates

2.2. Multivariable Differentiation

Let $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ be a function such that each of its first order partial derivatives exists on $\mathbb{R}^n$. Then the **Jacobian matrix** of $\mathbf{f}$, denoted $\mathbf{J_f}$, is the $m \times n$ matrix whose (i, j) entry is $\frac{\partial f_i}{\partial x_j}$:

$$\mathbf{J_f} = \begin{pmatrix} \frac{\partial \mathbf{f}}{\partial x_1} & \cdots & \frac{\partial \mathbf{f}}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \nabla^T f_1 \\ \vdots \\ \nabla^T f_m \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

Then, the Jacobian is the best linear approximation to $\mathbf{f}$ at $\mathbf{x}$; the linear map $\mathbf{h} \to \mathbf{J}(\mathbf{x}) \cdot \mathbf{h}$ is called the **total derivative** of $\mathbf{f}$ at $\mathbf{x}$.

In our case, we will frequently be dealing with **scalar-valued** multivariable functions, meaning $m = 1$. In particular, it will often be the case that each $v \in \mathbb{R}^n$ is can be written as functions of a single variable, t. That is, we write $f = f(x_1(t), ..., x_n(t))$. Thus, for $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}$, we have

$$\mathbf{J_f} = \nabla^T f = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{pmatrix} \implies \mathrm{d}f = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \, \mathrm{d}x_i$$

is our total derivative, with $\mathbf{h} = (\mathrm{d}x_1, ..., \mathrm{d}x_n)$. Thus we have

$$\boxed{\frac{\mathrm{d}f}{\mathrm{d}t} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{\mathrm{d}x_i}{\mathrm{d}t}.}$$

We also want to deal with the case where we have a composition of functions.

***Theorem:* Multivariable Chain Rule**

In single variable, calculus, recall we had

$$\frac{\mathrm{d}}{\mathrm{d}x}(f(g(x))) = f'(g(x))g'(x).$$

For general differentiable functions $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^k$ and $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$, the total derivative is

$$\mathbf{J_{f \circ g}}(\mathbf{x}) = \mathbf{J_f}(\mathbf{g}(\mathbf{x}))\mathbf{J_g}(\mathbf{x}).$$

Here, we are taking the product of a $k \times m$ matrix with an $m \times n$ matrix, so our result is a $k \times n$ matrix.

Again, the case that will be important for us in the future is the $k = 1$ case. Suppose $f = f(x_1, ..., x_m)$ is a differentiable function of m independent variables, where each $x_i = x_i(t_1, ..., t_n)$ is a differentiable function of n independent variables. Then

$$\mathbf{J_{f \circ g}} = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{pmatrix} = \mathbf{J_f}(\mathbf{g})\mathbf{J_g} = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_m} \end{pmatrix} \begin{pmatrix} \frac{\partial x_1}{\partial t_1} & \cdots & \frac{\partial x_1}{\partial t_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_m}{\partial t_1} & \dots & \frac{\partial x_m}{\partial t_n} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{m} \frac{\partial f}{\partial x_k} \frac{\partial x_k}{\partial t_1} & \cdots & \sum_{k=1}^{m} \frac{\partial f}{\partial x_k} \frac{\partial x_k}{\partial t_n} \end{pmatrix}.$$

In particular, the general result is

$$\boxed{\frac{\partial f}{\partial t_j} = \sum_{k=1}^{m} \frac{\partial f}{\partial x_k} \frac{\partial x_k}{\partial t_j}.}$$

2.3. Calculus of Variations

Define a **functional** F to be a map $F : X \to F$, where X is some function space and the field F is either $\mathbb{R}$ or $\mathbb{C}$. Define an **operator** f to be a map $f : X \to X$, a mapping from a set of functions to itself.

A typical example of a functional is a definite integral

$$F[x] = \int_{t_0}^{t_1} x(t)\,\mathrm{d}t$$

for some $x \in X$. In particular, we are often interested in functionals which act as a metric on our space, such as the path length function in $\mathbb{R}^2$:

$$\ell[y] = \int_{x=x_a}^{x=x_b} \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2} = \int_{x_a}^{x_b} \sqrt{1 + y'(x)^2}\,\mathrm{d}x\,.$$

Lemma: **Fundamental Lemma of the Calculus of Variations**

If a continuous function f on an open interval (a, b) satisfies the equality

$$\int_a^b f(x)h(x)\,\mathrm{d}x = 0$$

for all compactly supported smooth functions h on (a, b), then f is identically zero.

Proof: Suppose $f(\xi) \neq 0$ for some $\xi \in (a, b)$. Since f is nonzero, it is nonzero with the same sign for some c, d with $a < c < \xi < d < b$. Without loss of generality take $f(\xi) > 0$. Then we can find an h that is positive on (c, d) and zero elsewhere, so that the integral is nonzero, a contradiction.

□

2.4. Euler-Lagrange Equation

We want to find extrema of a general functional $J[f]$ with respect to f. We are particularly interested in an integrand of the form $L(x, f(x), f'(x))$, where we assume L is twice continuously differentiable.

Consider that if f is maximal, then any boundary-preserving perturbation to f_0 increases J (if f_0 is a minimizer) or decreases J (if f_0 is a maximizer).

Let our perturbation be given by

$$f_\varepsilon = f + \varepsilon\eta$$

where $\varepsilon > 0$, $f : \mathbb{R} \to \mathbb{R}$ is fixed, and we can vary the compactly supported smooth function $\eta : \mathbb{R} \to \mathbb{R}$. Note that we impose the condition $\eta(a) = \eta(b) = 0$ so that we only consider paths which go between the endpoints. Now define

$$\Phi(\varepsilon) = J[f + \varepsilon\eta] = \int_a^b L(x, f(x) + \varepsilon\eta(x), f'(x) + \varepsilon\eta'(x))\,\mathrm{d}x\,.$$

We want to find a sufficient condition for f being an extremum. Thus, assume f is an extremum of $J[f]$; that is, $\Phi(0)$ is an extremum. Thus $\varepsilon = 0 \implies \frac{\mathrm{d}\Phi}{\mathrm{d}\varepsilon} = 0 \forall \eta$.

Following our characterization of extrema, differentiate with respect to ε:

$$\begin{aligned}\frac{\mathrm{d}\Phi}{\mathrm{d}\varepsilon} &= \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\int_a^b L(x, f(x) + \varepsilon\eta(x), f'(x) + \varepsilon\eta'(x))\,\mathrm{d}x \\ &= \int_a^b \frac{\partial}{\partial\varepsilon} L(x, f(x) + \varepsilon\eta(x), f'(x) + \varepsilon\eta'(x))\,\mathrm{d}x\end{aligned}$$

where in the previous step, we used Leibniz's Integral Rule to interchange limits. To evaluate this partial derivative, recall the final result in the previous section:

$$\begin{aligned}\frac{\partial L}{\partial\varepsilon} &= \frac{\partial L}{\partial x}\frac{\partial x}{\partial\varepsilon} + \frac{\partial L}{\partial(f(x) + \varepsilon\eta(x))}\frac{\partial(f(x) + \varepsilon\eta(x))}{\partial\varepsilon} + \frac{\partial L}{\partial(f'(x) + \varepsilon\eta'(x))}\frac{\partial(f'(x) + \varepsilon\eta'(x))}{\partial\varepsilon} \\ &= \eta(x)\frac{\partial L}{\partial(f(x) + \varepsilon\eta(x))} + \eta'(x)\frac{\partial L}{\partial(f'(x) + \varepsilon\eta'(x))}\end{aligned}$$

so our integral becomes

$$\frac{\mathrm{d}\Phi}{\mathrm{d}\varepsilon} = \int_a^b \left[\eta(x)\frac{\partial L}{\partial(\cdots)}(x, f(x) + \varepsilon\eta(x), f'(x) + \varepsilon\eta'(x)) + \eta'(x)\frac{\partial L}{\partial(\cdots)}(x, f(x) + \varepsilon\eta(x), f'(x) + \varepsilon\eta'(x))\right]\mathrm{d}x\,.$$

By our assumption, an extremum is achieved when ε vanishes. Now we integrate the second term by parts and apply the boundary conditions:

$$\begin{aligned}0 &= \int_a^b \left[\frac{\partial L}{\partial f}(x, f(x), f'(x)) - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial L}{\partial f'}(x, f(x), f'(x))\right]\eta(x)\,\mathrm{d}x + \left[\eta(x)\frac{\partial L}{\partial f'}(x, f(x), f'(x))\right]_a^b \\ &= \int_a^b \left[\frac{\partial L}{\partial f}(x, f(x), f'(x)) - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial L}{\partial f'}(x, f(x), f'(x))\right]\eta(x)\,\mathrm{d}x\end{aligned}$$

which, by the Fundamental Lemma, yields the **Euler-Lagrange Equation**:

$$\boxed{\frac{\partial L}{\partial f}(x, f(x), f'(x)) = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial L}{\partial f'}(x, f(x), f'(x)).}$$

Note for a given extremum, satisfying the Euler-Lagrange Equation is necessary but not sufficient. I.e., every extremum satisfies Euler-Lagrange, but the converse is false.

