Variational Calculus

José Figueroa-O'Farrill
⟨j.m.figueroa@ed.ac.uk⟩

Version of 20th September 2018

Preface

This is the second draft of the lecture notes for **Variational Calculus**, a course offered in the School of Mathematics of the University of Edinburgh.

Although nominally a Mathematical Physics course, the bulk of the material is on the *calculus of variations*, which finds wide applicability in many branches of Mathematics, both pure and applied. Our preferred context for the calculus of variations is Hamilton's principle of least action, which promotes newtonian mechanics to the level of a mathematical principle and provides a new language (that of action functionals) for describing physical models. We will however also discuss other contexts taken from other branches of Mathematics, such as geodesics and minimal surfaces in differential geometry and the principle of maximum entropy in probability and statistics.

One advantage of the variational approach is that it provides a very transparent link between symmetries and conservation laws via the celebrated Noether's Theorem. In the context of hamiltonian dynamics, which we will discuss briefly in this course, the conserved quantities derived from symmetries themselves generate the symmetries on the phase space of the system.

There are some exercises of varying difficulty included in the notes and typeset in this way.

Scholia are typeset in this way. They contain background and/or explanatory material whose inclusion in the bulk of the text might otherwise interrupt the flow of the narrative.

Caveats are typeset in this way.

Notation will be introduced as we go. Term which are being defined for the first time appear in **bold sans-serif** type and are listed in the index at the end of the document.

Contents

1 Introduction

> Le Calcul des Variations est, pour les opérations fonctionnelles, ce que le Calcul différentiel est pour les fonctions.[a]
>
> — Jacques Hadamard

[a] The calculus of variations is to functionals what the differential calculus is to functions.

The calculus of variations gives us precise analytical techniques to answer questions of the following type:

- Find the shortest path between two given points on a surface.
- Find the curve between two given points in the plane that yields a surface of revolution of minimum area when revolved around a given axis.
- Find the curve along which a bead will slide (under the effect of gravity) in the shortest time.

It also underpins much of modern mathematical physics, via Hamilton's principle of least action. This is one of several important variational principles, such as Fermat's principle of least time (of use in geometric optics) and Jaynes's principle of maximum entropy (of use in statistics). Variational principles can be used both to generate interesting differential equations, and also to prove the existence of solutions, even when these cannot be found analytically. A recent example is the unexpected discovery of figure-eight periodic solution to the three-body problem (see, e.g., [1]).

The calculus of variations is concerned — although not uniquely — with the problem of finding extrema of *functionals*. This problem is a generalisation of the problem of finding extrema of functions of several variables. In a sense to be made more precise below, it is the problem of finding extrema of functions of an *infinite* number of variables. In fact, these variables will turn out to be themselves functions and we will be finding extrema of "functions of functions" or **functionals**.

This generalisation is actually quite straight-forward, provided we understand the finite-dimensional case, which we will review below. Before that, though, we start with a motivating example.

2 A motivating example: geodesics

> In this enormous volume, the author never succeeds in proving that the shortest distance between two points is a straight line.
> — GH Hardy

As a motivating example, let us consider the problem of finding the shortest path between two points P and Q, say, in the plane. It is well-known that the answer is the straight line joining these two points, but let us see how we can go about deriving this.

2.1 Setting up the problem

By a **(regular) path** between P and Q we mean a continuously differentiable curve (a C^1 **curve** for short)

$$x : [0,1] \to \mathbb{R}^2 \qquad t \mapsto (x^1(t), x^2(t))$$

with the condition that $x(0) = \mathrm{P}$ and $x(1) = \mathrm{Q}$ and such that $\dot{x}$ is nowhere zero. The arclength of such a path is obtained by integrating the norm of the velocity vector (i.e., the **speed**) along the path:

$$\mathrm{S}[x] = \int_0^1 \|\dot{x}(t)\| dt \,,$$

where

$$\|\dot{x}(t)\| = \sqrt{(\dot{x}^1(t))^2 + (\dot{x}^2(t))^2} \,.$$

Here and in what follows we shall use Newton's *fluxion* notation for derivatives: $\dot{x} := \frac{dx}{dt}$.

Let $C_{\mathrm{P,Q}}$ denote the space of paths between P and Q. Then the arclength functional defines a function $\mathrm{S} : C_{\mathrm{P,Q}} \to \mathbb{R}$. Finding the shortest path (i.e., the **geodesic**) between P and Q is equivalent to finding the minimum of S on $C_{\mathrm{P,Q}}$. In this case, it is geometrically obvious that there is a unique extremum, which is also a minimum, but this is not necessarily the case for other functionals.

