Lectures on Classical Mechanics

John C. Baez
Derek K. Wise

(version of January 15, 2018)

Preface

These are notes for a mathematics graduate course on classical mechanics at U.C. Riverside. I've taught this course three times recently. Twice I focused on the Hamiltonian approach. In 2005 I started with the Lagrangian approach, with a heavy emphasis on action principles, and derived the Hamiltonian approach from that. This approach seems more coherent.

Derek Wise took beautiful handwritten notes on the 2005 course, which can be found on my website:

http://math.ucr.edu/home/baez/classical/

Later, Blair Smith from Louisiana State University miraculously appeared and volunteered to turn the notes into LaTeX . While not yet the book I'd eventually like to write, the result may already be helpful for people interested in the mathematics of classical mechanics.

The chapters in this LaTeX version are in the same order as the weekly lectures, but I've merged weeks together, and sometimes split them over chapter, to obtain a more textbook feel to these notes. For reference, the weekly lectures are outlined here.

Week 1: (Mar. 28, 30, Apr. 1)—The Lagrangian approach to classical mechanics: deriving $F = ma$ from the requirement that the particle's path be a critical point of the action. The prehistory of the Lagrangian approach: D'Alembert's "principle of least energy" in statics, Fermat's "principle of least time" in optics, and how D'Alembert generalized his principle from statics to dynamics using the concept of "inertia force".

Week 2: (Apr. 4, 6, 8)—Deriving the Euler–Lagrange equations for a particle on an arbitrary manifold. Generalized momentum and force. Noether's theorem on conserved quantities coming from symmetries. Examples of conserved quantities: energy, momentum and angular momentum.

Week 3 (Apr. 11, 13, 15)—Example problems: (1) The Atwood machine. (2) A frictionless mass on a table attached to a string threaded through a hole in the table, with a mass hanging on the string. (3) A special-relativistic free particle: two Lagrangians, one with reparametrization invariance as a gauge symmetry. (4) A special-relativistic charged particle in an electromagnetic field.

Week 4 (Apr. 18, 20, 22)—More example problems: (4) A special-relativistic charged particle in an electromagnetic field in special relativity, continued. (5) A general-relativistic free particle.

Week 5 (Apr. 25, 27, 29)—How Jacobi unified Fermat's principle of least time and Lagrange's principle of least action by seeing the classical mechanics of a particle in a potential as a special case of optics with a position-dependent index of refraction. The ubiquity of geodesic motion. Kaluza-Klein theory. From Lagrangians to Hamiltonians.

Week 6 (May 2, 4, 6)—From Lagrangians to Hamiltonians, continued. Regular and strongly regular Lagrangians. The cotangent bundle as phase space. Hamilton's equations. Getting Hamilton's equations directly from a least action principle.

Week 7 (May 9, 11, 13)—Waves versus particles: the Hamilton-Jacobi equation. Hamilton's principal function and extended phase space. How the Hamilton-Jacobi equation foreshadows quantum mechanics.

Week 8 (May 16, 18, 20)—Towards symplectic geometry. The canonical 1-form and the symplectic 2-form on the cotangent bundle. Hamilton's equations on a symplectic manifold. Darboux's theorem.

Week 9 (May 23, 25, 27)—Poisson brackets. The Schrödinger picture versus the Heisenberg picture in classical mechanics. The Hamiltonian version of Noether's theorem. Poisson algebras and Poisson manifolds. A Poisson manifold that is not symplectic. Liouville's theorem. Weil's formula.

Week 10 (June 1, 3, 5)—A taste of geometric quantization. Kähler manifolds.

If you find errors in these notes, please email me! I thank Sheeyun Park and Curtis Vinson for catching lots of errors.

John C. Baez

Contents

Chapter 1

From Newtonian to Lagrangian Mechanics

Classical mechanics is a peculiar branch of physics with a long history. It used to be considered the sum total of our theoretical knowledge of the physical universe (Laplace's daemon, the Newtonian clockwork), but now it is known as an idealization, a toy model if you will. The astounding thing is that probably all professional applied physicists still use classical mechanics. So it is still an indispensable part of any physicist's or engineer's education.

It is so useful because the more accurate theories that we know of (general relativity and quantum mechanics) make corrections to classical mechanics generally only in extreme situations (black holes, neutron stars, atomic structure, superconductivity, and so forth). Given that general relativity and quantum mechanics are much harder theories to apply, it is no wonder that scientists revert to classical mechanics whenever possible.

So, what is classical mechanics?

1.1   Lagrangian and Newtonian Approaches

We begin by comparing the Newtonian approach to mechanics to the subtler approach of Lagrangian mechanics. Recall Newton's law:

$$F = ma \tag{1.1}$$

wherein we consider a particle moving in $\mathbb{R}^n$. Its position, say q, depends on time $t \in \mathbb{R}$, so it defines a function,

$$q \colon \mathbb{R} \longrightarrow \mathbb{R}^n.$$

From this function we can define **velocity**,

$$v = \dot{q} \colon \mathbb{R} \longrightarrow \mathbb{R}^n$$

where $\dot{q} = \frac{dq}{dt}$, and also **acceleration**,

$$a = \ddot{q} \colon \mathbb{R} \longrightarrow \mathbb{R}^n.$$

Now let $m > 0$ be the **mass** of the particle, and let F be a vector field on $\mathbb{R}^n$ called the **force**. Newton claimed that the particle satisfies $F = ma$. That is:

$$m\, a(t) = F\left(q(t)\right). \tag{1.2}$$

This is a 2nd-order differential equation for $q \colon \mathbb{R} \to \mathbb{R}^n$ which will have a unique solution given some $q(t_0)$ and $\dot{q}(t_0)$, provided the vector field F is 'nice' — by which we technically mean smooth and bounded (i.e., $|F(x)| < B$ for some $B > 0$, for all $x \in \mathbb{R}^n$).

We can then define a quantity called **kinetic energy**:

$$K(t) \ := \ \frac{1}{2} m\, v(t) \cdot v(t). \tag{1.3}$$

This quantity is interesting because

$$\begin{aligned}
\frac{d}{dt} K(t) &= m\, v(t) \cdot a(t) \\
&= F(q(t)) \cdot v(t).
\end{aligned}$$

So, kinetic energy goes up when you push an object in the direction of its velocity, and goes down when you push it in the opposite direction. Moreover,

$$\begin{aligned}
K(t_1) - K(t_0) &= \int_{t_0}^{t_1} F(q(t)) \cdot v(t)\, dt \\
&= \int_{t_0}^{t_1} F(q(t)) \cdot \dot{q}(t)\, dt.
\end{aligned}$$

So, the change of kinetic energy is equal to the **work** done by the force, that is, the integral of F along the curve $q \colon [t_0, t_1] \to \mathbb{R}^n$. In 3 dimensions, Stokes' theorem relating line integrals to surface integrals of the curl implies that the change in kinetic energy $K(t_1) - K(t_0)$ is independent of the curve going from $q(t_0) = a$ to $q(t_1) = b$ iff

$$\nabla \times F = 0.$$

This in turn is true iff

$$F = -\nabla V \tag{1.4}$$

for some function $V \colon \mathbb{R}^n \to \mathbb{R}$.

In fact, this conclusion is true even when $n \neq 3$, using a more general version of Stokes' theorem: the integral of F along a curve in $\mathbb{R}^n$ depends only on the endpoints of this curve iff $F = -\nabla V$ for some function V. Moreover, this function is then unique up

to an additive constant; we call this function the **potential**. A force with this property is called **conservative**. Why? Because in this case we can define the **total energy** of the particle by

$$E(t) := K(t) + V(q(t)) \tag{1.5}$$

where $V(t) := V(q(t))$ is called the **potential energy** of the particle, and then we can show that E is **conserved**: that is, constant as a function of time. To see this, note that $F = ma$ implies

$$\frac{d}{dt}[K(t) + V(q(t))] = F(q(t)) \cdot v(t) + \nabla V(q(t)) \cdot v(t)$$
$$= 0, \qquad (\text{because } F = -\nabla V).$$

Conservative forces let us apply a bunch of cool techniques. In the Lagrangian approach we define a quantity

$$L := K(t) - V(q(t)) \tag{1.6}$$

called the **Lagrangian**, and for any curve $q: [t_0, t_1] \to \mathbb{R}^n$ with $q(t_0) = a$, $q(t_1) = b$, we define the **action** to be

$$S(q) := \int_{t_0}^{t_1} L(t) \, dt. \tag{1.7}$$

From here one can go in two directions. One is to claim that nature causes particles to follow paths of least action, and derive Newton's equations from that principle. The other is to start with Newton's principles and find out what conditions, if any, on $S(q)$ follow from this. We will use the shortcut of hindsight, bypass the philosophy, and simply use the mathematics of variational calculus to show that particles follow paths that are 'critical points' of the action $S(q)$ if and only if Newton's law $F = ma$ holds. To do this,
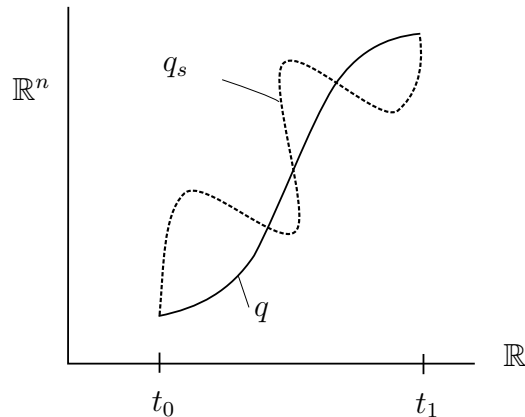

Figure 1.1: A particle can sniff out the path of least action.

let us look for curves (like the solid line in Fig. 1.1) that are **critical points** of S, namely:

$$\left.\frac{d}{ds}S(q_s)\right|_{s=0} = 0 \tag{1.8}$$

where

$$q_s = q + s\delta q$$

for all $\delta q \colon [t_0, t_1] \to \mathbb{R}^n$ with

$$\delta q(t_0) = \delta q(t_1) = 0.$$

To show that

$$F = ma \quad \Leftrightarrow \quad \left.\frac{d}{ds}S(q_s)\right|_{s=0} = 0 \text{ for all } \delta q \colon [t_0, t_1] \to \mathbb{R}^n \text{ with } \delta q(t_0) = \delta q(t_1) = 0 \tag{1.9}$$

we start by using the definition of the action and the chain rule:

$$\left.\frac{d}{ds}S(q_s)\right|_{s=0} = \left.\frac{d}{ds}\int_{t_0}^{t_1}\frac{1}{2}m\dot{q}_s(t)\cdot\dot{q}_s(t) - V(q_s(t))\,dt\right|_{s=0}$$

$$= \left.\int_{t_0}^{t_1}\frac{d}{ds}\left[\frac{1}{2}m\dot{q}_s(t)\cdot\dot{q}_s(t) - V(q_s(t))\right]dt\right|_{s=0}$$

$$= \left.\int_{t_0}^{t_1}\left[m\dot{q}_s\cdot\frac{d}{ds}\dot{q}_s(t) - \nabla V(q_s(t))\cdot\frac{d}{ds}q_s(t)\right]dt\right|_{s=0}$$

Next note that

$$\frac{d}{ds}q_s(t) = \delta q(t)$$

so

$$\frac{d}{ds}\dot{q}_s(t) = \frac{d}{ds}\frac{d}{dt}q_s(t) = \frac{d}{dt}\frac{d}{ds}q_s(t) = \frac{d}{dt}\delta q(t).$$

Thus we have

$$\left.\frac{d}{ds}S(q_s)\right|_{s=0} = \int_{t_0}^{t_1}\left[m\dot{q}\cdot\frac{d}{dt}\delta q(t) - \nabla V(q(t))\cdot\delta q(t)\right]dt.$$

Next we can integrate by parts, noting the boundary terms vanish because $\delta q = 0$ at t_1 and t_0:

$$\frac{d}{ds}S(q_s)|_{s=0} = \int_{t_0}^{t_1}[-m\ddot{q}(t) - \nabla V(q(t))]\cdot\delta q(t)dt\ .$$

It follows that variation in the action is zero for *all* variations δq iff the term in brackets is identically zero, that is,

$$-m\ddot{q}(t) - \nabla V(q(t)) = 0.$$

So, the path q is a critical point of the action S iff

$$F = ma.$$

The above result applies only for conservative forces, i.e., forces that can be written as minus the gradient of some potential. This is not true for all forces in nature: for example, the force on a charged particle in a magnetic field depends on its velocity as well as its position. However, when we develop the Lagrangian approach further we will see that it applies to this force as well!

1.1.1 Lagrangian versus Hamiltonian Approaches

I am not sure where to mention this, but before launching into the history of the Lagrangian approach may be as good a time as any. In later chapters we will describe another approach to classical mechanics: the Hamiltonian approach. Why do we need two approaches, Lagrangian and Hamiltonian?

They both have their own advantages. In the simplest terms, the Hamiltonian approach focuses on *position and momentum*, while the Lagrangian approach focuses on *position and velocity*. The Hamiltonian approach focuses on *energy*, which is a function of position and momentum — indeed, 'Hamiltonian' is just a fancy word for energy. The Lagrangian approach focuses on the *Lagrangian*, which is a function of position and velocity. Our first task in understanding Lagrangian mechanics is to get a gut feeling for what the Lagrangian means. The key is to understand the integral of the Lagrangian over time – the 'action', S. We shall see that this describes the 'total amount that happened' from one moment to another as a particle traces out a path. And, peeking ahead to quantum mechanics, the quantity $\exp(iS/\hbar)$, where $\hbar$ is Planck's constant, will describe the 'change in phase' of a *quantum* system as it traces out this path.

In short, while the Lagrangian approach takes a while to get used to, it provides invaluable insights into classical mechanics and its relation to quantum mechanics. We shall see this in more detail soon.

1.2 Prehistory of the Lagrangian Approach

We've seen that a particle going from point a at time t_0 to a point b at time t_1 follows a path that is a critical point of the action,

$$S = \int_{t_0}^{t_1} K - V \, dt$$

so that slight changes in its path do not change the action (to first order). Often, though not always, the action is minimized, so this is called the **Principle of Least Action**.

Suppose we did not have the hindsight afforded by the Newtonian picture. Then we might ask, "Why does nature like to minimize the action? And why *this* action $\int K - V \, dt$? Why not some other action?"

'Why' questions are always tough. Indeed, some people say that scientists should never ask 'why'. This seems too extreme: a more reasonable attitude is that we should only ask a 'why' question if we expect to learn something scientifically interesting in our attempt to answer it.

There are certainly some interesting things to learn from the question "why is action minimized?" First, note that total energy is conserved, so energy can slosh back and forth between kinetic and potential forms. The Lagrangian $L = K - V$ is big when most of the energy is in kinetic form, and small when most of the energy is in potential form. Kinetic energy measures how much is 'happening' — how much our system is moving around. Potential energy measures how much *could* happen, but isn't yet — that's what the word 'potential' means. (Imagine a big rock sitting on top of a cliff, with the potential to fall down.) So, the Lagrangian measures something we could vaguely refer to as the 'activity' or 'liveliness' of a system: the higher the kinetic energy the more lively the system, the higher the potential energy the less lively. So, we're being told that nature likes to minimize the total of 'liveliness' over time: that is, the total action.

In other words, nature is as lazy as possible!

For example, consider the path of a thrown rock in the Earth's gravitational field, as in Fig. 1.2. The rock traces out a parabola, and we can think of it as doing this in order

$K - V$ small here. . . good!

Spend as much time as possible here

PSfrag replacements

$K - V$ big. . . bad!

Get this over with quick!
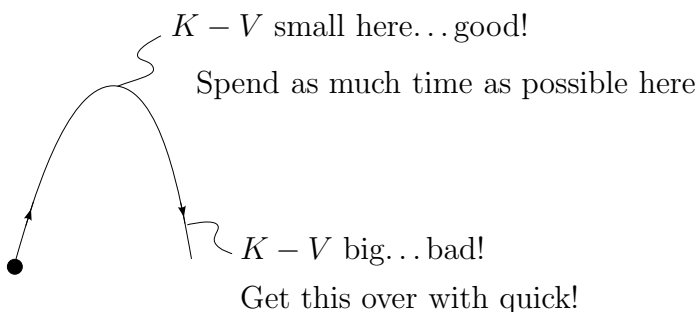
Figure 1.2: A particle's "lazy" motion minimizes the action.

to minimize its action. On the one hand, it wants to spend a lot much time near the top of its trajectory, since this is where the kinetic energy is least and the potential energy is greatest. On the other hand, if it spends *too* much time near the top of its trajectory, it will need to really rush to get up there and get back down, and this will take a lot of action. The perfect compromise is a parabolic path!

Here we are anthropomorphizing the rock by saying that it 'wants' to minimize its action. This is okay if we don't take it too seriously. Indeed, one of the virtues of the Principle of Least Action is that it lets us put ourselves in the position of some physical system and imagine what we would do to minimize the action.

There is another way to make progress on understanding 'why' action is minimized: history. Historically there were two principles that were fairly easy to deduce from observations of nature: (i) the principle of least time, used in optics, and (ii) the principle of minimum energy, used in statics. By putting these together, we can guess the principle of least action. So, let us recall these earlier minimum principles.