Remark

Because it will be important later on, we wonder what happens if L depends on additional variables; consider $L(x, f_1(x), f_1'(x), ..., f_n(x), f_n'(x)))$. Now, consider a perturbation in one of the coordinates,

$$J[f_1, ..., f_i + \varepsilon\eta, ..., f_n] = \int_a^b L(x, f_1, ..., f_i + \varepsilon\eta, f_i' + \varepsilon\eta', ..., f_n')\,\mathrm{d}x\,.$$

Following the logic, we see the same equation holds for the new L—but crucially, this is predicated upon the assumption that it is possible to vary f_i independently in the first place. If there were a relationship between the f_i's, we would get a more complicated equation.

2.5. Geodesics

Armed with the Euler-Lagrange equation, we now wish to tackle general shortest path problems.

Definition: Geodesic

A **geodesic** is a curve representing the locally shortest path between two points on a Riemannian manifold.

Formally, a curve $\gamma : I \to M$ from an interval $I \subseteq \mathbb{R}$ to a metric space M is a geodesic if $\exists v \geq 0$ such that $\forall t \in I \exists$ a neighborhood J of t in I such that $\forall t_1, t_2 \in J$ we have

$$\frac{d(\gamma(t_1), \gamma(t_2))}{|t_1 - t_2|} = v.$$

We interpret this as follows: For every t, if we zoom in far enough, the path γ becomes a straight line up to some constant factor. In fact, if we zoom in so that Δt becomes infinitesimal, we see what is effectively a derivative, albeit with respect to two different metrics in the general case. Thus, the curve has constant velocity, which should make sense, as a curve that speeds up or slows down should not be optimal.

In general, geodesics can be found by using Euler-Lagrange and identifying local minima.

Example

Suppose we want to find the shortest path between two points on the surface of a sphere.

Take a polar angle θ and azimuthal angle φ as the coordinates, with constant radius R. The infinitesimal distances are $\mathrm{d}s_\theta = R\,\mathrm{d}\theta$ and $\mathrm{d}s_\varphi = R\sin\theta\,\mathrm{d}\varphi$, implying

$$\mathrm{d}s^2 = R^2\,\mathrm{d}\theta^2 + R^2\sin^2\theta\,\mathrm{d}\varphi^2\,.$$

It turns out to be easier to use θ as a variable of integration, since the first term in the Euler-Lagrange equation vanishes. In particular we write

$$s = R\int_a^b \sqrt{1 + \sin^2\theta\varphi'^2}\,\mathrm{d}\theta$$

with corresponding Euler-Lagrange equation

$$\frac{\partial F}{\partial \varphi} - \frac{\mathrm{d}}{\mathrm{d}\theta}\frac{\partial F}{\partial \varphi'} = 0 \text{ with } F = \sqrt{1 + \sin^2\theta\varphi'^2}.$$

We could solve this for a particular

2.6. Fourier Transform

2.7. Legendre Transform

3. Lagrangian Mechanics

3.1. The Lagrangian

We want to now use the Euler-Lagrange equation to solve mechanics problems. To do this, we want to find some function $L(t, x, \dot{x})$ such that evaluating the Euler-Lagrange equation gives Newton's Second Law for a particle subject to conservative forces,

$$\frac{\mathrm{d}}{\mathrm{d}t}(m\dot{x}) = -\frac{\mathrm{d}U}{\mathrm{d}x}.$$

Comparing with the form of Euler-Lagrange, we see we must have

$$\frac{\partial L}{\partial \dot{x}} = m\dot{x}; \quad \frac{\partial L}{\partial x} = -\frac{\mathrm{d}U}{\mathrm{d}x}.$$

Solving the first equation by separation of variables gives

$$L = \frac{1}{2}m\dot{x}^2 + g(t, x).$$

Now since U is purely a function of x, our second PDE means we do not have to consider t dependence in our solution, meaning without loss of generality let g be a function of x alone. From here we deduce

$$\frac{\partial g}{\partial x} = -\frac{\mathrm{d}U}{\mathrm{d}x} \Rightarrow g(x) = -U(x),$$

implying that

$$L = T - U$$

works. This is called the **Lagrangian** of our system, and it gives us a powerful new formulation of mechanics. Importantly, because we did not consider our PDE solution in full generality, it is not unique in its implication of Newton's Second Law.

Definition: Least Action Principle

Given a mechanical system described by N dynamical generalized coordinates $q_k(t)$, with $k = 1, 2, ..., N$, define its **action** by

$$S[q_k(t)] = \int_{t_a}^{t_b} \mathrm{d}t\, L(t, q_1, q_2, ..., \dot{q}_1, \dot{q}_2, ...).$$

(We assume the particle begins at some position $(q_1, q_2, ...)_a$ at time t_a and ends at position $(q_1, q_2, ...)_b$ at time t_b.) Also note that this is a generalized version of the function $J[f]$ shown previously.

Now the **least action principle** states that, for trajectories $q_k(t)$ where S is stationary, i.e.,

$$\delta S = \delta \int_{t_a}^{t_b} L(t, q_k \dot{q}_k)\, \mathrm{d}t = 0,$$

then the $q_k(t)$'s satisfy the equations of motions for the system with the given boundary conditions. Note this is just the proof we did earlier a—if to first order $\delta S = 0$, then S is an extrema, and thus the Euler-Lagrange equations apply.

3.2. Cyclic Coordinates

When deriving the form of L, recall we assigned

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

which is the particle's momentum. This inspires us to define

$$p_k \equiv \frac{\partial L}{\partial \dot{q}_k}$$

to be the **generalized momentum** of the particle.

Now, consider the situation in which q_k is not present in the Lagrangian. Then q_k is called a **cyclic coordinate**. This is important because we observe that in the Euler-Lagrange equation,

$$0 = \frac{\partial L}{\partial q_k} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_k}$$

that if

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

we must have

$$\frac{\mathrm{d}p_k}{\mathrm{d}t} = \frac{\partial L}{\partial \dot{q}_k} = 0.$$

so p_k is constant through time, i.e., it's a **conserved quantity**.

3.3. The Hamiltonian

Take the total derivative of the Lagrangian with respect to t:

$$\frac{\mathrm{d}L(q_k, \dot{q}_k, t)}{\mathrm{d}t} = \frac{\partial L}{\partial t} + \frac{\partial L}{\partial q_k}\dot{q}_k + \frac{\partial L}{\partial \dot{q}_k}\ddot{q}_k$$

Note we are using Eisenstein summation convention; i.e., a sum over k is implied for each term. Notice it shares two terms in common with

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{q}_k\frac{\partial L}{\partial \dot{q}_k}\right) = \ddot{q}_k\frac{\partial L}{\partial \dot{q}_k} + \dot{q}_k\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_k}\right) = \ddot{q}_k\frac{\partial L}{\partial \dot{q}_k} + \dot{q}_k\frac{\partial L}{\partial q_k}.$$

The difference of the equations is

$$\begin{aligned}\frac{\mathrm{d}L(q_k, \dot{q}_k, t)}{\mathrm{d}t} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{q}_k\frac{\partial L}{\partial \dot{q}_k}\right) &= \frac{\partial L}{\partial t} \\ \implies \frac{\partial L}{\partial t} - \frac{\mathrm{d}}{\mathrm{d}t}\left(L - \dot{q}_k\frac{\partial L}{\partial \dot{q}_k}\right) &= 0.\end{aligned}$$

This inspires us to define the **Hamiltonian** H of the particle to be

$$\boxed{H \equiv \dot{q}_k p_k - L}$$

so that we can rewrite our previous equation as

$$\frac{\partial L}{\partial t} = -\frac{\mathrm{d}H}{\mathrm{d}t}.$$

Importantly, if L is not a function of time, then the Hamiltonian is conserved.

What does the quantity actually mean? Observe

$$H = \sum_k \dot{q}_k \frac{\partial (T-U)}{\partial \dot{q}_k} - (T-U) = \sum_k \dot{q}_k \frac{\partial T}{\partial \dot{q}_k} - T + U.$$

Now, let $\mathbf{r}(q_k, t)$ be a position vector of a particle in an inertial frame. We can say that

$$\mathbf{v} = \frac{\mathrm{d}\mathbf{r}(q_1, \dots q_n, t)}{\mathrm{d}t} = \frac{\partial \mathbf{r}}{\partial t} + \sum_k \frac{\partial \mathbf{r}}{\partial q_k}\dot{q}_k$$

so

$$T = \frac{1}{2}mv^2 = \frac{1}{2}m\mathbf{v}\cdot\mathbf{v} = \frac{1}{2}m\left[\frac{\partial \mathbf{r}}{\partial t}\cdot\frac{\partial \mathbf{r}}{\partial t} + 2\frac{\partial \mathbf{r}}{\partial t}\cdot\sum_k \frac{\partial \mathbf{r}}{\partial q_k}\dot{q}_k + \sum_k \frac{\partial \mathbf{r}}{\partial q_k}\dot{q}_k \sum_l \frac{\partial \mathbf{r}}{\partial q_l}\dot{q}_l\right]$$

which we rewrite as

$$2T - m\left[\frac{\partial \mathbf{r}}{\partial t}\cdot\frac{\partial \mathbf{r}}{\partial t} - \frac{\partial \mathbf{r}}{\partial t}\sum_k \frac{\partial \mathbf{r}}{\partial q_k}\dot{q}_k\right] = m\left[\frac{\partial \mathbf{r}}{\partial t}\cdot\sum_k \frac{\partial \mathbf{r}}{\partial q_k}\dot{q}_k + \sum_k \frac{\partial \mathbf{r}}{\partial q_k}\dot{q}_k \sum_l \frac{\partial \mathbf{r}}{\partial q_l}\dot{q}_l\right].$$

Now, indexing over the terms in the original expression by i, we find

$$\begin{aligned}
\sum_i \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} &= \sum_i \dot{q}_i \frac{1}{2}m\left[2\frac{\partial \mathbf{r}}{\partial t}\cdot\frac{\partial \mathbf{r}}{\partial q_i} + 2\sum_k \frac{\partial \mathbf{r}}{\partial q_i}\cdot\frac{\partial \mathbf{r}}{\partial q_k}\dot{q}_k\right] \\
&= m\left[\frac{\partial \mathbf{r}}{\partial t}\cdot\sum_i \frac{\partial \mathbf{r}}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial \mathbf{r}}{\partial q_i}\dot{q}_i \sum_k \cdot\frac{\partial \mathbf{r}}{\partial q_k}\dot{q}_k\right] \\
&= 2T - m\frac{\partial \mathbf{r}}{\partial t}\cdot\frac{\partial \mathbf{r}}{\partial t} - m\frac{\partial \mathbf{r}}{\partial t}\sum_k \frac{\partial \mathbf{r}}{\partial q_k}\dot{q}_k \\
&= 2T - m\frac{\partial \mathbf{r}}{\partial t}\cdot\left[\frac{\partial \mathbf{r}}{\partial t} + \sum_k \frac{\partial \mathbf{r}}{\partial q_k}\dot{q}_k\right] \\
&= 2T - m\mathbf{v}\cdot\frac{\partial \mathbf{r}}{\partial t}.
\end{aligned}$$

Thus overall,

$$\boxed{H = T + U - m\mathbf{v}\cdot\frac{\partial \mathbf{r}}{\partial t}.}$$

So, we see that the Hamiltonian is really the total energy, minus the dot product of momentum and some change in position with respect to time. This term becomes zero when, for example, there are fixed constraints on the problem.

4. Symmetry and Noether's Theorem

4.1. Constraint Forces

Consider a mechanical system parameterized by N coordinates q_k, with P algebraic relations between the coordinates (due to constraint forces), given by

$$C_l(q_1, q_2, ..., q_N, t) = 0,$$

with $l \in \{1, ..., P\}$. These are called **holonomic constraints**. Notice we effectively have $N - P$ degrees of freedom, rather than N.

Example

For example, suppose we have a block on a table. With the height at the table defined to be $z = 0$, we introduce the constraint that the block always lies on the table, i.e. $C(z) = z = 0$. Thus the Lagrangian becomes

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

leaving us with two unknowns.

However, immediately doing this is not necessarily great, for a few reasons:

1. It may be difficult to eliminate the coordinate from the Lagrangian if the relations are more complex;
2. We may be interested in finding the constraint force at play rather than discarding it. I.e., what if we wanted to know the normal force magnitude for the block?

To get around this, we introduce a way to delay implementing the constraints. Recall from the section on the Calculus of Variations that the Euler-Lagrange equation only holds if we are able to vary the equations independently, meaning that we can not directly plug the new equations into the Lagrangian.

Instead, we consider a new Lagrangian defined by

$$L' = L + \sum_{l=1}^{P} \lambda_l C_l$$

where we have introduced P degrees of freedom labeled λ_l with $l = 1, ..., P$. These are called **Lagrange multipliers**. In the previous example, our Langrangian would become

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - mgz + \lambda_1 z$$

Now assuming the constraint equations are not satisfied *a priori*, we have $N + P$ degrees of freedom. The equations of motion are thus:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L'}{\partial \dot{\lambda}_l}\right) - \frac{\partial L'}{\partial \lambda_l} = -C_l = 0 \implies C_l = 0$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L'}{\partial \dot{q}_k}\right) - \frac{\partial L'}{\partial q_k} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} - \sum_{l=1}^{p} \lambda_l \frac{\partial C_l}{\partial q_k} = 0.$$

This last term can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = \sum_{l=1}^{p} \lambda_l \frac{\partial C_l}{\partial q_k} \equiv \mathcal{F}_k$$

where we have defined the $\mathcal{F}_k$'s as the **generalized constraint forces** for our system. It turns out that this definition of L' gives exactly the solutions we desire, but the details are out of scope.

4.2. Symmetry

Define a **symmetry** be to be transformation that leaves the action unchanged in time. Start with the typical action

$$S = \int \mathrm{d}t\, L(q, \dot{q}, t).$$

Now apply some transformation given by $\Delta q_k(t, q)$ and $\delta t(t, q)$. This is very similar to our derivation of the Euler-Lagrange equation, but slightly more general, since we consider variations in time as well as variations in q_k.

Being slightly less formal this time, the variation in the action is

$$\delta S = \delta\left(\int L\,\mathrm{d}t\right) = \int \delta(L\,\mathrm{d}t) = \int \mathrm{d}t\,\delta(L) + \int \delta(\mathrm{d}t) L$$

Now we have that the variation in L due to the step Δq_k is

$$\delta_{\Delta q_k} L = \frac{\partial L}{\partial t} \cdot 0 + \frac{\partial L}{\partial q}\Delta q_k + \frac{\partial L}{\partial \dot{q}}\Delta \dot{q}_k = \frac{\partial L}{\partial q}\Delta q_k + \frac{\partial L}{\partial \dot{q}}\frac{\mathrm{d}}{\mathrm{d}t}\Delta q_k$$

while the variation in L due to the step δt is

$$\delta_{\delta t} L = \frac{\mathrm{d}L}{\mathrm{d}t}\delta t.$$

Then, the variation in $\mathrm{d}t$ due to the time step δt is

$$\delta(\mathrm{d}t) = \mathrm{d}t\,\frac{\delta(\mathrm{d}t)}{\mathrm{d}t} = \mathrm{d}t\,\frac{\mathrm{d}}{\mathrm{d}t}(\delta t)$$

Thus overall we have

$$\begin{aligned}\delta S &= \int \mathrm{d}t \left[\sum_k \left(\frac{\partial L}{\partial q_k}\Delta q_k + \frac{\partial L}{\partial \dot{q}_k}\frac{\mathrm{d}}{\mathrm{d}t}\Delta q_k\right) + \delta t \frac{\mathrm{d}L}{\mathrm{d}t} + L\frac{\mathrm{d}}{\mathrm{d}t}(\delta t)\right] \\ &= \int \mathrm{d}t \left[\sum_k \left(\frac{\partial L}{\partial q_k}\Delta q_k + \frac{\partial L}{\partial \dot{q}_k}\frac{\mathrm{d}}{\mathrm{d}t}\Delta q_k\right) + \frac{\mathrm{d}}{\mathrm{d}t}(L\delta t)\right]\end{aligned}$$

as a general variation in the action due to the transformations $\Delta q_k(t, q)$ and $\delta t(t, q)$.