1.2.1 The Principle of Least Time

In 1662, Pierre Fermat pointed out that light obeys the **principle of least time**: when a ray of light goes from one point to another, it chooses the path that takes the last time. It was known much earlier that moving freely through the air light moves in straight lines, which in Euclidean space are the shortest paths from one point to another. But more interesting than straight lines are piecewise straight paths and curves. Consider reflection of light from a mirror:
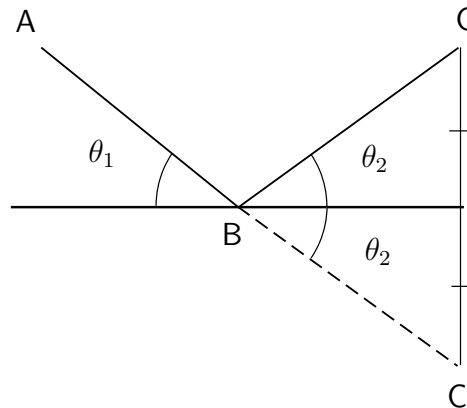
What path does the light take? The empirical answer was known at least since Euclid: it chooses B such that $\theta_1 = \theta_2$, angle of incidence equals the angle of reflection. But Hero of Alexandria pointed out that this is precisely the path that minimizes the length of the trajectory subject to the condition that it must hit the mirror (at least at one point). In fact light traveling from A to B takes *both* the straight paths ABC and AC. Why is ABC the shortest path hitting the mirror? This follows from some basic Euclidean geometry:

$$B \text{ minimizes } AB + BC \Leftrightarrow B \text{ minimizes } AB + BC'$$
$$\Leftrightarrow A, B, C' \text{ lie on a line}$$
$$\Leftrightarrow \theta_1 = \theta_2.$$

Note the introduction of the fictitious image C′ "behind" the mirror. A similar trick is now used in solving electrostatic problems: a conducting surface can be replaced by fictitious mirror image charges to satisfy the boundary conditions. (We also see this method in geophysics when one has a geological fault, and in hydrodynamics when there is a boundary between two media.)

However, the big clue came from refraction of light. Consider a ray of light passing from one medium to another: In 984 AD, the Persian scientist Ibn Sahl pointed out that
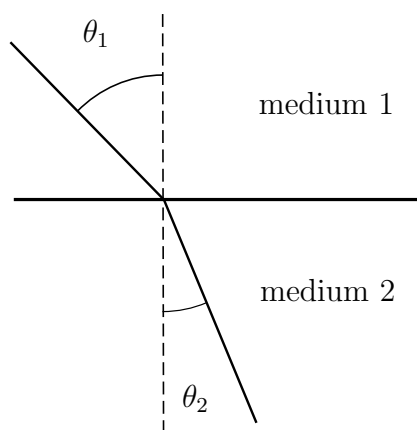
each medium has some number n associated with it, called its **index of refraction**, such that

$$n_1 \sin \theta_1 = n_2 \sin \theta_2.$$

This principle was rediscovered in the 1600s by the Dutch astronomer Willebrord Snellius, and is usually called **Snell's law**. It is fundamental to the design of lenses.

In 1662, Pierre de Fermat pointed out in a letter to a friend that Snell's law would follow if the speed of light were proportional to $1/n$ and light minimized the time it takes to get from A to C. Note: in this case it is the *time* that is important, not the length of the path. But the same is true for the law of reflection, since in that case the path of minimum length gives the same results as the path of minimum time.

So, not only is light the fastest thing around, it's also always taking the quickest path from here to there!

In fact, this idea seems to go back at least to 1021, when Ibn al-Haytham, a scientist in Cairo, mentioned it in his *Book of Optics*. But the French physicists who formulated the principle of least action were much more likely to have been influenced by Fermat.

Fermat's friend, Cureau de la Chambre, was unconvinced:

> The principle which you take as the basis for your proof, namely that Nature always acts by using the simplest and shortest paths, is merely a moral, and not a physical one. It is not, and cannot be, the cause of any effect in Nature.

The same philosophical objection is often raised against the principle of least action. That is part of what makes the principle so interesting: how does nature "know" how to take the principle of least action? The best explanation so far involves quantum mechanics. But I am getting ahead of myself here.

1.2.2 The Principle of Minimum Energy

Another principle foreshadowing the principle of least action was the "principle of minimum energy". Before physicists really got going in studying dynamics they thought a lot about statics. **Dynamics** is the study of moving objects, while **statics** is the study of objects at rest, or in equilibrium.

PSfrag replacements
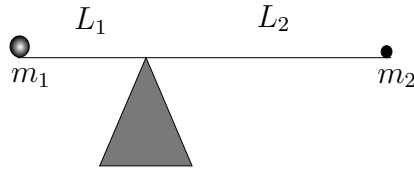
Figure 1.3: A principle of energy minimization determines a lever's balance.

For example, Archimedes studied the laws of a see-saw or lever (Fig. 1.3), and he found that this would be in equilibrium if

$$m_1 L_1 = m_2 L_2.$$

This can be understood using the "principle of virtual work", which was formalized quite nicely by Johann Bernoulli in 1715. Consider moving the lever slightly, i.e., infinitesimally, In equilibrium, the infinitesimal work done by this motion is zero! The reason is that the

PSfrag replacements
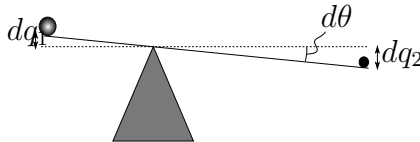

Figure 1.4: A principle of energy minimization determines a lever's balance.

work done on the ith body is

$$dW_i = F_i dq_i$$

and gravity pulls down with a force $m_i g$, so

$$dW_1 = (0, 0, -m_1 g) \cdot (0, 0, -L_1 d\theta)$$
$$= m_1 g L_1 \, d\theta$$

and similarly

$$dW_2 = -m_2 g L_2 \, d\theta$$

The total "virtual work" $dW = dW_1 + dW_2$ vanishes for all $d\theta$ (that is, for all possible infinitesimal motions) precisely when

$$m_1 L_1 - m_2 L_2 = 0$$

which is just as Archimedes wrote.

1.2.3   Virtual Work

Let's go over the above analysis in more detail. I'll try to make it clear what we mean by **virtual work**.

The forces and constraints on a system may be time dependent. So equal small infinitesimal displacements of the system might result in the forces $\mathbf{F}_i$ acting on the system doing different amounts of work at different times. To *displace a system by $\delta\mathbf{r}_i$ for each position coordinate, and yet remain consistent with all the constraints and forces at a given instant of time t*, without any time interval passing is called a *virtual displacement*. It's called 'virtual' because it cannot be realized: any actual displacement would occur over a finite time interval and possibly during which the forces and constraints might change. Now call the work done by this hypothetical virtual displacement, $\mathbf{F}_i \cdot \delta\mathbf{r}_i$, the **virtual work**. Consider a system in the special state of being in equilibrium, i.e., when $\sum \mathbf{F}_i = 0$. Then because by definition the virtual displacements do not change the forces, we must deduce that the virtual work vanishes for a system in equilibrium,

$$\sum_i \mathbf{F}_i \cdot \delta\mathbf{r}_i = 0, \qquad \text{(when in equilibrium)} \tag{1.10}$$

Note that in the above example we have two particles in $\mathbb{R}^3$ subject to a constraint (they are pinned to the lever arm). However, a number n of particles in $\mathbb{R}^3$ can be treated as a single quasi-particle in $\mathbb{R}^{3n}$, and if there are constraints it can move in some submanifold of $\mathbb{R}^{3n}$. So ultimately we need to study a particle on an arbitrary *manifold*. But, we'll postpone such sophistication for a while.

For a particle in $\mathbb{R}^n$, the principle of virtual work says

$$q(t) = q_0 \quad \text{satisfies} \quad F = ma, \qquad\qquad \text{(it's in equilibrium)}$$
$$\Updownarrow$$
$$dW = F \cdot dq \quad \text{vanishes for all} \quad dq \in \mathbb{R}^n, \qquad \text{(virtual work is zero for } \delta q \to 0\text{)}$$
$$\Updownarrow$$
$$F = 0, \qquad\qquad\qquad\qquad\qquad\qquad \text{(no force on it!)}$$

If the force is conservative ($F = -\nabla V$) then this is also equivalent to,

$$\nabla V(q_0) = 0$$

that is, we have equilibrium at a critical point of the potential. The equilibrium will be **stable** if q_0 is a local minimum of the potential V.
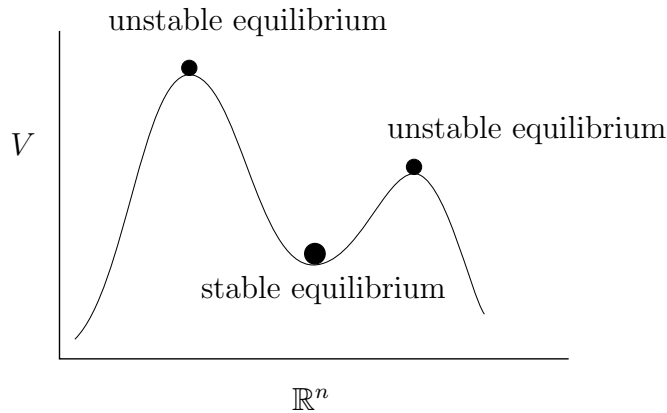
Figure 1.5: A principle of energy minimization determines a lever's balance.

We can summarize all the above by proclaiming that we have a "principle of least energy" governing stable equilibria. We also have an analogy between statics and dynamics:

Statics	Dynamics
equilibrium, $F = 0$	$F = ma$
potential, V	action, $S = \int_{t_0}^{t_1} K - V \, dt$
critical points of V	critical points of S

1.2.4   From Virtual Work to the Principle of Least Action

Sometimes laws of physics are just guessed using a bit of intuition and a gut feeling that nature must be beautiful or elegantly simple (though occasionally awesomely complex in beauty). One way to make good guesses is to *generalize*.

The principle of virtual work for statics says that equilibrium occurs when

$$F(q_0) \cdot \delta q = 0, \qquad \forall \delta q \in \mathbb{R}^n$$

Around 1743, D'Alembert generalized this principle to dynamics in his *Traite de Dynamique.* He did it by inventing what he called the "inertia force", $-m\,a$, and postulating

that in dynamics equilibrium occurs when the so-called **total force**, $F - ma$, vanishes. Of course this is just a restatement of Newton's law $F = ma$. But this allowed him to generalize the principle of virtual work from statics to dynamics. Namely, a particle will trace out a path $q \colon [t_0, t_1] \to \mathbb{R}^n$ obeying

$$[F(q(t)) - ma(t)] \cdot \delta q(t) = 0 \tag{1.11}$$

for all $\delta q \colon [t_0, t_1] \to \mathbb{R}^n$ with

$$\delta q(t_0) = \delta q(t_1) = 0.$$

Let us see how this principle implies the principle of least action. We create a family of paths parameterized by s in the usual way

$$q_s(t) = q(t) + s\,\delta q(t)$$

and define the **variational derivative** of any function f on the space of paths by

$$\delta f(q) = \frac{d}{ds} f(q_s)\Big|_{s=0}. \tag{1.12}$$

Then D'Alembert's generalized principle of virtual work implies

$$\int_{t_0}^{t_1} [(F(q(t)) - m\ddot{q}(t)] \cdot \delta q(t)\, dt = 0$$

for all δq, so if $F = -\nabla V$ we have

$$\begin{aligned}
0 &= \int_{t_0}^{t_1} [-\nabla V(q(t))) - m\ddot{q}(t)] \cdot \delta q(t)\, dt \\
&= \int_{t_0}^{t_1} [-\nabla V(q(t)) \cdot \delta q(t) + m\dot{q}(t) \cdot \delta \dot{q}(t)]\ dt
\end{aligned}$$

where in the second step we did an integration by parts, which has no boundary terms since $\delta q(t_0) = \delta q(t_1) = 0$. Next, using

$$\delta V(q(t)) \;\; = \frac{d}{ds} V(q_s(t))\Big|_{s=0} \;\; = \nabla V(q) \cdot \frac{dq_s(t)}{ds}\Big|_{s=0} \;\; = \nabla V(q(t)) \cdot \delta q(t)$$

and

$$\delta(\dot{q}(t)^2) = 2\dot{q}(t) \cdot \delta \dot{q}(t)$$

we obtain

$$
\begin{aligned}
0 &= \int_{t_0}^{t_1} \left[-\nabla V(q(t)) \cdot \delta q(t) + m\dot{q}(t) \cdot \delta\dot{q}(t) \right] dt \\
&= \int_{t_0}^{t_1} \left[-\delta V(q(t)) + \frac{m}{2}\delta(\dot{q}(t)^2) \right] dt \\
&= \delta\left(\int_{t_0}^{t_1} \left[-V(q(t)) + \frac{m}{2}\dot{q}(t)^2 \right] dt \right) \\
&= \delta\left(\int_{t_0}^{t_1} \left[K(t) - V(q(t)) \right] dt \right)
\end{aligned}
$$

and thus

$$
\delta S(q) = 0
$$

where

$$
S(q) = \int_{t_0}^{t_1} \left[K(t) - V(q(t)) \right] dt
$$

is the action of the path q.

In fact, Joseph-Louis Lagrange presented a calculation of this general sort in 1768, and this idea underlies his classic text *Mécanique Analitique*, which appeared 20 years later. This is why $K - V$ is called the Lagrangian.

I hope you now see that the principle of least action is a natural generalization of the principle of minimum energy from statics to dynamics. Still, there's something unsatisfying about the treatment so far. I did not really explain why one must introduce the "inertia force"—except, of course, that we need it to obtain agreement with Newton's $F = ma$.

We conclude with a few more words about this mystery. Recall from undergraduate physics that in an accelerating coordinate system there is a fictional force $-ma$, which is called the **centrifugal force**. We use it, for example, to analyze simple physics in a rotating reference frame. If you are *inside* the rotating system and you throw a ball straight ahead it will appear to curve away from your target, and if you did not know that you were rotating relative to the rest of the universe then you'd think there was a force on the ball equal to the centrifugal force. If you are inside a big rapidly rotating drum then you'll also feel pinned to the walls. This is an example of an inertia force which comes from using a funny coordinate system. In fact, in general relativity one sees that—in a certain sense—*gravity* is an inertia force! But more about this later.

Chapter 2

Lagrangian Mechanics

In this chapter we'll look at Lagrangian mechanics in more generality, and show the principle of least action is equivalent to some equations called the Euler–Lagrange equations.

2.1 The Euler–Lagrange Equations

We are going to start thinking of a general classical system as a set of points in an abstract *configuration space* or phase space[1]. So consider an arbitrary classical system as living in a space of points in some manifold Q. For example, the space for a spherical double pendulum would look like Fig. 2.1, where $Q = S^2 \times S^2$. So our system is "a particle

$$Q = S^2 \times S^2$$

PSfrag replacements
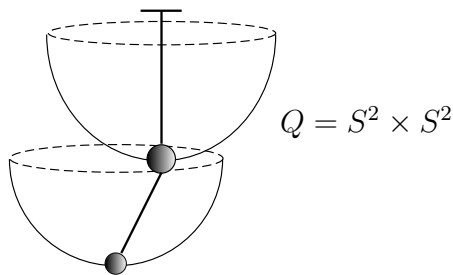
Figure 2.1: Double pendulum configuration space.

in Q", which means you have to disabuse yourself of the notion that we're dealing with real particles: what we're really dealing with is a *single abstract particle* in an abstract higher dimensional space. This single abstract particle represents two real particles if we are talking about the classical system in Fig. 2.1. Sometimes to make this clear we'll talk

---

[1]The tangent bundle TQ will be referred to as *configuration space*, later on when we get to the chapter on Hamiltonian mechanics we'll find a use for the cotangent bundle T^*Q, and normally we call this the *phase space.*

about "the system taking a path", instead of "the particle taking a path". It is then clear that when we say, "the system follows a path $q(t)$" that we're referring to the point q in configuration space Q that represents all of the particles in the real system.

So as time passes, "the system" traces out a path

$$q\colon [t_0, t_1] \longrightarrow Q$$

and we define its **velocity**

$$\dot{q}(t) \in T_{q(t)}Q$$

to be the tangent vector at $q(t)$ given by the equivalence class $[\sigma]$ of curves through $q(t)$ with derivatives $\dot{\sigma}(t) = dq(s)/ds|_{s=t}$. We'll just write is as $\dot{q}(t)$.

Let Γ be the space of smooth paths from $a \in Q$ to $b \in Q$,

$$\Gamma = \{q\colon [t_0, t_1] \to Q | q(t_0) = a, q(t_1) = b\}$$

(Γ is an infinite dimensional manifold, but we won't go into that for now.) Let the **Lagrangian** for the system be *any* smooth function of position and velocity:

$$L : TQ \longrightarrow \mathbb{R}$$

and define the **action**

$$S\colon \Gamma \longrightarrow \mathbb{R}$$

by

$$S(q) = \int_{t_0}^{t_1} L(q, \dot{q}) \, dt \tag{2.1}$$

The path that our abstract particle will actually take is a critical point of S. In other words, it will choose a path $q \in \Gamma$ such that for any smooth 1-parameter family of paths $q_s \in \Gamma$ with $q_0 = q$, we have

$$\frac{d}{ds} S(q_s) \Big|_{s=0} = 0. \tag{2.2}$$

For any function f on the space of paths we define its **variational derivative** by

$$\delta f(q) = \frac{d}{ds} f(q) \Big|_{s=0}$$

so that Eq. (2.2) can be rewritten simply as

$$\delta S(q) = 0. \tag{2.3}$$

What is a "1-parameter family of paths"? It is nothing more nor less than a set of well-defined paths $\{q_s\}$, each one labeled by a parameter s. For a "smooth" 1-parameter family of paths, $q_s(t)$ depends smoothly on both s and t. Thus, in Fig. 2.2 we can go from q_0 to q_s by smoothly varying the parameter from 0 to a given value s.
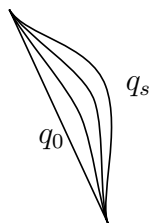
Figure 2.2: Schematic of a 1-parameter family of curves.
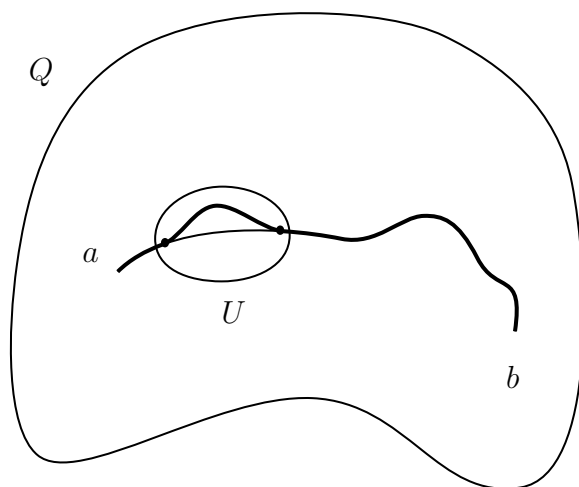
Figure 2.3: Local path variation.

Since Q is a manifold, it admits a covering by coordinate charts. For now, let's pick coordinates in a neighborhood U of some point $q(t) \in Q$. Next, consider only variations q_s such that $q_s = q$ outside U. A cartoon of this looks like Fig. 2.3 Then we restrict attention to a subinterval $[t_0', t_1'] \subseteq [t_0, t_1]$ such that $q_s(t) \in U$ for $t_0' \le t \le t_1'$.

Let's just go ahead and rename t_0' and t_1' as " t_0 and t_1" to drop the primes. We can use the coordinate charts on U,

$$\varphi \colon U \longrightarrow \mathbb{R}^n$$
$$x \longmapsto \varphi(x) = (x^1, x^2, \ldots, x^n)$$

and we also have coordinates for the tangent vectors

$$d\varphi \colon TU \longrightarrow T\mathbb{R}^n \cong \mathbb{R}^n \times \mathbb{R}^n$$
$$(x, y) \longmapsto d\varphi(x, y) = (x^1, \ldots, x^n, y^1, \ldots, y^n)$$

where $y \in T_x Q$. We can restrict $L \colon TQ \to \mathbb{R}$ to $TU \subseteq TQ$, and then we can describe L using the coordinates x^i, y^i on TU. The x^i are *position* coordinates, while the y^i are the

associated *velocity* coordinates. Using these coordinates we have

$$\delta S = \delta \int_{t_0}^{t_1} L(q(t), \dot{q}(t))\, dt$$

$$= \int_{t_0}^{t_1} \delta L(q, \dot{q})\, dt$$

$$= \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial x^i} \delta q^i + \frac{\partial L}{\partial y^i} \delta \dot{q}^i \right) dt$$

where we've used the smoothness of L and the Einstein summation convention for repeated indices i. Note that we can write δL as above using a local coordinate patch because the path variations δq are entirely trivial outside the patch for U. Continuing, using the Leibniz rule

$$\frac{d}{dt}\left(\frac{\partial L}{\partial y} \delta q \right) = \frac{d}{dt}\frac{\partial L}{\partial y} \delta q + \frac{\partial L}{\partial y} \delta \dot{q}$$

we have

$$\delta S = \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial x^i} - \frac{d}{dt}\frac{\partial L}{\partial y^i} \right) \delta q^i(t)\, dt$$

$$= 0.$$

If this integral is to vanish as demanded by $\delta S = 0$, then it must vanish for all path variations δq, further, the boundary terms vanish because we deliberately chose δq that vanish at the endpoints t_0 and t_1 inside U. That means the term in brackets must be identically zero, or

$$\frac{d}{dt}\frac{\partial L}{\partial y^i} - \frac{\partial L}{\partial x^i} = 0 \tag{2.4}$$

This is necessary to get $\delta S = 0$, for all δq, but in fact it's also sufficient. Physicists always give the coordinates x^i, y^i on TU the names "q^i" and "$\dot{q}^i$", despite the fact that these symbols also have another meaning, namely the x^i and y^i coordinates of the point

$$\big(q(t), \dot{q}(t)\big) \in TU.$$

Thus, physicists write

$$\boxed{\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = \frac{\partial L}{\partial q^i}}$$

and they call these the **Euler–Lagrange equations**.

   Our derivation of these equations was fairly abstract: we used the terms "position" and "velocity", but we did not assume these were the usual notions of position and velocity

for a particle in $\mathbb{R}^3$, or even $\mathbb{R}^n$. So, to bring things down to earth, consider the good old familiar case where the configuration space Q is $\mathbb{R}^n$ and the Lagrangian is

$$L(q, \dot{q}) = \frac{1}{2}m\dot{q} \cdot \dot{q} - V(q)$$

In this case

$$\frac{\partial L}{\partial q^i} = -\frac{\partial V}{\partial q^i} = F_i$$

are the components of the the force on the particle, while

$$\frac{\partial L}{\partial \dot{q}^i} = m\dot{q}^i$$

are the components of its mass times its velocity, also known as its **momentum**. In physics momentum is denoted by p for some obscure reason, so we say

$$\frac{\partial L}{\partial \dot{q}^i} = p_i$$

and the Euler–Lagrange equations say simply

$$\frac{dp}{dt} = F.$$

The time derivative of momentum is force! Since dp/dt is also mass times acceleration, this is just another way of stating Newton's law

$$F = ma.$$

Based on this example, we can make up nice names for the quantities in the Euler–Lagrange equation in general, for any Lagrangian $L \colon TQ \to \mathbb{R}$. We define

$$F_i = \frac{\partial L}{\partial q^i}$$

and call this quantity the **force**, and we define

$$p_i = \frac{\partial L}{\partial \dot{q}^i}$$

and call this quantity the **momentum**. The Euler–Lagrange equations then say

$$\frac{dp_i}{dt} = F_i.$$

Written this way, the general Euler–Lagrange equations are revealed to be a generalization of Newton's law, with ma replaced by the time derivative of momentum.

Term	Meaning for a particle in a potential	Meaning in general
$\dfrac{\partial L}{\partial \dot{q}^i}$	mv_i	momentum: p_i
$\dfrac{\partial L}{\partial q^i}$	$-\dfrac{\partial V}{\partial q^i}$	force: F_i

2.2   Noether's Theorem

If the form of a system of dynamical equations does not change under spatial translations then the momentum is a conserved quantity. When the form of the equations is similarly invariant under time translations then the total energy is a conserved quantity (a constant of the equations of motion). Time and space translations are examples of 1-parameter groups of transformations. *Invariance under a group of transformations* is precisely what we mean by a symmetry in group theory. So symmetries of a dynamical system give conserved quantities or conservation laws. The rigorous statement of all this is the content of *Noether's theorem.*

2.2.1   Time Translation

To handle time translations we need to replace our paths $q \colon [t_0, t_1] \to Q$ by paths $q : \mathbb{R} \to Q$, and then define a new space of paths,

$$\Gamma = \{q : \mathbb{R} \to Q\}.$$

The bad news is that the action

$$S(q) = \int_{-\infty}^{\infty} L\Big(q(t), \dot{q}(t)\Big)\, dt$$

typically will not converge, so S is then no longer a function of the space of paths. Nevertheless, if $\delta q = 0$ outside of some finite interval, then the functional variation,

$$\delta S := \int_{-\infty}^{\infty} \frac{d}{ds} L\Big(q_s(t), \dot{q}_s(t)\Big)\bigg|_{s=0}\, dt$$

will converge, since the integral is smooth and vanishes outside this interval. Moreover, demanding that this δS vanishes for all such variations δq is enough to imply the Euler–

Lagrange equations:

$$\delta S = \int_{-\infty}^{\infty} \frac{d}{ds} L\Big(q_s(t), \dot{q}_s(t)\Big)\Big|_{s=0} dt$$

$$= \int_{-\infty}^{\infty} \left(\frac{\partial L}{\partial q_i}\delta q_i + \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i\right) dt$$

$$= \int_{-\infty}^{\infty} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}\right) \delta q_i \, dt$$

where again the boundary terms have vanished since $\delta q = 0$ near $t = \pm\infty$. To be explicit, the first term in

$$\frac{\partial L}{\partial \dot{q}^i}\delta \dot{q}^i = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\delta q^i\right) - \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}\right)\delta q$$

vanishes when we integrate. Then the whole thing vanishes for all compactly supported smooth δq iff

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

So, we get Euler–Lagrange equations again.

Generalized Coordinates

2.2.2   Symmetries

First, let's give a useful definition that will make it easy to refer to a type of dynamical system symmetry. We want to refer to symmetry transformations (of the Lagrangian) governed by a single parameter.

Definition 2.1 (one-parameter family of symmetries)**.** *A 1-parameter family of symmetries of a Lagrangian system $L : TQ \to \mathbb{R}$ is a smooth map,*

$$F : \mathbb{R} \times \Gamma \longrightarrow \Gamma$$

$$(s, q) \longmapsto q_s, \qquad \text{with } q_0 = q$$

such that there exists a function $\ell(q, \dot{q})$ for which

$$\delta L = \frac{d\ell}{dt}$$

for some $\ell : TQ \to \mathbb{R}$, that is,

$$\frac{d}{ds} L\Big(q_s(t), \dot{q}_s(t)\Big)\Big|_{s=0} = \frac{d}{dt}\ell\Big(q_s(t), \dot{q}_s(t)\Big)$$

for all paths q.

Remark: The simplest case is $\delta L = 0$, in which case we really have a way of moving paths around ($q \mapsto q_s$) that doesn't change the Lagrangian—i.e., a symmetry of L in the most obvious way. But $\delta L = \frac{d}{dt}\ell$ is a sneaky generalization whose usefulness will become clear.

2.2.3   Noether's Theorem

Here's a statement of the theorem. Note that ℓ in this theorem is the function associated with F in Definition 2.1.

Theorem 2.1 (Noether's Theorem). *Suppose F is a one-parameter family of symmetries of the Lagrangian system, $L : TQ \to \mathbb{R}$. Then,*

$$p_i \delta q^i - \ell$$

is conserved, that is, its time derivative is zero for any path $q \in \Gamma$ satisfying the Euler–Lagrange equations. In other words, in boring detail:

$$\frac{d}{dt}\left[\frac{\partial L}{\partial y^i}\big(q(s)\dot{q}(s)\big) \left.\frac{d}{ds}q_s^i(t)\right|_{s=0} - \ell\big(q(t), \dot{q}(t)\big)\right] = 0$$

Proof.

$$\frac{d}{dt}\left(p^i \delta q_i - \ell \right) = \dot{p}_i \delta q^i + p_i \delta \dot{q}^i - \frac{d}{dt}\ell$$

$$= \frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial \dot{q}^i}\delta \dot{q}^i - \delta L$$

$$= \delta L - \delta L = 0. \qquad \square$$

"Okay, big deal" you might say. Before this can be of any use we'd need to find a symmetry F. Then we'd need to find out what this $p_i \delta q^i - \ell$ business is that is conserved. So let's look at some examples.

2.2.4   Conservation of Energy

1. **Conservation of Energy.** (The most important example!)

   All of our Lagrangian systems will have time translation invariance (because the laws of physics do not change with time, at least not to any extent that we can tell). So we have a one-parameter family of symmetries

$$q_s(t) = q(t + s)$$

   This indeed gives,

$$\delta L = \dot{L}$$

for

$$\frac{d}{ds}L(q_s)\Big|_{s=0} = \frac{d}{dt}L = \dot{L}$$

so here we take $\ell = L$ simply! We then get the conserved quantity

$$p_i\delta q^i - \ell = p_i\dot{q}^i - L$$

which we normally call the **energy**. For example, if $Q = \mathbb{R}^n$, and if

$$L = \frac{1}{2}m\dot{q}^2 - V(q)$$

then this quantity is

$$m\dot{q}\cdot\dot{q} - \left(\frac{1}{2}m\dot{q}\cdot\dot{q} - V\right) = \frac{1}{2}m\dot{q}^2 + V(q).$$

The term in parentheses is $K - V$, and the left-hand side is $K + V$.

Let's repeat this example, this time with a specific Lagrangian. It doesn't matter what the Lagrangian is, if it has 1-parameter families of symmetries then it'll have conserved quantities, guaranteed. The trick in physics is to write down a correct Lagrangian in the first place! (Something that will accurately describe the system of interest.)

2.3   Conserved Quantities from Symmetries

We've seen that any 1-parameter family

$$F_s : \Gamma \longrightarrow \Gamma$$
$$q \longmapsto q_s$$

which satisfies

$$\delta L = \dot{\ell}$$

for some function $\ell = \ell(q, \dot{q})$ gives a conserved quantity

$$p_i\delta q^i - \ell.$$

As usual we've defined

$$\delta L = \frac{d}{ds}L\Big(q_s(t), \dot{q}_s(t)\Big)\Big|_{s=0}.$$

Let's see how we arrive at a conserved quantity from a symmetry.

2.3.1   Time Translation Symmetry

For any Lagrangian system, $L : TQ \to \mathbb{R}$, we have a 1-parameter family of symmetries

$$q_s(t) = q(t + s)$$

because

$$\delta L = \dot{L}$$

so we get a conserved quantity called the **total energy** or **Hamiltonian**,

$$H = p_i \dot{q}^i - L \tag{2.5}$$

(You might prefer "Hamiltonian" to "total energy" because in general we are not in the same configuration space as Newtonian mechanics, if you are doing Newtonian mechanics then "total energy" is appropriate.)

For example: a particle on $\mathbb{R}^n$ in a potential V has $Q = \mathbb{R}^n$, $L(q, \dot{q}) = \frac{1}{2}m\dot{q}^2 - V(q)$. This system has

$$p_i \dot{q}^i = \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i = m\dot{q}^2 = 2K$$

so

$$H = p_i \dot{q}^i - L = 2K - (K - V) = K + V$$

as you'd have hoped.

2.3.2   Space Translation Symmetry

For a free particle in $\mathbb{R}^n$, we have $Q = \mathbb{R}^n$ and $L = K = \frac{1}{2}m\dot{q}^2$. This has **spatial translation** symmetries, so that for any $v \in \mathbb{R}^n$ we have the symmetry

$$q_s(t) = q(t) + s\,v$$

with

$$\delta L = 0$$

because $\delta\dot{q} = 0$ and L depends only on $\dot{q}$ not on q in this particular case. (Since L does not depend upon q^i we'll call q^i an **ignorable coordinate**; as above, these ignorables always give symmetries, hence conserved quantities. It is often useful therefore, to change coordinates so as to make some of them ignorable if possible!)

In this example we get a conserved quantity called **momentum in the v direction**:

$$p_i \delta q^i = m\dot{q}_i v^i = m\dot{q} \cdot v.$$

Aside: Note the subtle difference between two uses of the term "momentum"; here it is a conserved quantity derived from space translation invariance, but earlier it was a different thing, namely the momentum $\partial L/\partial \dot{q}^i = p_i$ conjugate to q^i. These two different "momentum's" happen to be the same in this example!

Since this is conserved for all v we say that $m\dot{q} \in \mathbb{R}^n$ is conserved. (In fact that whole Lie group $G = \mathbb{R}^n$ is acting as a translation symmetry group, and we're getting a $\mathfrak{q}(=\mathbb{R}^n)$-valued conserved quantity!)

2.3.3 Rotational Symmetry

The free particle in $\mathbb{R}^n$ also has rotation symmetry. Consider any $X \in \mathfrak{so}(n)$ (that is a skew-symmetric $n \times n$ matrix), then for all $s \in \mathbb{R}$ the matrix e^{sX} is in $SO(n)$, that is, it describes a rotation. This gives a 1-parameter family of symmetries

$$q_s(t) = e^{sX}q(t)$$

which has

$$\delta L = \frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial \dot{q}^i}\delta \dot{q}^i = m\dot{q}_i\delta \dot{q}^i.$$

Now q_i is ignorable, so $\partial L/\partial q^i = 0$, and $\partial L/\partial \dot{q}^i = p_i$, and

$$
\begin{aligned}
\delta \dot{q}^i &= \frac{d}{ds}\dot{q}_s^i\bigg|_{s=0} \\
&= \frac{d}{ds}\frac{d}{dt}\left(e^{sX}q\right)\bigg|_{s=0} \\
&= \frac{d}{dt}X\,q \\
&= X\,\dot{q}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\delta L &= m\dot{q}_i X_j^i \dot{q}^j \\
&= m\dot{q} \cdot (X\,\dot{q}) \\
&= 0
\end{aligned}
$$

since X is skew symmetric as stated previously ($X \in \mathfrak{so}(n)$). So we get a conserved quantity, the **angular momentum** in the X direction.

(Note: this whole bunch of math above for δL just says that the kinetic energy doesn't change when the velocity is *rotated*, without changing its magnitude.)

We write,

$$p_i\delta q^i = m\dot{q}_i \cdot (X\,q)^i$$

($\delta q^i = Xq$ just as $\delta \dot{q}^i = X\dot{q}$ in our previous calculation), or if X has zero entries except in ij and ji positions, where it's ± 1, then we get

$$m(\dot{q}_i q^j - \dot{q}_j q^i)$$

the "ij component of angular momentum". If $n = 3$ we can write these as

$$m\dot{\mathbf{q}} \times \mathbf{q}.$$

Note that above we have assumed one can construct a basis for $\mathfrak{so}(n)$ using matrices of the form assumed for X, i.e., skew symmetric with ± 1 in the respectively ij and ji elements, otherwise zero.

I mentioned earlier that we can do mechanics with any Lagrangian, but if we want to be useful we'd better pick a Lagrangian that actually describes a real system. But how do we do that? All this theory is fine but is useless unless you know how to apply it. The above examples were for a particularly simple system, a free particle, for which the Lagrangian is just the kinetic energy, since there is no potential energy variation for a free particle. We'd like to know how to solve more complicated dynamics.