4.3. Noether's Theorem

This leads to one of the most important results in mathematical physics.

Theorem: **Noether's Theorem**

For every symmetry, there exists a quantity that is conserved under time evolution.

Proof: Suppose we have a given symmetry $\{\delta t(t,q), \Delta q_k(t,q)\}$. Then we get

$$\delta S = \int \mathrm{d}t \left[\sum_k \left(\frac{\partial L}{\partial q_k} \Delta q_k + \frac{\partial L}{\partial \dot{q}_k} \frac{\mathrm{d}}{\mathrm{d}t} \Delta q_k \right) + \frac{\mathrm{d}}{\mathrm{d}t}(L\delta t) \right] = 0.$$

Now pick the curves q_k that satisfy the Euler-Lagrange equation:

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}_k} = \frac{\partial L}{\partial q_k}.$$

Thus

$$\begin{aligned}
\delta S &= \int \mathrm{d}t \left[\sum_k \left(\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}_k} \right) \Delta q_k + \frac{\partial L}{\partial \dot{q}_k} \frac{\mathrm{d}}{\mathrm{d}t} \Delta q_k \right) + \frac{\mathrm{d}}{\mathrm{d}t}(L\delta t) \right] = 0 \\
&= \int \mathrm{d}t \left[\sum_k \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}_k} \Delta q_k \right) + \frac{\mathrm{d}}{\mathrm{d}t}(L\delta t) \right] = 0 \\
&= \int \mathrm{d}t \, \frac{\mathrm{d}}{\mathrm{d}t} \left[\sum_k \left(\frac{\partial L}{\partial \dot{q}_k} \Delta q_k \right) + L\delta t \right] = 0
\end{aligned}$$

Now since the integration interval is arbitrary, we must have

$$\frac{\mathrm{d}}{\mathrm{d}t} Q = 0,$$

where

$$\boxed{Q \equiv \frac{\partial L}{\partial \dot{q}_k} \Delta q_k + L\delta t.}$$

We call this quantity Q the **Noether charge**.

□

We also have the following partial converse:

Corollary: **Noether's Theorem Partial Converse**

Suppose

$$\delta S = \int \mathrm{d}t \, \frac{\mathrm{d}K}{\mathrm{d}t} \text{ for some } K.$$

Then $\delta t(t,q), \delta q_k(t,q)$ is a symmetry, and $Q - K$ is a conserved quantity.

Proof:

□

5. Central Forces

5.1. Two Body Problem

A **central force** on a particle is directed away or toward a fixed point and is spherically symmetric about that point. Thus, the force and potential have no dependence upon θ or φ: we can write $\mathbf{F} = \mathbf{F}(r)$ and $U = U(r)$. A typical example is gravitational attraction

$$\mathbf{F} = -G\frac{m_1 m_2}{r^2}\hat{\boldsymbol{r}}.$$

Now consider a general two body central force problem. Define the coordinates $\mathbf{r}_1 = (x_1, y_1, z_1)$ and $\mathbf{r}_2 = (x_2, y_2, z_2)$ with corresponding **center of mass coordinates**

$$\mathbf{R}_{\text{cm}} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2}.$$

Also define the relative coordinates

$$\mathbf{r} \equiv \mathbf{r}_2 - \mathbf{r}_1.$$

Then we have

$$\mathbf{r}_1 = \mathbf{R}_{\text{cm}} - \frac{m_2}{M}\mathbf{r} \quad \text{and} \quad \mathbf{r}_2 = \mathbf{R}_{\text{cm}} + \frac{m_1}{M}\mathbf{r},$$

with $M = m_1 + m_2$. The total kinetic energy is

$$\begin{aligned} T &= \frac{1}{2}m_1\dot{\mathbf{r}}_1^2 + \frac{1}{2}m_2\dot{\mathbf{r}_2}^2 \\ &= \frac{1}{2}m_1\left(\dot{\mathbf{R}}_{\text{cm}} - \frac{m_2}{M}\dot{\mathbf{r}}\right)^2 + \frac{1}{2}m_2\left(\dot{\mathbf{R}}_{\text{cm}} + \frac{m_1}{M}\dot{\mathbf{r}}\right)^2 \\ &= \frac{1}{2}m_1\left(\dot{\mathbf{R}}_{\text{cm}}^2 - \frac{2m_2}{M}\dot{\mathbf{r}} + \frac{m_2^2}{M^2}\dot{\mathbf{r}}^2\right) + \frac{1}{2}m_2\left(\dot{\mathbf{R}}_{\text{cm}}^2 + \frac{2m_1}{M}\dot{\mathbf{r}} + \frac{m_1^2}{M^2}\dot{\mathbf{r}}^2\right) \\ &= \frac{1}{2}M\dot{\mathbf{R}}_{\text{cm}}^2 + \frac{1}{2}\mu\dot{\mathbf{r}}^2 \end{aligned}$$

where

$$\mu = \frac{m_1 m_2}{M}$$

is called the **reduced mass** of the system.

6. Non-Inertial Reference Frames

7. Coupled Oscillators

7.1. Introductory example

Consider a system of two masses with mass m, where we use x_1 to measure the position of the first mass from the left and x_2 to measure the position of the second mass from the left. The first mass is attached to the left wall by a spring with constant k, the second mass is attached to the right wall by a spring with constant k, and they are attached to each other with a third spring of constant k'. Suppose the distance between the walls is L, so that our Lagrangian becomes

$$\mathcal{L} = \frac{1}{2}m\dot{x_1}^2 + \frac{1}{2}m\dot{x_2}^2 - \frac{1}{2}k(x_1 - a_1)^2 - \frac{1}{2}(L - x_2 - a_2)^2 - \frac{1}{2}k'(x_2 - x_1 - a_{12})^2$$

where a_1, a_2, a_{12} are the rest lengths. Suppose the masses are at equilibrium. Then we can write

$$k\Delta\ell_1 + k'\Delta\ell_{12} = 0$$

as the horizontal force on mass 1. If we move mass 1 some $\Delta x_1 = q_1$ from equilibrium, we get

$$k(\Delta\ell_1 + q_1) + k'(\Delta\ell_{12} + q_1) = F_1 = kq_1 + k'q_1$$

so we can treat the system more simply as offsets from equilibrium. (Note we could have also just done this from the beginning, and chosen our coordinates so that displacements were relative to equilibrium points, but this illustrates why doing this works.) With this new coordinate system in place, our Lagrangian becomes

$$\mathcal{L} = \frac{1}{2}m\dot{q}_1^2 + \frac{1}{2}m\dot{q}_2^2 - \frac{1}{2}kq_1^2 - \frac{1}{2}kq_2^2 - \frac{1}{2}k'(q_2 - q_1)^2$$

where we noted that $\dot{x}_1 = \dot{q}_1$ since x and q differ by a constant offset. Thus the Euler-Lagrange equation gives

$$\begin{aligned} 0 &= -kq_1 + k'(q_2 - q_1) - m\ddot{q}_2 \\ 0 &= -kq_2 - k'(q_2 - q_1) - m\ddot{q}_1 \end{aligned}$$

This implies

$$0 = -k(q_1 + q_2) - m(\ddot{q}_1 + \ddot{q}_2).$$

If we define $\xi_1 \equiv q_1 + q_2$, we get

$$0 = -k\xi_1 - m\ddot{\xi}_1$$

which is finally tractable. Solving this ODE clearly gives

$$\xi_1 = c_1 e^{i\sqrt{\frac{k}{m}}t} + c_1^* e^{-i\sqrt{\frac{k}{m}}t} = A_1 \cos\left(\sqrt{\frac{k}{m}}t + \delta_1\right).$$

Now if we subtracted the equations instead of adding them, we should get

$$\begin{aligned} 0 &= -k(q_2 - q_1) - 2k'(q_2 - q_1) - m(\ddot{q}_2 - \ddot{q}_1) \\ &= -k\xi_2 - 2k'\xi_2 - m\ddot{\xi}_2 \end{aligned}$$

where we assigned $\xi_2 \equiv q_2 - q_1$. Then solving the ODE gives

$$\xi_2 = c_2 e^{i\sqrt{\frac{k+2k'}{m}}t} + c_2^* e^{-i\sqrt{\frac{k+2k'}{m}}t} = A_2 \cos\left(\sqrt{\frac{k+2k'}{m}}t + \delta_2\right).$$

Observe that we can now recover q_1 and q_2:

$$q_1 = \frac{A_1}{2}\cos(\omega_1 t + \delta_1) - \frac{A_2}{2}\cos(\omega_2 t + \delta_2)$$
$$q_2 = \frac{A_2}{2}\cos(\omega_1 t + \delta_1) + \frac{A_2}{2}\cos(\omega_2 t + \delta_2).$$

Note it turns out that ξ_1 and ξ_2, considering the system as an aggregate, are often easier to work with and interpret. I.e., we can think of ξ_1 as a center of mass coordinate and ξ_2 as a relative motion coordinate.

7.2. Generic Lagrangian as an Oscillator

Suppose we have a Lagrangian

$$L = T - U = \sum_{i,j} f_{ij}(q_1, ..., q_N)\dot{q}_i\dot{q}_j - U_{\text{eff}}(q_1, ..., q_N)$$

We can Taylor expand around equilibrium ($\nabla U = 0$) as follows:

$$L \approx f_{ij}\big|_{\text{eq.}}\dot{q}_i\dot{q}_j + \left.\frac{\partial f_{ij}}{\partial q_k}\right|_{\text{eq.}}(q_k - q_{\text{k, eq}})\dot{q}_i\dot{q}_j + \cdots - U_{\text{eff}}\big|_{\text{eq.}} - \frac{1}{2}\left.\frac{\partial^2 U_{\text{eff}}}{\partial q_i \partial q_j}\right|_{\text{eq.}}(q_i - q_{i,\text{ eq}})(q_j - q_{j,\text{ eq}}) + \cdots$$
$$= \frac{1}{2}M_{ij}\dot{q}_i\dot{q}_j - \frac{1}{2}K_{ij}q_i q_j \qquad \text{with } M_{ij} = 2f_{ij}\big|_{\text{eq}} \text{ and } K_{ij} = \frac{\partial^2 U_{\text{eff}}}{\partial q_i \partial q_j}$$

so we can treat this as a generic coupled oscillator. We call M_{ij} the "mass matrix" and K_{ij} the "spring constant matrix". With matrix notation, we write this

$$L = \frac{1}{2}\dot{\mathbf{q}}^T \cdot \widehat{M}\dot{\mathbf{q}} - \frac{1}{2}\dot{\mathbf{q}}^T \hat{K}\dot{\mathbf{q}}$$

with $\dot{\mathbf{q}}^T = (q_1, q_2, ..., q_N)$. In our previous example, we had

$$\widehat{M} = \begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix}, \qquad \hat{K} = \begin{pmatrix} k+k' & -k' \\ -k' & k+k' \end{pmatrix}.$$

7.3. Continuum Limit

Consider a system of n masses connected by springs, similar to the example in the previous section except arbitrarily large. It seems like an interesting math problem to figure out what happens; physically, this is also interesting because we can model a continuous material as a large number of tiny masses connected by springs. For example, a solid object is made up of molecules held together by some forces we can model as springs. Explicitly, suppose we take a mass and continuously split it into two half masses connected by a spring, and repeat recursively. This is the idea of the **continuum limit**.

Let's consider the Lagrangian of the system. Indexing the masses by $\mathbb{Z}$, it should be

$$L = \sum_i \left(\frac{1}{2}m\dot{x}_i^2 - \frac{1}{2}k(x_{i+1} - x_i)^2\right)$$

Now, make the replacement

$$\{x_i(t)\} \to \eta(t,x); \qquad x_n(t) \to \eta(t, x \equiv na),$$

so we rewriting our many functions in terms of a single multivariable function. Suppose $x_{i=0} = 0$ at equilibrium, and $x_n = an$ at equilibrium. Thus the Lagrangian becomes

$$L = \sum_{\text{rod}} \left(\frac{1}{2} m \dot{\eta}^2(t,x)^2 - \frac{1}{2} k (\eta(t, x+a) - \eta(t,x))^2 \right).$$

Now Taylor expanding η, we have

$$\eta(t, x+a) \approx \eta(t,x) + \frac{\partial \eta}{\partial x}(t,x) a + \cdots$$

so

$$\begin{aligned} L &= \sum_{\text{rod}} \left(\frac{1}{2} m \dot{\eta}^2 - \frac{1}{2} k (\eta'(a))^2 \right) \\ &= \int \mathrm{d}x \left(\frac{1}{2} \frac{m}{a} \dot{\eta}^2 - \frac{1}{2} k a \eta'^2 \right) \end{aligned}$$

where $a = \mathrm{d}x$ and we define $\eta' \equiv \frac{\partial \eta}{\partial x}$. Now define the mass density $\mu = \frac{m}{a}$ and the Young's Modulus $Y = ka$ (which is the like the stiffness, making it so we have a fixed spring constant as we divide into smaller pieces.)

Thus the action is

$$S = \iint \mathrm{d}t \, \mathrm{d}x \left(\frac{1}{2} \mu \dot{\eta}^2 - \frac{1}{2} Y \eta'^2 \right)$$

We call the integrand here the **Lagrangian density**. The Euler-Lagrange equation for a multivariable function is

$$0 = \frac{\partial \mathcal{L}}{\partial \eta} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \dot{\eta}} - \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial \mathcal{L}}{\partial \eta'}.$$

In our case we have

$$0 = 0 - \mu \ddot{\eta} + Y \eta''$$

which becomes the **wave equation**

$$\boxed{\frac{\partial^2 \eta}{\partial x^2} - \frac{1}{v^2} \frac{\partial^2 \eta}{\partial t^2} = 0}$$

where $v = \sqrt{\frac{Y}{m}}$. Note this has the general solution

$$\eta(t,x) = f(x - vt) + g(x + vt).$$

Using a finite difference method, we can approximate

$$\frac{\partial^2 t}{\partial x_0^2} \approx \frac{\eta(t, x_0 + \varepsilon) + \eta(t, x_0 - \varepsilon) - 2\eta(t, x_0)}{\varepsilon^2}$$

using a second order central difference. Now define

$$\xi_i = P_{ii'} q_{i'}$$

with some transformation

$$\tilde{\eta}(k) = \int \mathrm{d}x\, f(x,k)\eta(x),$$

where we are looking for a transformation that will simplify our system. Thus we have

$$0 = \int \mathrm{d}x\, \frac{\partial^2 f}{\partial x^2}\eta - \frac{1}{v^2}\ddot{\tilde{\eta}}.$$

This means we need

$$\frac{\partial^2 f}{\partial x^2} = A(k) f(k,x)$$

The standard example that satisfies this is to choose

$$f(x,k) = \frac{1}{\sqrt{2\pi}} e^{-ikx} \qquad \text{(Fourier transform)}$$

Thus our PDE becomes

$$0 = -k^2\tilde{\eta} - \frac{1}{v^2}\ddot{\tilde{\eta}},$$

which we note is exactly the harmonic oscillator. Thus

$$\tilde{\eta} = Ce^{-ikvt+i\varphi}.$$

Note this means that $\tilde{\eta}$'s are exactly the normal modes. Further, the inverse Fourier transform is

$$\eta = \int \mathrm{d}k\, \frac{i}{\sqrt{2\pi}} e^{ikx}\tilde{\eta}.$$

8. Scattering

Consider a setup where we have some target, fire some particles at it, and we observe how the particles react. Then, we can learn some information about the target. We call this general experiment **scattering**. For example, vision is a scattering experiment, or even tossing a ball into a dark crevice to determine if it's a hole.