The general idea is to guess the kinetic energy and potential energy of the particle (as functions of your generalized positions and velocities) and then let,

$$L = K - V$$

So we are not using Lagrangians directly to tell us what the fundamental physical laws should be, instead we plug in some assumed physics and use the Lagrangian approch to solve the system of equations. If we like, we can then compare our answers with experiments, which indirectly tells us something about the physical laws—but only provided the Lagrangian formulation of mechanics is itself a valid procedure in the first place.

Chapter 3

Examples

To see how Lagrangian mechanics and Noether's theorem works in practise, let's do some problems. The Lagrangian approach is often vastly superior to the simplistic $F = ma$ formulation of mechanics. The Lagrangian formulation allows the configuration space to be any manifold, and allows us to easily use any coordinates we wish.

3.1 The Atwood Machine

Consider a frictionless pulley with two masses, m_1 and m_2, hanging from it:

PSfrag replacements

x

m_1

$\ell - x$

m_2
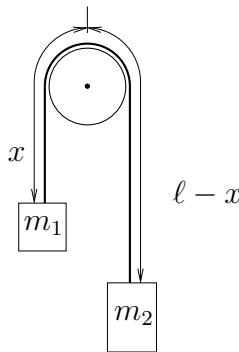
We have

$$K = \frac{1}{2}(m_1 + m_2)(\frac{d}{dt}(\ell - x))^2 = \frac{1}{2}(m_1 + m_2)\dot{x}^2$$
$$V = -m_1 gx - m_2 g(\ell - x)$$

so

$$L = K - V = \frac{1}{2}(m_1 + m_2)\dot{x}^2 + m_1 gx + m_2 g(\ell - x).$$

The configuration space is $Q = (0, \ell)$, and $x \in (0, \ell)$ (we could use the "owns" symbol $\ni$ here and write $Q = (0, \ell) \ni x$). Moreover $TQ = (0, \ell) \times \mathbb{R} \ni (x, \dot{x})$. As usual $L : TQ \to \mathbb{R}$. Note that solutions of the Euler–Lagrange equations will only be defined for *some* time $t \in \mathbb{R}$, as eventually the solutions reaches the "edge" of Q.

The momentum is:

$$p = \frac{\partial L}{\partial \dot{x}} = (m_1 + m_2)\dot{x}$$

and the force is:

$$F = \frac{\partial L}{\partial x} = (m_1 - m_2)g$$

The Euler–Lagrange equations say

$$\dot{p} = F$$
$$(m_1 + m_2)\ddot{x} = (m_1 - m_2)g$$
$$\ddot{x} = \frac{m_1 - m_2}{m_1 + m_2}g$$

So this is like a falling object in a downwards gravitational acceleration $a = \left(\frac{m_1 - m_2}{m_1 + m_2}\right)g$.

It is trivial to integrate the expression for $\ddot{x}$ twice (feeding in some appropriate initial conditions) to obtain the complete solution to the motion $x(t)$ and $\dot{x}(t)$. Note that $\ddot{x} = 0$ when $m_1 = m_2$, and $\ddot{x} = g$ if $m_2 = 0$.

3.2   Bead on a Rotating Rod

Now consider a bead of mass m sliding in a frictionless way on a rod rotating in a horizontal plane. The rod will go through the origin $(0, 0) \in \mathbb{R}^2$ and rotate at a constant angular velocity ω, so if the angle of the rod is $\theta(t)$ at time t, we may as well assume

$$\theta(t) = \omega t.$$

The bead's position on the rod will be given by a number $q(t) \in \mathbb{R}$ depending on time. What will the bead do?

Since the bead lies on a line, namely the rod, its configuration space is $Q = \mathbb{R}$, and its Lagrangian is a function $L : TQ \to \mathbb{R}$ where $TQ = \mathbb{R} \times \mathbb{R}$. Its position and velocity thus form a pair $(q(t), \dot{q}(t)) \in TQ$.

Since the rod lies in a horizontal plane, its gravitational potential energy is constant, and doesn't affect the Euler–Lagrange equations to assume this constant is zero. So, we take the bead's potential energy to be

$$V = 0.$$

Its kinetic energy is

$$K = \frac{1}{2}mv \cdot v$$

where v is its velocity. But what is its velocity? Its position in the plane at time t is

$$
\begin{aligned}
(x(t), y(t)) &= (r(t)\cos\theta(t), r(t)\sin\theta(t)) \\
&= (|q(t)|\cos(\omega t), |q(t)|\sin(\omega t)).
\end{aligned}
$$

The time derivative of $|q(t)|$ is $\pm q(t)$, with the plus sign if $q(t) > 0$ and the minus sign if $q(t) < 0$. If $q(t) = 0$ we seem to be in trouble, because the absolute value is not differentiable at zero, but we'll see a way around this later. Ignoring this case for now, the bead's velocity in the plane is thus

$$
\begin{aligned}
v(t) &= (\dot{x}(t), \dot{y}(t)) \\[2mm]
&= \frac{d}{dt}(|q(t)|\cos(\omega t), |q(t)|\sin(\omega t)) \\[2mm]
&= (\pm\dot{q}(t)\cos(\omega t) - |q(t)|\omega\sin(\omega t), \pm\dot{q}(t)\sin(\omega t) + |q(t)|\omega\cos(\omega t)).
\end{aligned}
$$

so we have

$$
\begin{aligned}
v(t) \cdot v(t) &= \dot{q}^2\cos^2(\omega t) \mp |q|\dot{q}\sin(\omega t)\cos(\omega(t)) + \omega^2 q^2\sin^2(\omega t) + \\
&\quad \dot{q}^2\sin^2(\omega t) \pm |q|\dot{q}\sin(\omega t)\cos(\omega(t)) + \omega^2 q^2\cos^2(\omega t) \\
&= \dot{q}^2 + \omega^2 q^2.
\end{aligned}
$$

Thus, the bead's kinetic energy is

$$K = \frac{1}{2}mv \cdot v = \dot{q}^2 + \omega^2 q^2.$$

This has a simple interpretation: the first term is the 'radial' part of the kinetic energy, while the second term is the 'angular' part.

The Lagrangian of the bead is

$$L(q, \dot{q}) = K - V = \dot{q}^2 + \omega^2 q^2.$$

Note that this is perfectly well-defined and smooth at $q = 0$. Thus, our problem at that point is easily dealt with: simply define the Lagrangian as above. A more careful analysis shows this is reasonable.

The force on the bead is

$$F = \frac{\partial L}{\partial q} = m\omega^2 q.$$

This is called **centrifugal force**, since it's caused by the rotating rod and tends to pull the bead out. The bead's momentum is

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

The Euler–Lagrange equation for the bead says

$$\frac{dp}{dt} = F$$

or

$$m\ddot{q}(t) = m\omega^2 q(t)$$

or simply

$$\ddot{q}(t) = \omega^2 q(t).$$

The mass of the bead does not affect its equation of motion. This equation is easy to solve, too:

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

Thus, the bead is likely to shoot off to infinity as $t \to +\infty$, due to the centrifugal force. If $\omega > 0$, the only exception is when $A = 0$: in this case the bead moves ever closer to the origin, or else just sits there if $A = B = 0$.

3.3 Disk Pulled by Falling Mass

Consider next a disk pulled across a table by a falling mass. The disk is free to move on a frictionless surface, and it can thus whirl around the hole to which it is tethered to the mass below.

PSfrag replacements

r m_1

$\ell - r$ — no swinging allowed!

m_2
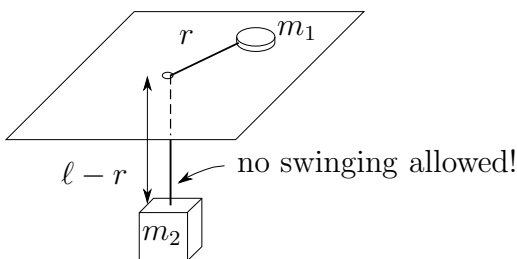
Here $Q = $ open disk of radius ℓ, minus its center
$$= (0, \ell) \times S^1 \ni (r, \theta)$$

$$TQ = (0, \ell) \times S^1 \times \mathbb{R} \times \mathbb{R} \ni (r, \theta, \dot{r}, \dot{\theta})$$

$$K = \frac{1}{2}m_1(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{1}{2}m_2(\frac{d}{dt}(\ell - r))^2$$
$$V = gm_2(r - \ell)$$
$$L = \frac{1}{2}m_1(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{1}{2}m_2\dot{r}^2 + gm_2(\ell - r)$$

having noted that ℓ is constant so $d/dt(\ell - r) = -\dot{r}$. For the momenta we get,

$$p_r = \frac{\partial L}{\partial \dot{r}} = (m_1 + m_2)\dot{r}$$

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = m_1 r^2 \dot{\theta}.$$

Note that θ is an "ignorable coordinate"—it doesn't appear in L—so there's a symmetry, rotational symmetry, and p_θ, the conjugate momentum, is conserved.

The forces are

$$F_r = \frac{\partial L}{\partial r} = m_1 r \dot{\theta}^2 - g m_2$$

$$F_\theta = \frac{\partial L}{\partial \theta} = 0 \quad (\theta \text{ is ignorable}).$$

Note: in F_r the term $m_1 r \dot{\theta}^2$ is recognizable as a centrifugal force, pushing m_1 radially *out*, while the term $-g m_2$ is gravity pulling m_2 down and thus pulling m_1 radially *in*.

So, the Euler–Lagrange equations give

$$\dot{p}_r = F_r, \quad (m_1 + m_2)\ddot{r} = m_1 r \dot{\theta}^2 - m_2 g$$

$$\dot{p}_\theta = 0, \quad p_\theta = m_1 r^2 \dot{\theta} = J = \text{ a constant.}$$

Let's use our conservation law here to eliminate $\dot{\theta}$ from the first equation:

$$\dot{\theta} = \frac{J}{m_1 r^2}$$

so

$$(m_1 + m_2)\ddot{r} = \frac{J^2}{m_1 r^3} - m_2 g$$

Thus effectively we have a particle on $(0, \ell)$ of mass $m = m_1 + m_2$ feeling a force

$$F_r = \frac{J^2}{m_1 r^3} - m_2 g$$

which could come from an "effective potential" $V(r)$ such that $dV/dr = -F_r$. So integrate $-F_r$ to find $V(r)$:

$$V(r) = \frac{J^2}{2 m_1 r^2} + m_2 g r$$

this is a sum of two terms that look like Fig. 3.1

If $\dot{\theta}(t = 0) = 0$ then there is no centrifugal force and the disk will be pulled into the hole until it gets stuck. At that time the disk reaches the hole, which is topologically the center of the disk that has been removed from Q, so then we've hit the boundary of Q and our solution is broken.

At $r = r_0$, the minimum of $V(r)$, our disc mass m_1 will be in a stable circular orbit of radius r_0 (which depends upon J). Otherwise we get orbits like Fig. 3.2.
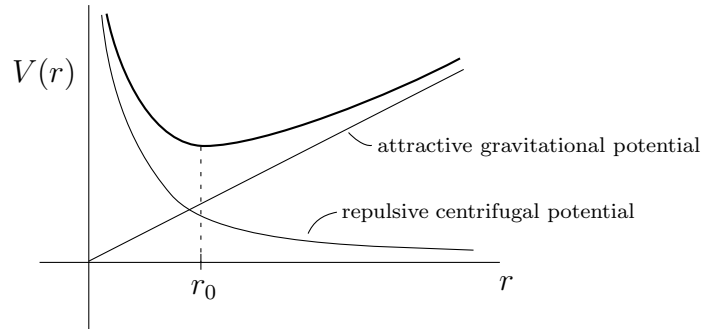
no swinging allowed!

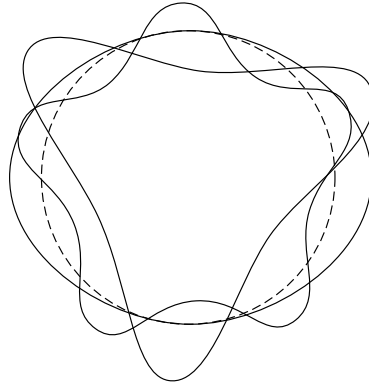Figure 3.1: Potential function for disk pulled by gravitating mass.

Figure 3.2: Orbits for the disc and gravitating mass system.

3.4 Free Particle in Special Relativity

In relativistic dynamics the parameter coordinate that parametrizes the particle's path in Minkowski spacetime need not be the "time coordinate", indeed in special relativity there are many allowed time coordinates.

Minkowski spacetime is,

$$\mathbb{R}^{n+1} \ni (x^0, x^1, \ldots, x^n)$$

if space is n-dimensional. We normally take x^0 as "time", and $(x^1, \ldots, x^n)$ as "space", but of course this is all relative to one's reference frame. Someone else traveling at some high velocity relative to us will have to make a Lorentz transformation to translate from our coordinates to theirs.

This has a Lorentzian metric

$$g(v, w) = v^0 w^0 - v^1 w^1 - \ldots - v^n w^n$$
$$= \eta_{\mu\nu} v^\mu w^\nu$$

where

$$
\eta_{\mu\nu} = \begin{pmatrix}
1 & 0 & 0 & \dots & 0 \\
0 & -1 & 0 & \dots & 0 \\
0 & 0 & -1 & & 0 \\
\vdots & \vdots & & \ddots & \vdots \\
0 & 0 & 0 & \dots & -1
\end{pmatrix}.
$$

In special relativity we take *spacetime* to be the configuration space of a single point particle, so we let Q be Minkowski spacetime, i.e., $\mathbb{R}^{n+1} \ni (x^0, \dots, x^n)$ with the metric $\eta_{\mu\nu}$ defined above. Then the path of the particle is,

$$
q : \mathbb{R}(\ni t) \longrightarrow Q
$$

where t is a completely arbitrary parameter for the path, not necessarily x^0, and not necessarily *proper time* either. We want some Lagrangian $L : TQ \to \mathbb{R}$, i.e., $L(q^i, \dot{q}^i)$ such that the Euler–Lagrange equations will dictate how our free particle moves at a constant velocity. Many Lagrangians do this, but the "best" one should give an action that is *independent* of the parameterization of the path—since the parameterization is "unphysical": it can't be measured. So the action

$$
S(q) = \int_{t_0}^{t_1} L\Big(q^i(t), \dot{q}^i(t)\Big)\, dt
$$

for $q \colon [t_0, t_1] \to Q$, should be independent of t. The obvious candidate for S is mass times arclength,

$$
S = m \int_{t_0}^{t_1} \sqrt{\eta_{ij}\dot{q}^i(t)\dot{q}^j(t)}\, dt
$$

or rather the Minkowski analogue of arclength, called **proper time**, at least when $\dot{q}$ is a timelike vector, i.e., $\eta_{ij}\dot{q}^i\dot{q}^j > 0$, which says $\dot{q}$ points into the future (or past) lightcone and makes S *real*, in fact it's then the time ticked off by a clock moving along the path $q : [t_0, t_1] \to Q$. By "obvious candidate" we are appealing somewhat to physical intuition and
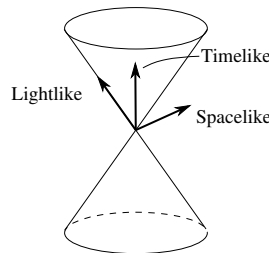
generalization. In Euclidean space, free particles follow straight paths, so the arclength or pathlength variation is an extremum, and we expect the same behavior in Minkowski

spacetime. Also, the arclength does not depend upon the parameterization, and lastly, the mass m merely provides the correct units for 'action'.

So let's take

$$L = m\sqrt{\eta_{ij}\dot{q}^i\dot{q}^j} \tag{3.1}$$

and work out the Euler–Lagrange equations. We have

$$\begin{aligned}
p_i = \frac{\partial L}{\partial \dot{q}^i} &= m\frac{\partial}{\partial \dot{q}^i}\sqrt{\eta_{ij}\dot{q}^i\dot{q}^j} \\
&= m\frac{2\eta_{ij}\dot{q}^j}{2\sqrt{\eta_{ij}\dot{q}^i\dot{q}^j}} \\
&= m\frac{\eta_{ij}\dot{q}^j}{\sqrt{\eta_{ij}\dot{q}^i\dot{q}^j}} = \frac{m\dot{q}_i}{\|\dot{q}\|}.
\end{aligned}$$

Now note that this p_i doesn't change when we change the parameter to accomplish $\dot{q} \mapsto \alpha\dot{q}$. The Euler–Lagrange equations say

$$\dot{p}_i = F_i = \frac{\partial L}{\partial q^i} = 0.$$

The meaning of this becomes clearer if we use "proper time" as our parameter (like parameterizing a curve by its arclength) so that

$$\int_{t_0}^{t_1} \|\dot{q}\|\,dt = t_1 - t_0, \quad \forall\, t_0, t_1$$

which fixes the parametrization up to an additive constant. This implies $\|\dot{q}\| = 1$, so that

$$p_i = m\frac{\dot{q}_i}{\|\dot{q}\|} = m\dot{q}_i$$

and the Euler–Lagrange equations say

$$\dot{p}_i = 0 \Rightarrow m\ddot{q}_i = 0$$

so our free particle moves unaccelerated along a straight line, as one would expect.

3.4.1   Comments

This Lagrangian from Eq.(3.1) has lots of symmetries coming from *reparameterizing the path*, so Noether's theorem yields lots of conserved quantities for the relativistic free particle. This is in fact called "the problem of time" in general relativity. Here we see it starting to show up in special relativity.