Suppose we have some incident particles N_{inc} with random initial conditions being fired at a target. Suppose the target is a sheet covered in holes, with some number density of targets n_{tar}. Our goal is to know how many particles were scattered N_{sc}, versus the number that made it through the hole. Note this is a similar setup to the Rutherford gold foil experiment, for example. The reason we use a sheet instead of a single hole is that practically speaking, we cannot isolate to a single hole or object we're interested in, so our best bet is to use a sheet and control for this.

Define the **cross-section** by

$$\sigma = \frac{1}{n_{\text{tar}}} \frac{N_{\text{sc}}}{N_{\text{inc}}}$$

For a more specific experiment, suppose we have incident particles modeled by spheres with radius r_1, and our sheet is modeled by a sequence of vertical spheres with radius r_2. Thus if we consider the particles as point particles, the effective cross sectional area covered by the sheet is

$$n_{\text{tar}} \pi (r_1 + r_2)^2$$

Then let A is the total area of the beam.

Now suppose σ is non-constant: we might have some $\sigma_{\text{tot}} = \sum_i \sigma_i$ in that case. Or, infinitesimally,

$$\sigma = \int \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \, \mathrm{d}\Omega = \int_0^{2\pi} \int_{-1}^{1} \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \mathrm{d}(\cos\theta) \, \mathrm{d}\varphi$$

where $\mathrm{d}\Omega$ is the differential solid angle (a measurement of the field of view a given angle covers on a sphere). I.e., take the cone generated by some square angle $\mathrm{d}\Omega$: the area of the spherical cap $\mathrm{d}A$ relates to $\mathrm{d}\Omega$ by $\mathrm{d}\Omega = \frac{\mathrm{d}A}{r^2}$. Solid angles are measured in steradians, equal to one square radian.

Now consider a ring with radius b and thickness $\mathrm{d}b$. Thus we have

$$\mathrm{d}\sigma_{\text{inc}} = 2\pi b \, \mathrm{d}b$$

and

$$\mathrm{d}\Omega = 2\pi \sin\theta \, \mathrm{d}\theta$$

since we have symmetry in φ. Now recall that from our work in central forces,

$$\varepsilon = \sqrt{1 + \frac{2E\ell^2}{k^2 \mu}}$$

with

$$r = \frac{r_0}{1 + \varepsilon \cos\varphi}$$

we can use the fact that $\ell = \mu v_0 b$ and

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{b}{\sin\theta} \frac{\mathrm{d}b}{\mathrm{d}\theta}.$$

Then we eventually get

$$b = \frac{k}{2E} \cot \frac{\theta}{2}$$

9. Hamiltonian Mechanics

9.1. Introduction

Consider the Euler-Lagrange equation

$$\frac{\partial L}{\partial q_k} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_k}\right).$$

Recall we called $\frac{\partial L}{\partial \dot{q}_k}$ the generalized momentum, and $\frac{\partial L}{\partial q_k}$ the generalized force, giving us a direct analogue of Newton's Second Law. Now when we work with Cartesian coordinates, we can simply say $\mathbf{p} = m\mathbf{v}$, so working with momentum is really the same as working with velocities. Thus, we could formulate mechanics in terms of forces causing accelerations; from the chapter on non inertial frames, however, we know that when we don't use Cartesian coordinates, or have a non-inertial frames, then we get terms that aren't necessarily accelerations.

Thus, it is better to think fundamentally as mechanics as forces causing changes in momentum. This is the core idea of Hamiltonian mechanics–rather than consider a spring system Lagrangian

$$L(t, q_k, \dot{q}_k) = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}kq^2$$

where velocity is one of our coordinates, we instead assign $p = m\dot{q} \Leftrightarrow \dot{q} = \frac{p}{m}$, so that the Hamiltonian becomes

$$H = p_k\dot{q}_k - L = p\frac{p}{m} - \frac{1}{2}m\left(\frac{p}{m}\right)^2 + \frac{1}{2}kq^2 = \frac{p^2}{2m} + \frac{1}{2}kq^2$$

This allows us to write the action as

$$S[q_k, p_k] = \int_{t_1}^{t_2} p_k\frac{\mathrm{d}q_k}{\mathrm{d}t} - H\,\mathrm{d}t$$

Thus the instantaneous action is

$$\begin{aligned}\delta S = S[q_k + \delta q_k, p_k + \delta p_k] - S[q_k, p_k] &\approx \int_{t_1}^{t_2} \frac{\mathrm{d}q_k}{\mathrm{d}t}\delta p_k + p_k\frac{\mathrm{d}(\delta q_k)}{\mathrm{d}t} - \frac{\partial H}{\partial q_k}\delta q_k - \frac{\partial H}{\partial p_k}\delta p_k\,\mathrm{d}t \\ &= \int_{t_1}^{t_2} \frac{\mathrm{d}q_k}{\mathrm{d}t}\delta p_k - \frac{\mathrm{d}p_k}{\mathrm{d}t}\delta q_k - \frac{\partial H}{\partial q_k}\delta q_k - \frac{\partial H}{\partial p_k}\delta p_k\,\mathrm{d}t + p_k\delta q_k|_{t_1}^{t_2}\end{aligned}$$

Note as opposed to Lagrangian mechanics, saying that position and momentum are independent is more justified than saying position and velocity are independent.

Note our boundary conditions are $\delta q_k(t_2) = \delta q_k(t_1) = 0$. By comparing coefficients, we find that

$$\boxed{\frac{\mathrm{d}q_k}{\mathrm{d}t} = \frac{\partial H}{\partial p_k}, \qquad \frac{\mathrm{d}p_k}{\mathrm{d}t} = -\frac{\partial H}{\partial q_k}}$$

These are called **Hamilton's Equations**. Thus we think of $\frac{\partial H}{\partial p_k}$ as the velocity, and $-\frac{\partial H}{\partial q_k}$ as the force.

In our case, these equations become

$$\dot{q} = \frac{p}{m}, \qquad \dot{p} = -kq$$

Note in practice, we will usually determine the Hamiltonian based on the Lagrangian, since it is not always $T+U$.

Thus instead of one simple equation and one very complicated one, we have two moderately complicated equations.

This formulation gives us a key advantage: rather than having n second order differential equations, Hamiltonian mechanics gives us $2n$ first order differential equations. It is significantly easier to numerically solve first order ODEs than second order ODEs.

Further, while in Lagrangian mechanics, we were working in **configuration space**, which is just position and velocities. Meanwhile, when we are doing Hamiltonian mechanics, we work in **phase space**, meaning that our coordinates are position and momentum. The reason that this is an advantage is that p is not necessarily $m\dot{q}$ - it may be another term, meaning that far more trajectories in phase space are possible than trajectories in configuration space.

Thus the tangent line to the physical trajectory in phase space is

$$\left(\frac{\partial H}{\partial p_k}, -\frac{\partial H}{\partial q_k}\right)$$

In the case of our old problem, we would plot the vector field

$$\left(\frac{p}{m}, -kq\right).$$

This gives an elliptical shape, implying that the motion is an ellipse. This tells us immediately that the motion of a spring system is a harmonic oscillator.

$$z = \frac{x^2}{2a}; \qquad L = \frac{1}{2}m\left(1+\left(\frac{x}{a}\right)^2\right)\dot{x}^2 - \frac{mg}{2a}x^2$$

$$\implies p = m\left(1+\left(\frac{x}{a}\right)^2\right)\dot{x} \quad H = \frac{p^2}{2m\left(1+\left(\frac{x}{a}\right)^2\right)} + \frac{mg}{2a}x^2$$

$$\left(\frac{p}{2m\left(1+\left(\frac{x}{a}\right)^2\right)}, -\frac{mg}{a}x + \frac{p^2}{2m\left(1+\left(\frac{x}{a}\right)^2\right)^2}\cdot\frac{2x}{a^2}\right)$$

9.2. Coordinate Transformations in Phase Space

Notice that in the previous derivation, we used $p = m\dot{q}$, which only applies in Cartesian coordinates. We often want to make some coordinate transformation, transforming our equations of motion appropriately to account for the fact that they're no longer Cartesian.

In Lagrangian mechanics, if we made a change of coordinates

$$Q_k = f_k(q_1, ..., q_N); \quad P_k = \frac{\partial L}{\partial \dot{Q}_k} = \sum_j \frac{\partial L}{\partial \dot{q}_j}\frac{\partial q_j}{\partial Q_k} = \sum_j \frac{\partial q_j}{\partial Q_k}P_j$$

Thus, we see the momentum is merely a function of the new coordinates, so that Q_k uniquely defines our system. However, this is not the case in Hamiltonian mechanics—here, position and momentum are separate coordinates.

Thus, to make an arbitrary transformation in Hamiltonian mechanics, we first want to remove the distinction between position and momentum in phase space, so that we can standardize notation.

Thus we develop **unified coordinates**. In particular, assign

$$\xi_k = \begin{cases} q_k \text{ if } 1 \le k \le N \\ p_{k-N} \text{ if } N+1 \le k \le 2N \end{cases}$$

For example, in the $N = 2$ case, we assign

$$\xi_1 = q_1, \quad \xi_2 = q_2, \quad \xi_3 = p_1, \quad \xi_4 = p_2$$

so that the equations of motion become

$$\dot{\xi}_1 = \frac{\partial H}{\partial \xi_3}, \quad \dot{\xi}_2 = \frac{\partial H}{\partial \xi_4}, \quad \dot{\xi}_3 = -\frac{\partial H}{\partial \xi_1}, \quad \dot{\xi}_4 = -\frac{\partial H}{\partial \xi_2}.$$

In general, the equations of motion are

$$\dot{\xi}_\alpha = w_{\alpha\beta}\frac{\partial H}{\partial \xi_\beta} \quad \text{with} \quad \omega_{\alpha\beta} = \begin{pmatrix} 0_{N\times N} & I_{N\times N} \\ -I_{N\times N} & 0_{N\times N} \end{pmatrix}.$$

For example for $N = 1$, $\hat{\omega} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and for $N = 2$,

$$\hat{\omega} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

Now consider some change of coordinates

$$\eta_\alpha(\xi_1, \ldots, \xi_{2N})$$

and then applying the chain rule gives

$$\dot{\eta}_\alpha = \sum_\beta \frac{\partial \eta_\alpha}{\partial \xi_\beta}\dot{\xi}_\beta = \sum_\beta \frac{\partial \eta_\alpha}{\partial \xi_\beta}\sum_\gamma \omega_{\beta\gamma}\frac{\partial H}{\partial \xi_\gamma} = \sum_\beta \frac{\partial \eta_\alpha}{\partial \xi_\beta}\sum_\gamma \omega_{\beta\gamma}\sum_\delta \frac{\partial H}{\partial \eta_\delta}\frac{\partial \eta_\delta}{\partial \xi_\gamma} = \sum_{\beta,\gamma,\delta} \omega_{\beta\gamma}\frac{\partial \eta_\alpha}{\partial \xi_\beta}\frac{\partial \eta_\delta}{\partial \xi_\gamma}\frac{\partial H}{\partial \eta_\delta} \equiv \hat{\chi}_{\alpha\delta}\frac{\partial H}{\partial \eta_\delta}$$

This is becoming an ungodly mess, so we restrict to transformations where $\hat{\chi} = \hat{\omega}$. We call this the **symplectic condition**, and corresponding transformations are called **canonical transformations**; the resulting coordinates are called **canonical coordinates**.

Note that this is the same kind of thing we do in special relativity: there, we restrict to working in inertial frames, only making Lorentz transformations, so that our equations of motion don't look insanely complicated. We call $\hat{\omega}$ the **symplectic matrix**.

Example: Simple Harmonic Oscillator

Consider the equations of motion

$$H = \frac{p^2}{2m} + \frac{m\omega^2}{2}q^2 \quad q = \sqrt{\frac{2\eta_2}{m\omega}}\sin\eta_1, \quad p = \sqrt{2m\omega\eta_2}\cos\eta_1$$

Here η_1 is our angular polar coordinate, and η_2 is our radial polar coordinate.

$$m\omega q^2 + \frac{p^2}{m\omega} = 2\eta_2\sin^2\eta_1 + 2\eta_2\cos^2\eta_1 = 2\eta_2 \implies H = \omega\eta_2 \quad m\omega\frac{q}{p} = \tan\eta_1$$

Then

$$\frac{\partial q}{\partial\eta_1} = \sqrt{\frac{2\eta_2}{m\omega}}\cos\eta_1, \quad \frac{\partial q}{\partial\eta_2} = \sqrt{\frac{1}{2m\omega\eta_2}}\sin\eta_1$$

and

$$\frac{\partial p}{\partial\eta_1} = -\sqrt{2m\omega\eta_2}\sin\eta_1, \quad \frac{\partial p}{\partial\eta_2} = \sqrt{\frac{m\omega}{2\eta_2}}\cos\eta_1$$

We should get

$$\frac{\partial q}{\partial\eta_1}\frac{\partial p}{\partial\eta_2} - \frac{\partial p}{\partial\eta_1}\frac{\partial q}{\partial\eta_2} = 1; \quad \frac{\partial p}{\partial\eta_1}\frac{\partial q}{\partial\eta_2} - \frac{\partial q}{\partial\eta_1}\frac{\partial q}{\partial\eta_2} = 0$$

which shows that the polar coordinates are indeed a canonical transformation.

Then

$$\dot\eta_1 = \frac{\partial H}{\partial\eta_2} = \omega; \quad \dot\eta_2 = -\frac{\partial H}{\partial\eta_1} = 0.$$

9.3. Poisson Brackets

Definition: Poisson Bracket

One particular quantity comes up a lot in Hamiltonian mechanics:

$$\{A, B\} = \omega_{\alpha\beta}\frac{\partial A}{\partial\xi_\alpha}\frac{\partial B}{\partial\xi_\beta} = \frac{\partial A}{\partial q_k}\frac{\partial B}{\partial p_k} - \frac{\partial A}{\partial p_k}\frac{\partial B}{\partial q_k}$$

We define this to be the **Poisson bracket**.

We can in fact describe our original coordinates using Poisson brackets with the Hamiltonian.

Example

$$\{\xi_\alpha, H\} = \omega_{\beta\gamma}\frac{\partial\xi_\alpha}{\partial\xi_\beta}\frac{\partial H}{\partial\xi_\gamma} = \omega_{\beta\gamma}\delta_{\alpha\beta}\frac{\partial H}{\partial\xi_\gamma} = \omega_{\alpha\gamma}\frac{\partial H}{\partial\xi_\gamma} = \dot\xi_\alpha$$

Now if η is a canonical coordinate, we have

$$\{A,B\} = \omega_{\alpha\beta}\frac{\partial A}{\partial \xi_\alpha}\frac{\partial B}{\partial \xi_\beta} = \omega_{\alpha\beta}\frac{\partial A}{\partial \eta_\gamma}\frac{\partial \eta_\gamma}{\partial \xi_\alpha}\frac{\partial B}{\partial \eta_\gamma}\frac{\partial \eta_\gamma}{\partial \xi_\beta} = \omega_{\gamma\delta}\frac{\partial A}{\partial \eta_\gamma}\frac{\partial B}{\partial \eta_\gamma}.$$

Thus we get a condition for having a canonical coordinate:

$$\boxed{\omega_{\alpha\beta} = \omega_{\gamma\delta}\frac{\partial \eta_\alpha}{\partial \xi_\gamma}\frac{\partial \eta_\beta}{\partial \xi_\delta} \iff \{\eta_\alpha, \eta_\beta\} = \omega_{\alpha\beta}}$$

***Proposition:* Properties of Poisson Brackets**

1. $\{A,B\} = -\{B,A\}$
2. $\{A,B+C\} = \{A,B\} + \{A,C\}$
3. $\{A,BC\} = B\{A,C\} + C\{A,B\}$
4. $\{A,\{B,C\}\} + \{B,\{C,A\}\} + \{C,\{A,B\}\} = 0$ (Jacobi identity)

Proof:

1. Follows from the antisymmetry of ω.
2. Linearity property of differentiation.
3. Follows from product rule.
4. Follows from first law of antisymmetry

□

Consider a coordinate transformation

$$\eta_\alpha = \xi_\alpha + \varepsilon\{\xi_\alpha, G\}$$

where $G(\xi_1, \ldots, \xi_{2n})$ is called the **generator**. This basically tells us the change in ξ_α under the transformation generated by G. Note that since $\{\xi_\alpha, G\} = -\{G, \xi_\alpha\}$, we can also view this is the change in G under the transformation generated by ξ_α.