These reparameterization symmetries work as follows. Consider any (smooth) 1-parameter family of reparameterizations, i.e., diffeomorphisms

$$f_s \colon \mathbb{R} \longrightarrow \mathbb{R}$$

with $f_0 = 1_{\mathbb{R}}$. These act on the space of paths $\Gamma = \{q : \mathbb{R} \to Q\}$ as follows: given any $q \in \Gamma$ we get

$$q_s(t) = q\big(f_s(t)\big)$$

where we should note that q_s is *physically indistinguishable* from q. Let's show that

$$\delta L = \dot{\ell}, \quad \text{(when Euler–Lagrange eqns. hold)}$$

so that Noether's theorem gives a conserved quantity

$$p_i \delta q^i - \ell$$

Here we go then:

$$
\begin{aligned}
\delta L &= \frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial \dot{q}^i}\delta \dot{q}^i \\
&= p_i \delta \dot{q}^i \\
&= \frac{m\dot{q}_i}{\|\dot{q}\|} \frac{d}{ds}\dot{q}^i\big(f_s(t)\big)\Big|_{s=0} \\
&= \frac{m\dot{q}_i}{\|\dot{q}\|} \frac{d}{dt}\frac{d}{ds}q^i\big(f_s(t)\big)\Big|_{s=0} \\
&= \frac{m\dot{q}_i}{\|\dot{q}\|} \frac{d}{dt}\dot{q}^i\big(f_s(t)\big)\frac{f_s(t)}{ds}\Big|_{s=0} \\
&= \frac{m\dot{q}_i}{\|\dot{q}\|} \frac{d}{dt}\big(\dot{q}^i\delta f_s\big) \\
&= \frac{d}{dt}\big(p_i\dot{q}^i\delta f\big)
\end{aligned}
$$

where in the last step we used the Euler–Lagrange equations, i.e. $\frac{d}{dt}p_i = 0$, so $\delta L = \dot{\ell}$ with $\ell = p_i\dot{q}^i\delta f$.

So to recap a little: we saw the free relativistic particle has

$$L = m\|\dot{q}\| = m\sqrt{\eta_{ij}\dot{q}^i\dot{q}^j}$$

and we've considered reparameterization symmetries

$$q_s(t) = q\Big(f_s(t)\Big), \qquad f_s : \mathbb{R} \to \mathbb{R}$$

we've used the fact that

$$\delta q^i = \frac{d}{ds} q^i \Big(f_s(t) \Big)\Big|_{s=0} = \dot{q}^i \delta f$$

so (repeating a bit of the above)

$$\begin{aligned}
\delta L &= \frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial \dot{q}^i}\delta \dot{q}^i \\
&= p_i \delta \dot{q}^i, \quad (\text{since } \partial L/\partial q^i = 0, \text{ and } \partial L/\partial \dot{q}^i = p) \\
&= p_i \delta \dot{q}^i \\
&= p_i \frac{d}{dt}\delta q^i \\
&= p_i \frac{d}{dt}\dot{q}^i \delta f \\
&= \frac{d}{dt} p_i \dot{q}^i \delta f, \quad \text{and set } p_i \dot{q}^i \delta f = \ell
\end{aligned}$$

so Noether's theorem gives a conserved quantity

$$\begin{aligned}
p_i \delta q^i - \ell &= p_i \dot{q}^i \delta f - p_i \dot{q}^i \delta f \\
&= 0
\end{aligned}$$

So these conserved quantities *vanish*! In short, we're seeing an example of what physicists call **gauge symmetries**. This is a good topic for starting a new section.

3.5   Gauge Symmetries

What are gauge symmetries?

1. These are symmetries that permute different mathematical descriptions of the same physical situation—in this case reparameterizations of a path.

2. These symmetries make it impossible to compute $q(t)$ given $q(0)$ and $\dot{q}(0)$: since if $q(t)$ is a solution so is $q(f(t))$ for any reparameterization $f : \mathbb{R} \to \mathbb{R}$. We have a high degree of non-uniqueness of solutions to the Euler–Lagrange equations.

3. These symmetries give conserved quantities that work out to equal zero!

Note that (1) is a subjective criterion, (2) and (3) are objective, and (3) is easy to test, so we often use (3) to distinguish *gauge* symmetries from *physical* symmetries.

3.5.1 Relativistic Hamiltonian

What then is the Hamiltonian for special relativity theory? We're continuing here with the example problem of §3.4. Well, the Hamiltonian comes from Noether's theorem from *time translation* symmetry,

$$q_s(t) = q(t + s)$$

and this is an example of a reparametrization (with $\delta f = 1$), so we see from the previous results that the Hamiltonian is *zero*!

$$H = 0.$$

Explicitly, $H = p_i \delta \dot{q}^i - \ell$ where under $q(t) \to q(t + s)$ we have $\delta \dot{q}^i = \dot{q}^i \delta f$, and so $\delta L = d\ell/dt$, which implies $\ell = p_i \delta q^i$. The result $H = 0$ follows.

Now you know why people talk about "the problem of time" in general relativity theory, its glimmerings are seen in the flat Minkowski spacetime of special relativity. You may think it's nice and simple to have $H = 0$, but in fact it means that there is no temporal evolution possible! So we can't establish a dynamical theory on this footing! That's bad news. (Because it means you might have to solve the static equations for the 4D universe as a whole, and that's impossible!)

But there is another conserved quantity deserving the title of "energy" which is *not* zero, and it comes from the symmetry,

$$q_s(t) = q(t) + s\,w$$

where $w \in \mathbb{R}^{n+1}$ and w points in some timelike direction.

PSfrag replacements w
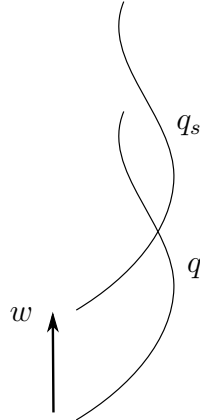
In fact *any* vector w gives a conserved quantity,

$$\delta L = \frac{\partial L}{\partial q^i} \delta q^i + \frac{\partial L}{\partial \dot{q}^i} \delta \dot{q}^i$$

$$= p_i \delta \dot{q}^i, \quad (\text{since } \partial L/\partial q^i = 0 \text{ and } \partial L/\partial \dot{q}^i = p_i)$$

$$= p_i 0 = 0$$

since $\delta q^i = w^i$, $\delta \dot{q}^i = \dot{w}^i = 0$. This is our $\dot{\ell}$ from Noether's theorem with $\ell = 0$, so Noether's theorem says that we get a conserved quantity

$$p_i \delta q^i - \ell = p_i w^i$$

namely, the **momentum in the w direction**. We know $\dot{p} = 0$ from the Euler–Lagrange equations, for our free particle, but here we see it coming from spacetime translation symmetry:

$$p = (p_0, p_1, \ldots, p_n)$$
$$p_0 \text{ is energy}, \quad (p_1, \ldots, p_n) \text{ is spatial momentum.}$$

We've just about exhausted all the basic stuff that we can learn from the free particle. So next we'll add some external force via an electromagnetic field.

3.6   Relativistic Particle in an Electromagnetic Field

The electromagnetic field is described by a 1-form A on spacetime, A is the **vector potential**, such that

$$dA = F \tag{3.2}$$

is a 2-form containing the electric and magnetic fields,

$$F_{ij} = \frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j} \tag{3.3}$$

We can write (for Q having local charts to $\mathbb{R}^{n+1}$),

$$A = A_0 dx^0 + A_1 dx^1 + \ldots A_n dx^n$$

and then because $d^2 = 0$

$$dA = dA_0 dx^0 + dA_1 dx^1 + \ldots dA_n dx^n$$

and since the A_j are just functions,

$$dA_j = \partial_i A_j dx^i$$

using the summation convention and $\partial_i := \partial/\partial x^i$. The reader can easily check that the components for $F = F_{01} dx^0 \wedge dx^1 + F_{02} dx^0 \wedge dx^2 + \ldots$, agrees with the matrix expression below (at least in 4 dimensions).

So, for example, in 4-dimensional spacetime

$$F = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & B_3 & -B_2 \\ -E_2 & -B_3 & 0 & B_1 \\ -E_3 & B_2 & -B_1 & 0 \end{pmatrix}$$

where E is the electric field and B is the magnetic field. The action for a particle of charge e is

$$S = m \int_{t_0}^{t_1} \|\dot{q}\| \, dt + e \int_q A$$

here

$$\int_{t_0}^{t_1} \|\dot{q}\| \, dt = \text{ proper time,}$$

$$\int_q A = \text{ integral of } A \text{ along the path } q.$$

Note that since A is a 1-form we can integrate it over an oriented manifold, but one can also write the path integral using time t as a parameter, with $A_i \dot{q}^i \, dt$ the differential, after $dq^i = \dot{q}^i dt$.

The Lagrangian in the above action, for a charge e with mass m in an electromagnetic potential A is

$$L(q, \dot{q}) = m\|\dot{q}\| + eA_i \dot{q}^i \tag{3.4}$$

so we can work out the Euler–Lagrange equations:

$$p_i = \frac{\partial L}{\partial \dot{q}^i} = m\frac{\dot{q}_i}{\|\dot{q}\|} + eA_i$$

$$= mv_i + e\,A_i$$

where $v \in \mathbb{R}^{n+1}$ is the velocity, normalized so that $\|v\| = 1$. Note that *now momentum is no longer mass times velocity*! That's because we're in $n + 1$-d spacetime, the momentum is an $n + 1$-vector. Continuing the analysis, we find the force

$$F_i = \frac{\partial L}{\partial q^i} = \frac{\partial}{\partial q^i}\left(e\,A_j \dot{q}^j\right)$$

$$= e\frac{\partial A_j}{\partial q^i}\dot{q}^j$$

So the Euler–Lagrange equations say (noting that $A_i = A_j\Big(q(t)\Big)$:

$$\dot{p} = F$$

$$\frac{d}{dt}\Big(mv_i + eA_i\Big) = e\frac{\partial A_j}{\partial q^i}\dot{q}^j$$

$$m\frac{dv_i}{dt} = e\frac{\partial A_j}{\partial q^i}\dot{q}^j - e\frac{dA_i}{dt}$$

$$m\frac{dv_i}{dt} = e\frac{\partial A_j}{\partial q^i}\dot{q}^j - e\frac{\partial A_i}{\partial q^j}\dot{q}^j$$

$$= e\left(\frac{\partial A_j}{\partial q^i} - \frac{\partial A_i}{\partial q^j}\right)\dot{q}^j$$

the term in parentheses is F_{ij} = the electromagnetic field, $F = dA$. So we get the following equations of motion

$$m\frac{dv_i}{dt} = eF_{ij}\dot{q}^j, \quad \text{(Lorentz force law)} \tag{3.5}$$

(Usually called the "Lorenz" force law.)

3.7   Lagrangian for a String

So we've looked at a point particle and tried

$$S = m \cdot (\text{arclength}) + \int A$$

or with 'proper time' instead of 'arclength', where the 1-from A can be integrated over a 1-dimensional path. A generalization (or specialization, depending on how you look at it) would be to consider a Lagrangian for an extended object.

In string theory we boost the dimension by +1 and consider a string tracing out a 2D surface as time passes (Fig. 3.3).
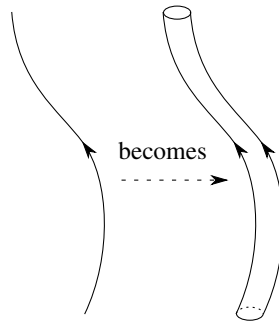

Figure 3.3: Worldtube of a closed string.

Can you infer an appropriate action for this system? Remember, the physical or physico-philosophical principle we've been using is that the path followed by physical objects minimizes the "activity" or "aliveness" of the system. Given that we presumably cannot tamper with the length of the closed string, then the worldtube quantity analogous to arclength or proper time would be the area of the worldtube (or worldsheet for an open string). If the string is also assumed to be a source of electromagnetic field then we need a 2-form to integrate over the 2D worldtube analogous to the 1-form integrated over the pathline of the point particle. In string theory this is usually the "Kalb-Ramond field",

call it B. To recover electrodynamic interactions it should be antisymmetric like A, but its tensor components will have two indices since it's a 2-form. The string action can then be written

$$S = \alpha \cdot (area) + e \int B \tag{3.6}$$

We've also replaced the point particle mass by the string tension α [mass·length^{-1}] to obtain the correct units for the action (since replacing arclength by area meant we had to compensate for the extra length dimension in the first term of the above string action).

This may still seem like we've pulled a rabbit out of a hat. But we haven't checked that this action yields sensible dynamics yet! But supposing it does, then would it justify our guesswork and intuition in arriving at Eq.(3.6)? Well by now you've probably realized that one can have more than one form of action or Lagrangian that yields the same dynamics. So provided we supply reasonabe physically realistic heuristics then whatever Lagrangian or action that we come up with will stand a good chance of describing some system with a healthy measure of physical verisimilitude.

That's enough about string for now. The point was to illustrate the type of reasoning that one can use in conjuring up a Lagrangian. It's particularly useful when Newtonian theory cannot give us a head start, i.e., in relativistic dynamics and in the physics of extended particles.

3.8 Another Lagrangian for Relativistic Electrodynamics

In § 3.6, Eq.(3.4) we saw an example of a Lagrangian for relativistic electrodynamics that had awkward reparametrization symmetries, meaning that $H = 0$ and there were non-unique solutions to the Euler–Lagrange equations arising from applying gauge transformations. This freedom to change the gauge can be avoided.

Recall Eq.(3.4), which was a Lagrangian for a charged particle with reparametrization symmetry

$$L = m\|\dot{q}\| + eA_i\dot{q}^i$$

just as for an uncharged relativistic particle. But there's another Lagrangian we can use that doesn't have this gauge symmetry:

$$L = \frac{1}{2}m\dot{q} \cdot \dot{q} + eA_i\dot{q}^i \tag{3.7}$$

This one even has some nice features.

- It looks formally like "$\frac{1}{2}mv^2$", familiar from nonrelativistic mechanics.

- There's no ugly square root, so it's everywhere differentiable, and there's no trouble with paths being timelike or spacelike in direction, they are handled the same.

What Euler–Lagrange equations does this Lagrangian yield?

$$p_i = \frac{\partial L}{\partial \dot{q}^i} = m\dot{q}_i + eA_i$$

$$F_i = \frac{\partial L}{\partial q^i} = e\frac{\partial A_j}{\partial q^i}\dot{q}^j$$

Very similar to before! The Euler–Lagrange equations then say

$$\frac{d}{dt}\left(m\dot{q}_i + eA_i\right) = e\frac{\partial A_j}{\partial q^i}\dot{q}^j$$

$$m\ddot{q}_i = e\mathsf{F}_{ij}\dot{q}^j$$

almost as before. (I've taken to using F here for the electromagnetic field tensor to avoid clashing with F for the generalized force.) The only difference is that we have $m\ddot{q}_i$ instead of $m\dot{v}_i$ where $v_i = \dot{q}_i/\|\dot{q}\|$. So the old Euler–Lagrange equations of motion reduce to the new ones if we pick a parametrization with $\|\dot{q}\| = 1$, which would be a parametrization by proper time for example.

Let's work out the Hamiltonian for this

$$L = \frac{1}{2}m\dot{q} \cdot \dot{q} + eA_i\dot{q}^i$$

for the relativistic charged particle in an electromagnetic field. Recall that for our reparametrization-invariant Lagrangian

$$L = m\sqrt{\dot{q}_i\dot{q}^i} + eA_iq^i$$

we got $H = 0$, time translation was a gauge symmetry. With the new Lagrangian it's not! Indeed

$$H = p_i\dot{q}^i - L$$

and now

$$p_i = \frac{\partial L}{\partial \dot{q}^i} = m\dot{q}^i + eA_i$$

so

$$H = (m\dot{q}_i + eA_i)\dot{q}^i - (\tfrac{1}{2}m\dot{q}_i\dot{q}^i + eA_i\dot{q}^i)$$
$$= \tfrac{1}{2}m\dot{q}_i\dot{q}^i$$

Comments. This is vaguely like how a nonrelativistic particle in a potential V has

$$H = p_i \dot{q}^i - L = 2K - (K - V) = K + V,$$

but now the "potential' $V = eA_i\dot{q}^i$ is *linear in velocity*, so now

$$H = p_i \dot{q}^i - L = (2K - V) - (K - V) = K.$$

As claimed H is not zero, and the fact that it's conserved says $\|\dot{q}(t)\|$ is constant as a function of t, so the particle's path is parameterized by proper time up to rescaling of t. That is, we're getting "conservation of speed" rather than some more familiar "conservation of energy". The reason is that this Hamiltonian comes from the symmetry

$$q_s(t) = q(t + s)$$

instead of spacetime translation symmetry

$$q_s(t) = q(t) + s\,w, \quad w \in \mathbb{R}^{n+1}$$

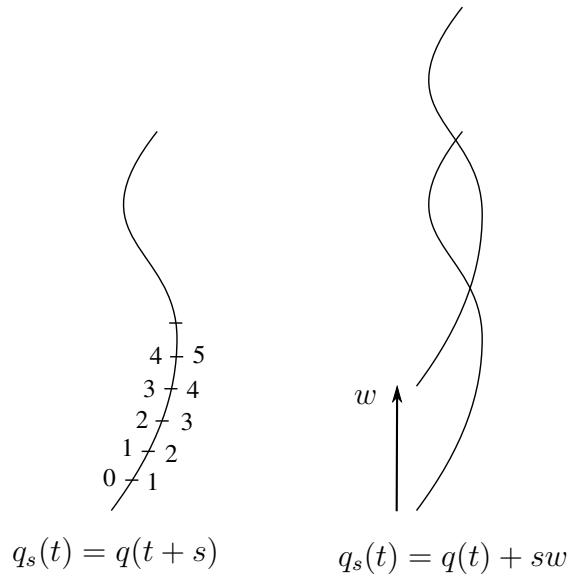the difference is illustrated schematically in Fig. 3.4.

Figure 3.4: Proper time rescaling vs spacetime translation.

Our Lagrangian

$$L(q, \dot{q}) = \tfrac{1}{2}m\|\dot{q}\|^2 + A_i(q)\dot{q}^i$$

has time translation symmetry iff A is translation invariant (but it's highly unlikely a given system of interest will have $A(q) = A(q + sw)$). In general then there's no conserved "energy" for our particle corresponding to translations in time.

3.9   The Free Particle in General Relativity

In general relativity, spacetime is an $(n + 1)$-dimensional Lorentzian manifold, namely a smooth $(n+1)$-dimensional manifold Q with a *Lorentzian metric g*. We define the metric as follows.

1. For each $x \in Q$, we have a bilinear map

$$g(x) : T_x Q \times T_x Q \longrightarrow \mathbb{R}$$
$$(v, w) \longmapsto g(x)(v, w)$$

or we could write $g(v, w)$ for short.

2. With respect to some basis of $T_x Q$ we have

$$g(v, w) = g_{ij} v^i w^j$$

$$g_{ij} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & -1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & -1 \end{pmatrix}$$

Of course we can write $g(v, w) = g_{ij} v^i w^j$ in any basis, but for different bases g_{ij} will have a different form.

3. $g(x)$ varies smoothly with x.

The Lagrangian for a free point particle in the spacetime Q is

$$L(q, \dot{q}) = m\sqrt{g(q)(\dot{q}, \dot{q})}$$
$$= m\sqrt{g_{ij} \dot{q}^i \dot{q}^j}$$

just like in special relativity but with η_{ij} replaced by g_{ij}. Alternatively we could just as well use

$$L(q, \dot{q}) = \tfrac{1}{2} m g(q)(\dot{q}, \dot{q})$$
$$= \tfrac{1}{2} m g_{ij} \dot{q}^i \dot{q}^j$$

The big difference between these two Lagrangians is that now spacetime translation symmetry (and rotation, and boost symmetry) is gone! So there is no conserved energy-momentum (nor angular momentum, nor velocity of center of energy) anymore!

Let's find the equations of motion. Suppose then Q is a Lorentzian manifold with metric g and $L : TQ \to \mathbb{R}$ is the Lagrangian of a free particle,

$$L(q, \dot{q}) = \tfrac{1}{2} m g_{ij} \dot{q}^i \dot{q}^j$$

We find equations of motion from the Euler–Lagrange equations, which in this case start from

$$p_i = \frac{\partial L}{\partial \dot{q}^i} = mg_{ij}\dot{q}^j$$

The velocity $\dot{q}$ here is a tangent vector, the momentum p is a cotangent vector, and we need the metric to relate them, via

$$g : T_q\mathcal{M} \times T_q\mathcal{M} \longrightarrow \mathbb{R}$$
$$(v, w) \longmapsto g(v, w)$$

which gives

$$T_q\mathcal{M} \longrightarrow T_q^*\mathcal{M}$$
$$v \longmapsto g(v, -).$$

In coordinates this would say that the tangent vector v^i gets mapped to the cotangent vector $g_{ij}v^j$. This is lurking behind the passage from $\dot{q}^i$ to the momentum $mg_{ij}v^j$.

Getting back to the Euler–Lagrange equations,

$$p_i = \frac{\partial L}{\partial \dot{q}^i} = mg_{ij}\dot{q}^j$$
$$F_i = \frac{\partial L}{\partial q^i} = \frac{\partial}{\partial q^i}\left(\tfrac{1}{2}mg_{jk}(q)\dot{q}^j\dot{q}^k\right)$$
$$= \tfrac{1}{2}m\partial_i g_{jk}\dot{q}^j\dot{q}^k, \quad (\text{where } \partial_i = \frac{\partial}{\partial q^i}).$$

So the Euler–Lagrange equations say

$$\frac{d}{dt}mg_{ij}\dot{q}^j = \tfrac{1}{2}m\partial_i g_{jk}\dot{q}^j\dot{q}^k.$$

The mass factors away, so the motion is independent of the mass! Essentially we have a geodesic equation.

We can rewrite this **geodesic equation** as follows

$$\frac{d}{dt}g_{ij}\dot{q}^j = \tfrac{1}{2}\partial_i g_{jk}\dot{q}^j\dot{q}^k$$
$$\therefore \quad \partial_k g_{ij}\dot{q}^k\dot{q}^j + g_{ij}\ddot{q}^j = \tfrac{1}{2}\partial_i g_{jk}\dot{q}^j\dot{q}^k$$
$$\therefore \quad g_{ij}\ddot{q}^j = \left(\tfrac{1}{2}\partial_i g_{jk} - \partial_k g_{ij}\right)\dot{q}^j\dot{q}^k$$
$$= \tfrac{1}{2}\left(\partial_i g_{jk} - \partial_k g_{ij} - \partial_j g_{ki}\right)\dot{q}^j\dot{q}^k$$

where the last line follows by symmetry of the metric, $g_{ik} = g_{ki}$. Now let,

$$\Gamma_{ijk} = -\left(\partial_i g_{jk} - \partial_k g_{ij} - \partial_j g_{ki}\right)$$

the minus sign being just a convention (so that we agree with everyone else). This defines what we call the Christoffel symbols Γ^i_{jk}. Then

$$\ddot{q}_i = g_{ij}\ddot{q}^j = -\Gamma_{ijk}\dot{q}^j\dot{q}^k$$
$$\therefore \quad \ddot{q}^i = -\Gamma^i_{jk}\dot{q}^j\dot{q}^k.$$

So we see that $\ddot{q}$ can be computed in terms of $\dot{q}$ and the Christoffel symbols Γ^i_{jk}, which is really a particular type of connection that a Lorentzian manifold has (the Levi-Civita connection), a "connection" is just the rule for parallel transporting tangent vectors around the manifold.

Parallel transport is just the simplest way to compare vectors at different points in the manifold. This allows us to define, among other things, a "covariant derivative".

3.10   A Charged Particle on a Curved Spacetime

We can now apply what we've learned in consideration of a charged particle, of charge e, in an electromagnetic field with potential A, in our Lorentzian manifold. The Lagrangian would be

$$L = \tfrac{1}{2}mg_{jk}\dot{q}^j\dot{q}^k + eA_i\dot{q}^i$$

which again was conjured up be replacing the flat space metric η_{ij} by the metric for GR g_{ij}. Not surprisingly, the Euler–Lagrange equations then yield the following equations of motion,

$$m\ddot{q}_i = -m\Gamma_{ijk}\dot{q}^j\dot{q}^k + e\mathsf{F}_{ij}\dot{q}^j.$$

If you want to know more about Lagrangians for general relativity we recommend the paper by Peldan [Pel94], and also the "black book" of Misner, Thorne & Wheeler [WTM71].

3.11   The Principle of Least Action and Geodesics

3.11.1   Jacobi and Least Time vs Least Action

We've mentioned that Fermat's principle of least time in optics is analogous to the principle of least action in particle mechanics. This analogy is strange, since in the principle of least action we fix the time interval $q : [0, 1] \to Q$. Also, if one imagines a force on a particle resulting from a potential gradient at an interface as analogous to light refraction then you also get a screw-up in the analogy (Fig. 3.5).

Nevertheless, Jacobi was able to reinterpret the mechanics of a particle as an optics problem and hence "unify" the two minimization principles. First, let's consider light in a medium with a varying index of refraction n (recall $1/n \propto$ speed of light). Suppose
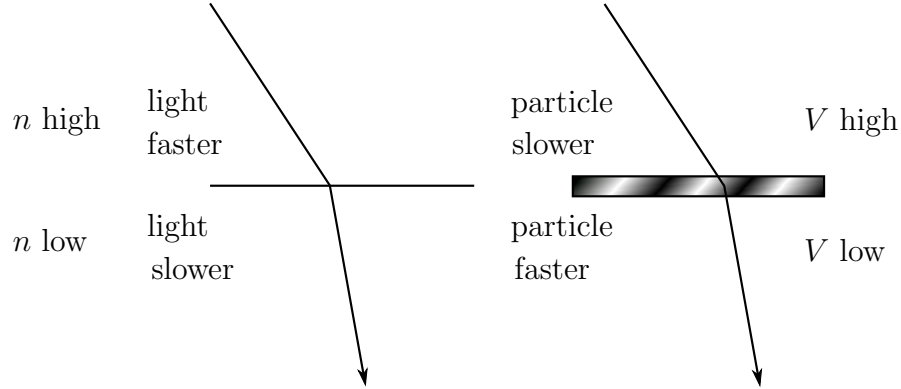
Figure 3.5: Least time versus least action.

it's in $\mathbb{R}^n$ with its usual Euclidean metric. If the light is trying to minimize the *time*, its trying to minimize the arclength of its path in the metric

$$g_{ij} = n^2 \delta_{ij}$$

that is, the index of refraction $n : \mathbb{R}^n \to (0, \infty)$, times the usual Euclidean metric

$$\delta_{ij} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

This is just like the free particle in general relativity (minimizing its proper time) except that now g_{ij} is a Riemannian metric

$$g(v, w) = g_{ij} v^i w^j$$
$$\text{where} \quad g(v, v) \geq 0$$

So we'll use the same Lagrangian:

$$L(q, \dot{q}) = \sqrt{g_{ij}(q) \dot{q}^i \dot{q}^j}$$

and get the same Euler–Lagrange equations:

$$\frac{d^2 q^i}{dt^2} + \Gamma^i_{jk} \dot{q}^j \dot{q}^k = 0 \tag{3.8}$$

if q is parameterized by arclength or more generally

$$\|\dot{q}\| = \sqrt{g_{ij}(q) \dot{q}^i \dot{q}^j} = \text{constant}.$$

As before the Christoffel symbols Γ are built from the derivatives of the metric g.

Now, what Jacobi did is show how the motion of a particle in a potential could be viewed as a special case of this. Consider a particle of mass m in Euclidean $\mathbb{R}^n$ with potential $V : \mathbb{R}^n \to \mathbb{R}$. It satisfies $F = ma$, i.e.,

$$m\frac{d^2q^i}{dt^2} = -\partial_i V \tag{3.9}$$

How did Jacobi see (3.9) as a special case of (3.8)? He considered a particle of energy E and he chose the index of refraction to be

$$n(q) = \sqrt{\frac{2}{m}\big(E - V(q)\big)}$$

which is just the **speed** of a particle of energy E when the potential energy is $V(q)$, since

$$\sqrt{\frac{2}{m}(E - V)} = \sqrt{\frac{2}{m}\frac{1}{2}m\|\dot{q}\|^2} = \|\dot{q}\|.$$

Note: this is precisely backwards compared to optics, where $n(q)$ is proportional to the *reciprocal* of the speed of light!! But let's see that it works.

$$\begin{aligned}
L &= \sqrt{g_{ij}(q)\dot{q}^i\dot{q}^j} \\
&= \sqrt{n^2(q)\dot{q}^i\dot{q}^j} \\
&= \sqrt{2/m(E - V(q))\dot{q}^2}
\end{aligned}$$

where $\dot{q}^2 = \dot{q} \cdot \dot{q}$ is just the usual Euclidean dot product, $v \cdot w = \delta_{ij}v^iw^j$. We get the Euler–Lagrange equations,

$$\begin{aligned}
p_i &= \frac{\partial L}{\partial \dot{q}^i} = \sqrt{\frac{2}{m}(E - V)} \cdot \frac{\dot{q}}{\|\dot{q}\|} \\
F_i &= \frac{\partial L}{\partial q^i} = \partial_i\sqrt{\frac{2}{m}(E - V(q))} \cdot \|\dot{q}\| \\
&= \frac{1}{2}\frac{-2/m\partial_i V}{\sqrt{2/m(E - Vq)}} \cdot \|\dot{q}\|
\end{aligned}$$

Then $\dot{p} = F$ says,

$$\frac{d}{dt}\sqrt{2/m(E - V(q))} \cdot \frac{\dot{q}_i}{\|\dot{q}\|} = -\frac{1}{m}\partial_i V\frac{\|\dot{q}\|}{\sqrt{2/m(E - V)}}$$

Jacobi noticed that this is just $F = ma$, or $m\ddot{q}_i = -\partial_i V$, that is, provided we reparameterize q so that,

$$\|\dot{q}\| = \sqrt{2/m(E - V(q))}.$$

Recall that our Lagrangian gives reparameterization invariant Euler–Lagrange equations! This is the unification between least time (from optics) and least action (from mechanics) that we sought.

3.11.2 The Ubiquity of Geodesic Motion

We've seen that many classical systems trace out paths that are geodesics, i.e., paths $q : [t_0, t_1] \to Q$ that are critical points of

$$S(q) = \int_{t_0}^{t_1} \sqrt{g_{ij} \dot{q}^i \dot{q}^j} \, dt$$

which is proper time when (Q, g) is a Lorentzian manifold, or arclength when (Q, g) is a Riemannian manifold. We have

1. The metric at $q \in Q$ is,

$$g(q) : T_q Q \times T_q Q \to \mathbb{R}$$
$$(v, w) \mapsto g(v, w)$$

   and it is bilinear.

2. w.r.t a basis of $T_q Q$

$$g(v, w) = \delta_{ij} v^i w^j$$

3. $g(q)$ varies smoothly with $q \in Q$.

An important distinction to keep in mind is that Lorentzian manifolds represent *space-times*, whereas Riemannian manifolds represent that we'd normally consider as just *space*.

We've seen at least three important things.

(1) In the geometric optics approximation, light in $Q = \mathbb{R}^n$ acts like particles tracing out geodesics in the metric

$$g_{ij} = n(q)^2 \delta_{ij}$$

where $n : Q \to (0, \infty)$ is the index of refraction function.

(2) Jacobi saw that a particle in $Q = \mathbb{R}^n$ in some potential $V : Q \to \mathbb{R}$ traces out geodesics in the metric

$$g_{ij} = \frac{2}{m}(E - V)\delta_{ij}$$

if the particle has energy E (where[1] $V < E$).

(3) A free particle in general relativity traces out a geodesic on a Lorentzian manifold (Q, q).

In fact all three of these results can be generalized to cover every problem that we've discussed!

---

[1]The case $V > E$, if they exist, would be classically forbidden regions.

(1′) Light on *any* Riemannian manifold (Q, q) with index of refraction $n : Q \to (0, \infty)$ traces out geodesics in the metric $h = n^2 g$.

(2′) A particle on a Riemannian manifold (Q, q) with potential $V : Q \to \mathbb{R}$ traces out geodesics w.r.t the metric

$$h = \frac{2}{m}(E - V)g$$

if it has energy E. Lots of physical systems can be described this way, e.g., the Atwood machine, a rigid rotating body $(Q = SO(3))$, spinning tops, and others. All of these systems have a Lagrangian which is a quadratic function of position, so they all fit into this framework.

(3′) **Kaluza-Klein Theory.** A particle with charge e on a Lorentzian manifold (Q, q) in an electromagnetic vector potential follows a path with

$$\ddot{q}_i = -\Gamma_{ijk}\dot{q}^j \dot{q}^k + \frac{e}{m}F_{ij}\dot{q}^j$$

where

$$F_{ij} = \partial_i A_j - \partial_j A_i$$

but this is actually geodesic motion on the manifold $Q \times U(1)$ where $U(1) = \{e^{i\theta} : \theta \in \mathbb{R}\}$ is a circle.

Let's examine this last result a bit further. To get the desired equations for motion on $Q \times U(1)$ we need to given $Q \times U(1)$ a cleverly designed metric built from g and A where the amount of "spiralling"—the velocity in the $U(1)$ direction is e/m. The metric h on

$Q \times U(1)$

PSfrag replacements

Q

a geodesic

the apparent path
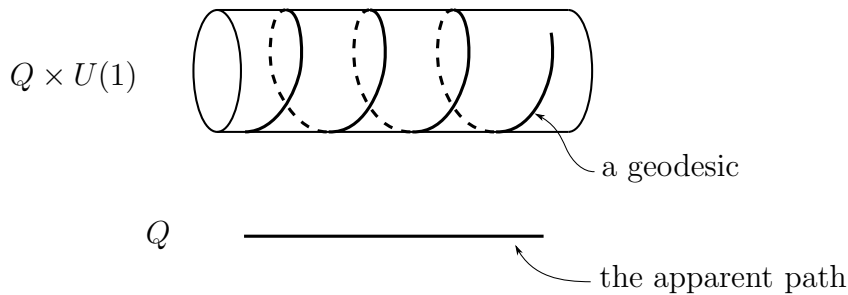
$Q \times U(1)$ is built from g and A in a very simple way. Let's pick coordinates x^i on Q where $i \in \{0, \ldots, n\}$ since we're in $n + 1$-dimensional spacetime, and θ is our local coordinate on S^1. The components of h are

$$h_{ij} = g_{ij} + A_i A_j$$
$$h_{\theta i} = h_{\theta i} = -A_i$$
$$h_{\theta\theta} = 1$$

Working out the equations for a geodesic in this metric we get

$$\ddot{q}_i = -\Gamma_{ijk}\dot{q}^j\dot{q}^k + \frac{e}{m}F_{ij}\dot{q}^j$$

$$\ddot{q}_\theta = 0,$$

if $\dot{q}_\theta = e/m$. Since F_{ij} is part of the Christoffel symbols for h.

To summarize this section on least time versus least action we can say that every problem that we've discussed in classical mechanics can be regarded as geodesic motion!