Example

Suppose

$$\xi_1 = q, \quad \xi_2 = p, \quad G = p.$$

Then

$$\eta_1 = q + \varepsilon\{q,p\} = q + \varepsilon\left[\frac{\partial q}{\partial q}\frac{\partial p}{\partial p} - \frac{\partial q}{\partial p}\frac{\partial p}{\partial q}\right] = q + \varepsilon$$

$$\eta_2 = p + \varepsilon\{p,p\} = p$$

Thus the transformation generated by momentum is translation.

Now consider a general transformation of the Poisson bracket of two transformations:

$$\begin{aligned}\{\eta_\alpha, \eta_\beta\} &= \{\xi_\alpha + \varepsilon\{\xi_\alpha, G\}, \xi_\beta + \varepsilon\{\xi_\beta, G\}\} \\ &= \{\xi_\alpha, \xi_\beta\} + \{\xi_\alpha, \varepsilon\{\xi_\beta, G\}\} + \{\varepsilon\{\xi_\alpha, G\}, \xi_\beta\} + O(\varepsilon^2) \\ &= \omega_{\alpha\beta} + \varepsilon\left[\{\xi_\alpha, \{\xi_\beta, G\}\} - \{\{G, \xi_\beta\}, \xi_\alpha\} - \{\{\xi_\beta, \xi_\alpha\}, G\}\right] \qquad \text{(via Jacobi identity)}\end{aligned}$$

Now $\{\xi_\beta, \xi_\alpha\} = \omega_{\beta\alpha}$ which is constant, so the final term goes to zero. Also swap $\{\{G, \xi_\beta\}, \xi_\alpha\} = \{\xi_\alpha, \{\xi_\beta, G\}\}$. Thus the entire thing becomes $\omega_{\alpha\beta}$.

Definition: Active vs Passive Transformation

An **active transformation** takes the system and leaves the coordinates fixed, whereas a **passive transformation** moves coordinates some way and leaving the system fixed. These are obviously linked: an active transformation is the inverse of a passive transformation.

Consider an arbitrary quantity A, some function of position and momentum, such as the energy of the system. Then under an active transformation, it changes by

$$\begin{aligned}\delta A = A(\eta_\alpha) - A(\xi_\alpha) &= \frac{\partial A}{\partial \xi_\alpha}\varepsilon\{\xi_\alpha, G\} \\ &= \varepsilon\frac{\partial A}{\partial \xi_\alpha}\omega_{\beta\gamma}\frac{\partial \xi_\alpha}{\partial \xi_\beta}\frac{\partial G}{\partial \xi_\gamma} = \varepsilon\omega_{\alpha\gamma}\frac{\partial A}{\partial \xi_\alpha}\frac{\partial G}{\partial \xi_\gamma} = \varepsilon\{A, G\}\end{aligned}$$

If we make some transformation and the Hamiltonian does not change, i.e., $0 = \delta H$, then we should get a symmetry. We can also consider using the Hamiltonian as the generator, in which case the change in any property has to do with time.

THe Hamiltonian generates time translations. In other words, the generator is conserved. That is Noether's Theorem in the Hamiltonian setting.

Note the final term is usually zero.

***Theorem:* Noether's Theorem (Hamiltonian Setting)**

Suppose $\{G, H\} = 0$ and $\frac{\partial G}{\partial t} = 0$. Then G is conserved, and the transformation generated by G is a symmetry.

Proof: Let A be some quantity of the system. Then

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{\partial A}{\partial \xi_\alpha}\frac{\mathrm{d}\xi_\alpha}{\mathrm{d}t} + \frac{\partial A}{\partial t} = \frac{\partial A}{\partial \xi_\alpha}\omega_{\alpha\beta}\frac{\partial H}{\partial \xi_\beta} + \frac{\partial A}{\partial t} = \{A, H\} + \frac{\partial A}{\partial t}$$

Now under our assumption, we have that A is conserved.

□

Note this version of Noether's Theorem allows us to go in the reverse direction as well, since

$$\delta\xi_\alpha = \{\xi_\alpha, G\} \implies \delta q_k = f_k(q, p) \implies f_k = \frac{\partial G}{\partial p_k} = \{q_k, G\}$$

9.4. Applications in Quantum Mechanics

Suppose we have some equation with Poisson brackets $\{\cdot, \cdot\}$ and replace it with the commutator $\frac{1}{i\hbar}[\cdot, \cdot]$. Then this is a valid quantum mechanical system.

Example

Suppose $\{q, p\} = 1$. Then

$$\frac{[\hat{q}, \hat{p}]}{i\hbar} = 1$$

which we recognize to be the canonical relation between position and momentum in quantum mechanics.

Example

Consider the transformation

$$\underbrace{\frac{\mathrm{d}A}{\mathrm{d}t} = \{A, H\} + \frac{\partial A}{\partial t}}_{\text{Hamiltonian formulation}} \quad \Longrightarrow \quad \underbrace{\frac{\partial \hat{A}}{\partial t} = \frac{\left[\hat{A}, \hat{H}\right]}{i\hbar} + \frac{\partial \hat{A}}{\partial t}}_{\text{Heisenberg Picture}}$$

which we recognize as the Heisenberg equation of motion. Note that the Heisenberg picture of quantum mechanics expresses everything in terms of operators, whereas the Schrödinger picture expresses everything in terms of a wave function.

In addition, the way in which the Schrödinger equation was developed was analogous to Hamiltonian mechanics following from Lagrangian mechanics. Physicists initially wanted to find a PDE in $2N + 1$ variables.

This is the idea behind the **Hamilton-Jacobi Equation**:

$$\boxed{H\left(q_i, \frac{\partial S}{\partial q_i}, t\right) + \frac{\partial S}{\partial t} = 0}$$

where S is some generating function. S is some function of position and time, and somehow encodes the state of the system. This is similar to how in quantum mechanics ψ encodes the state of the system.

Example

For a free particle, the equation becomes

$$\frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 + \frac{\partial S}{\partial t} = 0.$$

The new momenta P are the integration constants we get from solving, and the new coordinates Q are given by $\frac{\partial S}{\partial P}$.

Proposition

The Schrödinger equation falls out of the Hamilton-Jacobi equation.

Proof: Begin with the ansatz $S(q_i, t) = W(q_i) - Et$ and $H\left(q_i, \frac{\partial W}{\partial q_i}, t\right) - E = 0$. Then using Cartesian coordinates,

$$E = \frac{1}{2m}(\nabla w)^2 + V$$

Compare that to the Schrödinger equation,

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi = E\psi$$

Now if we suppose ψ is an arbitrary complex number

$$\psi(\mathbf{r}) = e^{A(\mathbf{r}) + \frac{i}{\hbar}W(\mathbf{r})}$$

and plug in we get

$$\nabla\psi = \psi\left(\nabla A + \frac{i}{\hbar}\nabla W\right).$$

Thus

$$\begin{aligned}
\Delta\psi &= \nabla \cdot \nabla\psi \\
&= \psi\left(\nabla^2 A + \frac{i}{\hbar}\nabla^2 W\right) + \psi\left(\nabla A + \frac{i}{\hbar}\nabla W\right)^2 \\
&= \psi\left[\nabla^2 A + \frac{i}{\hbar}\nabla^2 W + (\nabla A)^2 - \frac{1}{\hbar^2}(\nabla W)^2 + \frac{2i}{\hbar}(\nabla A)\cdot(\nabla W)\right] \\
&\approx \frac{1}{2m}(\nabla W)^2 + V\psi = E\psi
\end{aligned}$$

if $\hbar$ is small. This is called the **small $\hbar$ limit of ψ**. This tells us that the function W, which we want to solve for in classical mechanics, is exactly the phase W of the complex number. It turns out that S is in fact the action of the classical mechanical system.

This is one way in which quantum mechanics reproduces classical mechanics in the macro limit.

□

10. Chaos Theory

11. Special Relativity