Chapter 4

From Lagrangians to Hamiltonians

In the Lagrangian approach we focus on the *position* and *velocity* of a particle, and compute what the particle does starting from the Lagrangian $L(q, \dot{q})$, which is a function

$$L : TQ \longrightarrow \mathbb{R}$$

where the **tangent bundle** is the space of position-velocity pairs. But we're led to consider momentum

$$p_i = \frac{\partial L}{\partial \dot{q}^i}$$

since the equations of motion tell us how it changes:

$$\frac{dp_i}{dt} = \frac{\partial L}{\partial q^i}.$$

4.1   The Hamiltonian Approach

In the Hamiltonian approach we focus on *position* and *momentum*, and compute what the particle does starting from the energy

$$H = p_i \dot{q}^i - L(q, \dot{q})$$

reinterpreted as a function of position and momentum We call this the *Hamiltonian*

$$H : T^*Q \longrightarrow \mathbb{R}$$

where the **cotangent bundle** is the space of position-momentum pairs. In this approach, position and momentum will satisfy **Hamilton's equations**:

$$\frac{dq^i}{dt} = \frac{\partial H}{\partial p_i}, \qquad \frac{dp^i}{dt} = -\frac{\partial H}{\partial q_i}$$

where the latter is the Euler–Lagrange equation

$$\frac{dp^i}{dt} = \frac{\partial L}{\partial q_i}$$

in disguise (it has a minus sign since $H = p\dot{q} - L$).

To obtain this Hamiltonian description of mechanics rigorously we need to study this map

$$\lambda : TQ \longrightarrow T^*Q$$
$$(q, \dot{q}) \longmapsto (q, p)$$

where $q \in Q$, and $\dot{q}$ is any tangent vector in $T_q Q$ (*not* the time derivative of something), and p is a cotangent vector in $T_q^*Q := (T_q Q)^*$, given by

$$\dot{q} \xrightarrow{\lambda} p_i = \frac{\partial L}{\partial \dot{q}^i}.$$

So λ is defined using $L : TQ \to \mathbb{R}$. Despite appearances, λ can be defined in a coordinate-free way, as follows (referring to Fig. 4.1). To understand "$\frac{\partial L}{\partial \dot{q}^i}$" in a coordinate-free way,
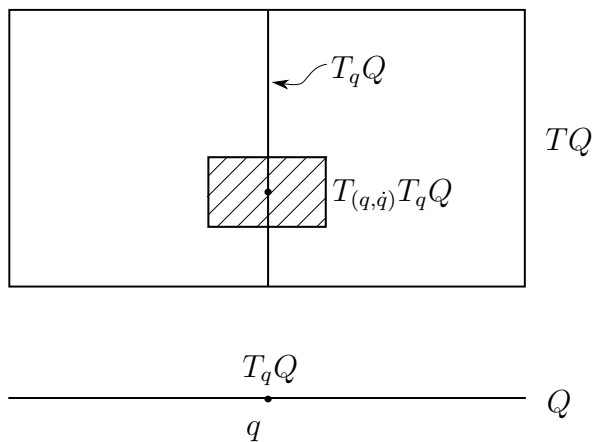

Figure 4.1:

we treat it as the "differential of L in the vertical direction"—i.e., the $\dot{q}^i$ directions. To do this, note that the differential of

$$\pi : TQ \longrightarrow Q$$
$$(q, \dot{q}) \longmapsto q$$

is a map

$$d\pi : TTQ \longrightarrow TQ.$$

We define the bundle of **vertical vectors** to be

$$VTQ = \{v \in TTQ : \ d\pi(v) = 0\} \subseteq TTQ.$$

That is, it consists of tangent vectors to TQ that are sent to zero by $d\pi$. Intuitively, they point 'straight up'. The differential of L at some point $(q, \dot{q}) \in TQ$ is a map from TTQ to $\mathbb{R}$, so we have

$$(dL)_{(q,\dot{q})} \in T^*_{(q,\dot{q})}TQ$$

that is,

$$dL_{(q,\dot{q})} : T_{(q,\dot{q})}TQ \longrightarrow \mathbb{R}.$$

We can restrict this to $VTQ \subseteq TTQ$, getting

$$f : V_{(q,\dot{q})}TQ \longrightarrow \mathbb{R}.$$

But note

$$V_{(q,\dot{q})}TQ \cong T_{q,\dot{q}}(T_qQ)$$

and since T_qQ is a vector space,

$$T_{(q,\dot{q})}T_qQ \cong T_qQ$$

in a canonical way[1]. So f gives a linear map

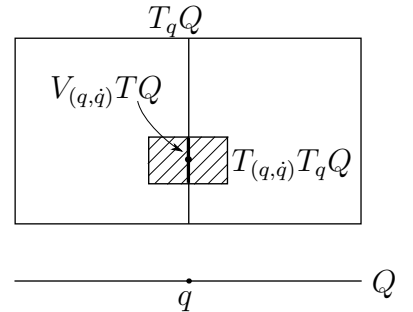$$p : T_qQ \longrightarrow \mathbb{R}$$

that is,

$$p \in T^*_qQ$$

this is the momentum!

Given $L \colon TQ \to T^*Q$, we now know a coordinate-free way of describing the map

$$\lambda \colon TQ \longrightarrow T^*Q$$
$$(q, \dot{q}) \longmapsto (q, p)$$

given in local coordinates by

$$p_i = \frac{\partial L}{\partial \dot{q}^i}.$$

---

[1] Each tangent space T_vV of a vector space V is isomorphic to V in a canonical way.

We say L is **regular** if λ is a diffeomorphism from TQ to some open subset $X \subseteq T^*Q$. In this case we can describe what our system is doing equally well by specifying position and velocity,

$$(q, \dot{q}) \in TQ$$

or position and momentum

$$(q, p) = \lambda(q, \dot{q}) \in X.$$

We call X the **phase space** of the system. In practice often $X = T^*Q$, then L is said to be **strongly regular**.

4.2   Regular and Strongly Regular Lagrangians

This section discusses some examples of the above theory.

4.2.1   Example: A Particle in a Riemannian Manifold with Potential $V(q)$

For a particle in a Riemannian manifold (Q, q) in a potential $V : Q \to \mathbb{R}$ has Lagrangian

$$L(q, \dot{q}) = \frac{1}{2} m g_{ij} \dot{q}^i \dot{q}^j - V(q)$$

Here

$$p_i = \frac{\partial L}{\partial \dot{q}^i} = m g_{ij} \dot{q}^j$$

so

$$\lambda(q, \dot{q}) = \big(q, mg(\dot{q}, -)\big)$$

L is strongly regular in this case because

$$T_q Q \longrightarrow T_q^* Q$$
$$v \longrightarrow g(v, -)$$

is 1-1 and onto, i.e., the metric is nondegenerate. Thus λ is a diffeomorphism, which in this case extends to all of T^*Q.

4.2.2   Example: General Relativistic Particle in an E-M Potential

For a general relativistic particle with charge e in an electromagnetic vector potential A the Lagrangian is

$$L(q, \dot{q}) = \frac{1}{2} m g_{ij} \dot{q}^i q^j - e A_i \dot{q}^i$$

and thus

$$p_i = \frac{\partial L}{\partial \dot{q}^i} = mg_{ij}\dot{q}^i q^j + eA_i.$$

This L is still strongly regular, but now each map

$$\lambda \mid_{T_q Q} : T_q Q \longrightarrow T_q^* Q$$
$$\dot{q} \longmapsto m\, g(\dot{q}, -) + eA(q)$$

is *affine* rather than linear[2].

4.2.3 Example: Free General Relativistic Particle with Reparameterization Invariance

The free general relativistic particle with reparameterization invariant Lagrangian has

$$L(q, \dot{q}) = m\sqrt{g_{ij}\dot{q}^i \dot{q}^j}.$$

This is terrible from the perspective of regularity—it's not differentiable when $g_{ij}\dot{q}^i\dot{q}^j$ vanishes, and undefined when the same is negative. Where it *is* defined

$$p_i = \frac{\partial L}{\partial \dot{q}^i} = \frac{mg_{ij}\dot{q}^j}{\|\dot{q}\|}$$

(where $\dot{q}$ is timelike), we can ask about regularity. Alas, the map λ is not 1-1 where defined since multiplying $\dot{q}$ by some number has *no effect* on p! (This is related to the reparameterization invariance—this always happens with reparameterization-invariant Lagrangians.)

4.2.4 Example: A Regular but not Strongly Regular Lagrangian

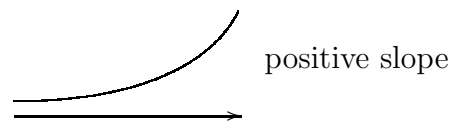Here's a Lagrangian that's regular but not strongly regular. Let $Q = \mathbb{R}$ and

$$L(q, \dot{q}) = f(\dot{q})$$

so that

$$p = \frac{\partial L}{\partial \dot{q}} = f'(\dot{q})$$

This will be regular but not strongly so if $f' : \mathbb{R} \to \mathbb{R}$ is a diffeomorphism from $\mathbb{R}$ to some proper subset $U \subset \mathbb{R}$. For example, take $f(\dot{q}) = e^{\dot{q}}$ so $f' : \mathbb{R} \xrightarrow{\sim} (0, \infty) \subset \mathbb{R}$. So
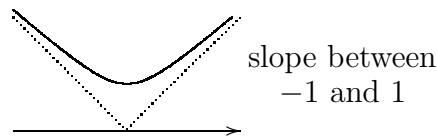
positive slope

$$L(q, \dot{q}) = e^{\dot{q}}$$

---

[2]All linear transforms are affine, but affine transformations include translations, which are nonlinear. In affine geometry there is no defined origin. For the example the translation is the "$+eA(q)$" part.

or

slope between
-1 and 1

$$L(q, \dot{q}) = \sqrt{1 + \dot{q}^2}$$

and so forth.

4.3   Hamilton's Equations

Now let's assume L is regular, so

$$\lambda \colon TQ \longrightarrow X \subseteq T^*Q$$
$$(q, \dot{q}) \longmapsto (q, p)$$

is a diffeomorphism. This lets us have the best of both worlds: we can identify TQ with X using λ. This lets us treat q^i, p^i, L, H, etc., all as functions on X (or TQ), thus writing

$$\dot{q}^i \qquad \text{(function on } TQ)$$

for the function

$$\dot{q}^i \circ \lambda^{-1} \qquad \text{(function on } X)$$

In particular

$$\dot{p}_i := \frac{\partial L}{\partial q^i} \qquad \text{(Euler–Lagrange equation)}$$

which is really a function on TQ, will be treated as a function on X. Now let's calculate:

$$dL = \frac{\partial L}{\partial q^i} dq^i + \frac{\partial L}{\partial \dot{q}^i} d\dot{q}^i$$
$$= \dot{p}_i dq^i + p_i d\dot{q}^i$$

while

$$dH = d(p_i \dot{q}^i - L)$$
$$= \dot{q}^i dp_i + p_i d\dot{q}^i - dL$$
$$= \dot{q}^i dp_i + p_i d\dot{q}^i - (\dot{p}_i dq^i + p_i d\dot{q}^i)$$
$$= \dot{q}^i dp_i - \dot{p}_i d\dot{q}^i$$

so

$$dH = \dot{q}^i dp_i - \dot{p}^i dq_i.$$

Assume the Lagrangian $L\colon TQ \to \mathbb{R}$ is regular, so

$$\lambda\colon TQ \overset{\sim}{\longrightarrow} X \subseteq T^*Q$$
$$(q, \dot{q}) \longmapsto (q, p)$$

is a diffeomorphism. This lets us regard both L and the Hamiltonian $H = p_i \dot{q}^i - L$ as functions on the phase space X, and use $(q^i, \dot{q}^i)$ as local coordinates on X. As we've seen, this gives us

$$dL = \dot{p}_i dq^i + p_i d\dot{q}^i$$
$$dH = \dot{q}^i dp_i - \dot{p}_i dq^i.$$

But we can also work out dH directly, this time using local coordinates (q^i, p_i), to get

$$dH = \frac{\partial H}{\partial p^i} dp_i + \frac{\partial H}{\partial q^i} dq^i.$$

Since dp_i, dq^i form a basis of 1-forms, we conclude:

$$\boxed{\dot{q}^i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}}$$

These are **Hamilton's Equations**.

4.3.1 Hamilton and Euler–Lagrange

Though $\dot{q}^i$ and $\dot{p}_i$ are just functions of X, when the Euler–Lagrange equations hold for some path $q\colon [t_0, t_1] \to Q$, they will be the time derivatives of q^i and p_i. So when the Euler–Lagrange equations hold, Hamilton's equations describe the motion of a point $x(t) = \big(q(t), p(t)\big) \in X$. In fact, in this context, Hamilton's equations are just the Euler–Lagrange equations in disguise. The equation

$$\dot{q}^i = \frac{\partial H}{\partial p_i}$$

really just lets us recover the velocity $\dot{q}$ as a function of q and p, inverting the formula

$$p_i = \frac{\partial L}{\partial \dot{q}^i}$$

which gave p as a function of q and $\dot{q}$. So we get a formula for the map

$$\lambda^{-1}\colon X \longrightarrow TQ$$
$$(q, p) \longmapsto (q, \dot{q}).$$

Given this, the other Hamilton equation

$$\dot{p}_i = -\frac{\partial H}{\partial q^i}$$

is secretly the Euler–Lagrange equation

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

or $\dot{p}_i = \frac{\partial L}{\partial q^i}$ for short. These are the same because

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

Example: Particle in a Potential $V(q)$

For a particle in $Q = \mathbb{R}^n$ in a potential $V : \mathbb{R}^n \to \mathbb{R}$ the system has Lagrangian

$$L(q, \dot{q}) = \frac{m}{2}\|\dot{q}\|^2 - V(q)$$

which gives

$$p = m\dot{q}$$

$$\dot{q} = \frac{p}{m}, \qquad \text{(though really that's } \dot{q} = \frac{g^{ij}p_j}{m})$$

and Hamiltonian

$$\begin{aligned} H(q, p) = p_i \dot{q}^i - L &= \frac{1}{m}\|p\|^2 - \left(\frac{\|p\|^2}{2m} - V\right) \\ &= \frac{1}{2m}\|p\|^2 + V(q). \end{aligned}$$

So Hamilton's equations say

$$\dot{q}^i = \frac{\partial H}{\partial p_i} \quad \Rightarrow \quad \dot{q} = \frac{p}{m}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q^i} \quad \Rightarrow \quad \dot{p} = -\nabla V$$

The first just recovers $\dot{q}$ as a function of p; the second is $F = ma$.

Note on Symplectic Structure

Hamilton's equations push us toward the viewpoint where p and q have equal status as coordinates on the phase space X. Soon, we'll drop the requirement that $X \subseteq T^*Q$ where Q is a configuration space. X will just be a manifold equipped with enough structure to write down Hamilton's equations starting from any $H \colon X \to \mathbb{R}$.

The coordinate-free description of this structure is the major 20th century contribution to mechanics: a symplectic structure.

This is important. You might have some particles moving on a manifold like S^3, which is not symplectic. So the Hamiltonian mechanics point of view says that the abstract manifold that you are really interested in is something different: it must be a symplectic manifold. That's the phase space X. We'll introduce symplectic geometry more completely in later chapters.

4.3.2 Hamilton's Equations from the Principle of Least Action

Before, we obtained the Euler–Lagrange equations by associating an "action" S with any $q \colon [t_0, t_1] \to Q$ and setting $\delta S = 0$. Now let's get Hamilton's equations *directly* by assigning an action S to any path $x \colon [t_0, t_1] \to X$ and setting $\delta S = 0$. Note: we don't impose any relation between p and q, $\dot{q}$! The relation will follow from $\delta S = 0$.

Let P be the space of paths in the phase space X and define the action

$$S \colon P \longrightarrow \mathbb{R}$$

by

$$S(x) = \int_{t_0}^{t_1} (p_i \dot{q}^i - H)dt$$

since $p_i \dot{q}^i - H$ is another way of writing the Lagrangian. More precisely, write our path x as $x(t) = \big(q(t), p(t)\big)$ and let

$$S(x) = \int_{t_0}^{t_1} \left[p_i(t) \frac{d}{dt} q^i(t) - H\big(q(t), p(t)\big) \right] dt$$

we write $\frac{d}{dt} q^i$ instead of $\dot{q}^i$ to emphasize that we mean the time derivative rather than a coordinate in phase space.

Let's show $\delta S = 0 \Leftrightarrow$ Hamilton's equations.

$$\delta S = \delta \int (p_i \dot{q}^i - H)dt$$

$$= \int \big(\delta p_i \dot{q}^i + p_i \delta \dot{q}^i - \delta H\big)dt$$

then integrating by parts,

$$= \int \left(\delta p_i \dot{q}^i - \dot{p}_i \delta q^i - \delta H \right) dt$$

$$= \int \left(\delta p_i \dot{q}^i - \dot{p}_i \delta q^i - \frac{\partial H}{\partial q^i} \delta q^i - \frac{\partial H}{\partial p_i} \delta p_i \right) dt$$

$$= \int \left(\delta p_i \left(\dot{q}^i - \frac{\partial H}{\partial p_i} \right) + \delta q^i \left(-\dot{p}_i - \frac{\partial H}{\partial q^i} \right) \right) dt$$

This vanishes $\forall \delta x = (\delta q, \delta p)$ if and only if Hamilton's equations

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \qquad p_i = -\frac{\partial H}{\partial q^i}$$

hold, just as we hoped.

We've seen two principles of "least action":

1. For paths in configuration space Q, $\delta S = 0 \Leftrightarrow$ Euler–Lagrange equations.

2. For paths in phase space X, $\delta S = 0 \Leftrightarrow$ Hamilton's equations.

Additionally, since $X \subseteq T^*Q$, we might consider a third version based on paths in position-velocity space TQ. But when our Lagrangian is regular we have a diffeomorphism $\lambda \colon TQ \xrightarrow{\sim} X$, so this third principle of least action is just a reformulation of principle 2. However, the *really* interesting principle of least action involves paths in the **extended phase space** where we have an additional coordinate for time: $X \times \mathbb{R}$.

Recall the action

$$S(x) = \int (p_i \dot{q}^i - H) \, dt$$

$$= \int p_i \frac{dq^i}{dt} dt - H \, dt$$

$$= \int p_i dq^i - H \, dt$$

We can interpet the integrand as a 1-form

$$\beta = p_i dq^i - H \, dt$$

on $X \times \mathbb{R}$, which has coordinates $\{p_i, q^i, t\}$. So any path

$$x \colon [t_0, t_1] \longrightarrow X$$

gives a path

$$\sigma \colon [t_0, t_1] \longrightarrow X \times \mathbb{R}$$
$$t \longmapsto (x(t), t)$$

and the action becomes the integral of a 1-form over a curve:

$$S(x) = \int p_i \, dq^i - H \, dt = \int_\sigma \beta$$

# 4.4 Waves versus Particles—The Hamilton-Jacobi Equations

In quantum mechanics we discover that every particle—electrons, photons, neutrinos, etc.—is a wave, and vice versa. Interestingly Newton already had a particle theory of light (his "corpuscules") and various physicists argued against it by pointing out that diffraction is best explained by a wave theory. We've talked about geometrized optics, an approximation in which light consists of particles moving along geodesics. Here we start with a Riemannian manifold (Q, g) as space, but we use the new metric

$$h_{ij} = n^2 g_{ij}$$

where $n \colon Q \to (0, \infty)$ is the index of refraction throughout space (generally not a constant).

4.4.1 Wave Equations

Huygens considered this same setup (in simpler language) and considered the motion of a wavefront:
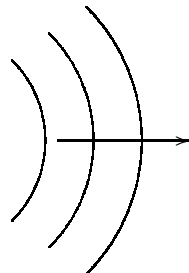
and saw that the wavefront is the envelope of a bunch of little wavelets centered at points along the big wavefront:

balls of radius ϵ
centered at points
of the old wavefront
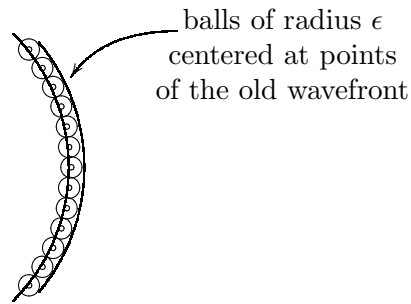
In short, the wavefront moves at unit speed in the normal direction with respect to the "optical metric" h. We can think about the distance function

$$d\colon Q \times Q \longrightarrow [0, \infty)$$

on the Riemannian manifold (Q, h), where

$$d(q_0, q_1) = \inf_{\Upsilon}(\text{arclength})$$

where $\Upsilon = \{\text{paths from } q_0 \text{ to } q_1\}$. (Secretly this $d(q_0, q_1)$ *is* the least action—the infimum of action over all paths from q_0 to q_1.) Using this we get the wavefronts centered at $q_0 \in Q$ as the level sets

$$\{q\colon d(q_0, q) = c\}$$

or at least for small $c > 0$, as depicted in Fig. 4.2. For larger c the level sets can cease
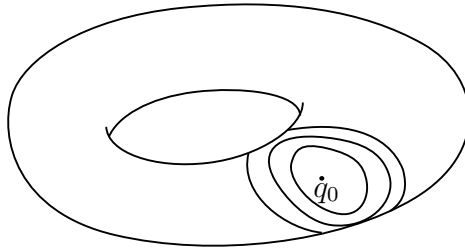

PSfrag replacements

Figure 4.2:

to be smooth—we say a **catastrophe** occurs—and then the wavefronts are no longer the level sets. This sort of situation can happen for topological reasons (as when the waves smash into each other in the back of Fig. 4.2) and it can also happen for geometrical
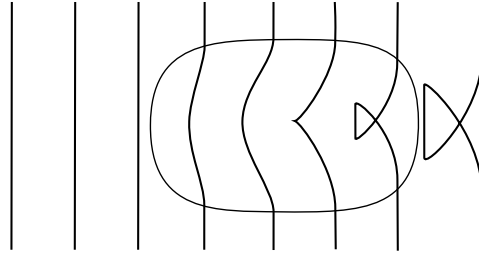
Figure 4.3:

reasons (Fig. 4.3). Assuming no such catastrophes occur, we can approximate the waves of light by a wavefunction:

$$\psi(q) = A(q)e^{ik\, d(q,q_0)}$$

where k is the wavenumber of the light (i.e., its color) and $A\colon Q \to \mathbb{R}$ describes the amplitude of the wave, which drops off far from q_0. This becomes the **eikonal approximation** in optics[3] once we figure out what A should be.

Hamilton and Jacobi focused on distance $d\colon Q \times Q \to [0, \infty)$ as a function of *two* variables. This is now called *Hamilton's principal function*, and we will denote it as W. They noticed that
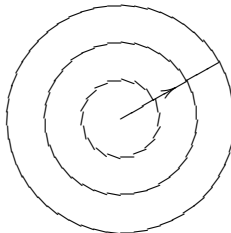
$$\frac{\partial}{\partial q_1^i} W(q_0, q_1) = (p_1)_i$$

where p_1 is a cotangent vector *pointing normal to the wavefronts.*

4.4.2 The Hamilton-Jacobi Equations

We've seen that in optics, particles of light move along geodesics, but wavefronts are level sets of the distance functions:

---

[3]Eikonal comes form the Greek word for 'image' or 'likeness', in optics the eikonal approximation is the basis for ray tracing methods.

at least while the level sets remain smooth. In the eikonal approximation, light is described by waves

$$\psi \colon Q \longrightarrow \mathbb{C}$$
$$\psi(q_1) = A(q_1)e^{ik\,W(q_0,q_1)}$$

where (Q, h) is a Riemannian manifold, h is the optical metric, $q_0 \in Q$ is the light source, k is the frequency and
$$W \colon Q \times Q \longrightarrow [0, \infty)$$
is the distance function on Q, or **Hamilton's principal function**:

$$W(q_0, q_1) = \inf_{q \in \Upsilon} S(q)$$

where Υ is the space of paths from q_0 to q and $S(q)$ is the action of the path q, i.e., its arclength. This is begging to be generalized to other Lagrangian systems! (At least it is retrospectively, with the advantage of our historical perspective.) We also saw that

$$\frac{\partial}{\partial q_1^i} W(q_0, q_1) = (p_1)_i,$$

"points normal to the wavefront"—really the tangent vector

$$p_1^i = h^{ij}(p_1)_j$$

points in this direction. In fact kp_{1i} is the *momentum* of the light passing through q_1. This foreshadows quantum mechanics! After all, in quantum mechanics, the momentum is an operator that acts to differentiating the wavefunction.

Jacobi generalized this to the motion of point particles in a potential $V \colon Q \to \mathbb{R}$, using the fact that a particle of energy E traces out geodesics in the metric

$$h_{ij} = \frac{2(E - V)}{m}\,g_{ij}.$$

We've seen this reduces point particle mechanics to optics—but only for particles of fixed energy E. Hamilton went further, and we now can go further still.

Suppose Q is any manifold and $L \colon TQ \to \mathbb{R}$ is any function (Lagrangian). Define Hamilton's principal function

$$W \colon Q \times \mathbb{R} \times Q \times \mathbb{R} \longrightarrow \mathbb{R}$$

by

$$W(q_0, t_0; q_1, t_1) = \inf_{q \in \Upsilon} S(q)$$

where

$$\Upsilon = \left\{ q \colon [t_0, t_1] \to Q,\ q(t_0) = q_0,\ \&\ q(t_1) = q_1 \right\}$$

and

$$S(q) = \int_{t_0}^{t_1} L\Big(q(t), \dot{q}(t) \Big)\, dt$$

Now W is just the *least action* for a path from (q_0, t_0) to (q_1, t_1); it'll be smooth if (q_0, t_0) and (q_1, t_1) are close enough—so let's assume that is true. In fact, we have

$$\frac{\partial}{\partial q_1^i} W(q_0, q_1) = (p_1)_i,$$

where p_1 is the momentum of the particle going from q_0 to q_1, at time t_1, and

$$\frac{\partial W}{\partial q_0^i} = -(p_0)_i, \qquad \text{(-momentum at time } t_0)$$

$$\frac{\partial W}{\partial t_1} = -H_1, \qquad \text{(-energy at time } t_1)$$

$$\frac{\partial W}{\partial t_0} = H_0, \qquad \text{(+energy at time } t_0)$$

($H_1 = H_0$ as energy is conserved). These last four equations are the **Hamilton–Jacobi equations**. The mysterious minus sign in front of energy was seen before in the 1-form,

$$\beta = p_i dq^i - H\, dt$$

on the extended phase space $X \times \mathbb{R}$. Maybe the best way to get the Hamilton–Jacobi equations is from this extended phase space formulation. But for now let's see how Hamilton's principal function W and variational principles involving least action also yield the Hamilton–Jacobi equations.
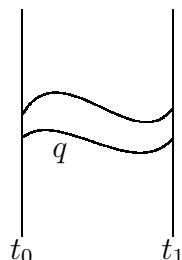
Given (q_0, t_0), (q_1, t_1), let

$$q \colon [t_0, t_1] \longrightarrow Q$$

be the action-minimizing path from q_0 to q_1. Then

$$W(q_0, t_0;\ q_1, t_1) = S(q)$$

Now consider varying q_0 and q_1 a bit

and thus vary the action-minimizing path, getting a variation δq which does not vanish at t_0 and t_1. We get

$$
\begin{aligned}
\delta W &= \delta S \\
&= \delta \int_{t_0}^{t_1} L(q, \dot{q})\, dt \\
&= \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial q^i} \delta q^i + \frac{\partial L}{\partial \dot{q}^i} \delta \dot{q}^i \right) dt \\
&= \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial q^i} \delta q^i - \dot{p}_i \delta q^i \right) dt + p_i \delta q^i \Big|_{t_0}^{t_1} \\
&= \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial q^i} - \dot{p}_i \right) \delta q^i\, dt
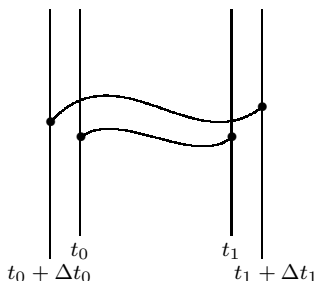\end{aligned}
$$

the term in parentheses is zero because q minimizes the action *and* the Euler–Lagrange equations hold. So we δq^i have

$$
\delta W = p_{1i} \delta q_1^i - p_{0i} \delta q_0^i
$$

and so

$$
\frac{\partial W}{\partial q_1^i} = p_{1i}, \quad \text{and} \quad \frac{\partial W}{\partial q_0^i} = -p_{0i}
$$

These are two of the four Hamilton–Jacobi equations! To get the other two, we need to vary t_0 and t_1:

Now change in W will involve Δt_0 and Δt_1

(you can imagine $\Delta t_0 < 0$ in this figure if you like).

We want to derive the Hamilton–Jacobi equations describing the derivatives of Hamilton's principal function

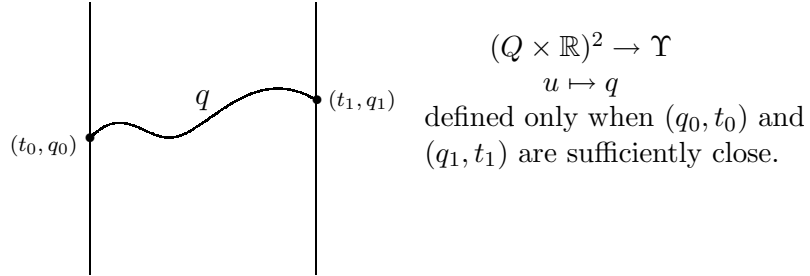$$W(q_0, t_0; q_1, t_1) = \inf_{q \in \Upsilon} S(q)$$

where Υ is the space of paths $q \colon [t_0, t_1] \to Q$ with $q(t_0) = q$, $q(t_1) = q_1$ and

$$S(q) = \int_{t_0}^{t_1} L(q, \dot{q}) \, dt$$

where the Lagrangian $L \colon TQ \to \mathbb{R}$ will now be assumed *regular*, so that

$$\lambda TQ \colon \longrightarrow X \subseteq T^*Q$$
$$(q, \dot{q}) \longmapsto (q, p)$$

is a diffeomorphism. We need to ensure that (q_0, t_0) is close enough to (q_i, t_1) that there is a unique $q \in \Upsilon$ that minimizes the action S, and assume that this q depends smoothly on $U = (q_0, t_0; q_1, t_1) \in (Q \times \mathbb{R})^2$. We'll think of q as a function of U:

$(Q \times \mathbb{R})^2 \to \Upsilon$
$u \mapsto q$
defined only when (q_0, t_0) and (q_1, t_1) are sufficiently close.

Then Hamilton's principal function is

$$W(u) := W(q_0, t_0; q_1, t_1) = S(q)$$
$$= \int_{t_0}^{t_1} L(q, \dot{q}) \, dt$$
$$= \int_{t_0}^{t_1} \Big(p\dot{q} - H(q, p) \Big) \, dt$$
$$= \int_{t_0}^{t_1} p \, dq - H \, dt$$
$$= \int_C \beta$$

where $\beta = pdq - H(q, p)dt$ is a 1-form on the extended phase space $X \times \mathbb{R}$, and C is a curve in the extended phase space:

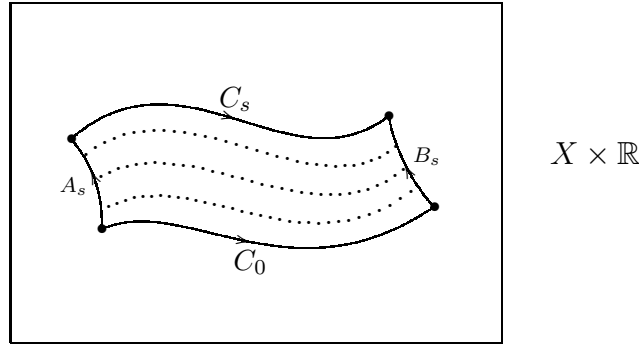$$C(t) = \Big(q(t), p(t), t \Big) \in X \times \mathbb{R}.$$

Note that C depends on the curve $q \in \Upsilon$, which in turn depends upon $u = (q_0, t_0; q_1, t_1) \in (Q \times \mathbb{R})^2$. We are after the derivatives of W that appear in the Hamilton–Jacobi relations, so let's differentiate

$$W(u) = \int_C \beta$$

with respect to u and get the Hamilton–Jacobi equations from β. Let u_s be a 1-parameter family of points in $(Q \times \mathbb{R})^2$ and work out

$$\frac{d}{ds} W(u_s) = \frac{d}{ds} \int_{C_s} \beta$$

where C_s depends on u_s as above

Let's compare

$$\int_{C_s} \beta \quad \text{and} \quad \int_{A_s + C_s + B_s} = \int_{A_s} \beta + \int_{C_s} \beta + \int_{B_s} \beta$$

Since C_0 minimizes the action among paths with the given end-points, and the curve $A_s + C_s + B_s$ has the same end-points, we get

$$\frac{d}{ds} \int_{A_s + C_s + B_s} \beta = 0$$

(although $A_s + C_s + B_s$ is not smooth, we can approximate it by a path that is smooth). So

$$\frac{d}{ds} \int_{C_s} \beta = \frac{d}{ds} \int_{B_s} \beta - \frac{d}{ds} \int_{A_s} \beta \quad \text{at } s = 0.$$

Note

$$\frac{d}{ds} \int_{A_s} \beta = \frac{d}{ds} \int \beta(A_r') \, dr$$
$$= \beta(A_0')$$

where $A_0' = v$ is the tangent vector of A_s at $s = 0$. Similarly,

$$\frac{d}{ds} \int_{B_s} \beta = \beta(w)$$

where $w = B_0'$. So,

$$\frac{d}{ds}W(u_s) = \beta(w) - \beta(v)$$

where w keeps track of the change of (q_1, p_1, t_1) as we move C_s and v keeps track of (q_0, p_0, t_0). Now since $\beta = p^i dq_i - H dt$, we get

$$\frac{\partial W}{\partial q_1^i} = p_1^i$$

$$\frac{\partial W}{\partial t_1} = -H$$

and similarly

$$\frac{\partial W}{\partial q_0^i} = -p_0^i$$

$$\frac{\partial W}{\partial t_0} = H$$

So, if we define a wavefunction:

$$\psi(q_0, t_0; q_1, t_1) = e^{iW(q_0, t_0; q_1, t_1)/\hbar}$$

then we get

$$\frac{\partial \psi}{\partial t_1} = -\frac{i}{\hbar}H_1\psi$$

$$\frac{\partial \psi}{\partial q_1^i} = \frac{i}{\hbar}p_1\psi$$

At the time of Hamilton and Jacobi's research this would have been new... but nowadays it is thoroughly familiar from *quantum mechanics*!

Bibliography

[Pel94]   Peter Peldan. Actions for gravity, with generalizations: A review. *Classical and Quantum Gravity*, 11:1087, 1994.

[WTM71] J. A. Wheeler, K. S. Thorne, and C. W. Misner. *Gravitation*. W. H. Freeman, New York, 1971.