By analogy with the case of several variable calculus, we may attempt to solve this problem by differentiating S, determining its critical points and then checking whether those critical points are extrema. This is indeed our aim, but before doing that, let's see why this problem can be interpreted as that of finding extrema of a function of an infinite number of variables.

Let us follow Euler's *method of finite differences*. We break up the interval $[0,1]$ into N equally spaced segments with endpoints $t_n = \frac{n}{\mathrm{N}}$, where $n = 0, 1, \ldots, \mathrm{N}$. Let $x_n = x(t_n) \in \mathbb{R}^2$. Of course, $x_0 = \mathrm{P}$ and $x_{\mathrm{N}} = \mathrm{Q}$ are fixed. The curve $x : [0,1] \to \mathbb{R}^2$ can be approximated by the piecewise linear curve consisting of the straight line segments from x_{n-1} to x_n, for $n = 1, \ldots, \mathrm{N}$, whose arclength is the function

$$\mathrm{S}_{\mathrm{N}}(x_1, \ldots, x_{\mathrm{N}-1}) = \sum_{n=1}^{\mathrm{N}} \|x_n - x_{n-1}\|$$

of the $2(N-1)$ variables $x_1, \dots, x_{N-1}$. Finding minima of S_N is thus a problem in several variable calculus: we find the critical points of S_N and then check whether they are minima. Clearly, as $N \to \infty$ the piecewise linear curve approximates the original curve better and better, and in this sense we can think of the variational problem as the problem of finding the minima of a function of an infinite number of variables: namely, the x_n for $n \geq 1$.

Euler's way is not the way we go about finding extrema of functionals today. In fact, Euler himself abandoned this approach for the more general *method of variations* due to Lagrange.

2.2 Review: finding extrema of functions of several variables

Let $U \subset \mathbb{R}^n$ be an open subset and let $f : U \to \mathbb{R}$ be a differentiable function (of n variables). Recall that a point $x_0 \in U$ is a **critical point** of the function f if $Df(x_0) = 0$, where $Df(x_0)$ is the derivative of f at x_0.

Recall that the derivative $Df(x_0)$ of f at x_0 is a gadget to which we feed a vector ζ and gives the directional derivative $Df(x_0)\zeta$ of f along ζ at x_0:

$$Df(x_0)\zeta = \frac{d}{ds} f(x_0 + s\zeta)\Big|_{s=0} .$$

Thus the derivative $Df(x_0)$ of f at x_0 is a linear map which sends the vector ζ to the number $Df(x_0)\zeta$.

☞ We leave it as an instructive exercise to the reader to convince herself that this definition of the derivative agrees with the (unfortunately) more common one:

$$Df(x_0)\zeta = \left\langle \nabla f|_{x_0}, \zeta \right\rangle ,$$

where ∇f is the gradient of the function f. This definition actually uses the inner product on $\mathbb{R}^n$, which is something we may not have at our disposal in some of the function spaces we deal with in the calculus of variations. The definition of the derivative given in terms of curves does generalise.

This definition of the directional derivative has a very natural geometric interpretation which I cannot resist to elaborate upon. The function $\gamma : (-\varepsilon, \varepsilon) \to U$, sending s to $\gamma(s) = x_0 + s\zeta$ is a curve in U through x_0. Its velocity is ζ. The image $f(x_0 + s\zeta)$ by f of this curve is a curve in $\mathbb{R}$ through $f(x_0)$ with velocity $Df(x_0)\zeta$ at $f(x_0)$:

$$f(x_0 + s\zeta) = f(x_0) + sDf(x_0)\zeta + o(s) .$$

Velocities of curves through $x_0 \in U$ define the **tangent space** $T_{x_0}U$ to U at x_0. It is an n-dimensional real vector space and $\zeta \in T_{x_0}U$. Then the derivative defines a linear map

$$Df(x_0) : T_{x_0}U \to \mathbb{R} ,$$

whence it defines an element of the *dual vector space* $\left(T_{x_0}U\right)^*$ to $T_{x_0}U$, which is called the **cotangent space** of U at x_0 and is denoted $T^*_{x_0}U$. It is again an n-dimensional real vector space.

Exercise 2.1. *Show that the tangent space $T_{x_0}\mathbb{R}^n$ is indeed an n-dimensional vector space and show, in case you have not seen this before, that the dual space of a real n-dimensional vector space is again a real n-dimensional vector space.*

A point $x_0 \in U$ is a **critical point** of f if $Df(x_0) = 0$; that is, if $Df(x_0)\zeta = 0$ for all $\zeta \in T_{x_0}U$. In other words, this condition is equivalent to

$$\boxed{Df(x_0)\zeta := \frac{d}{ds} f(x_0 + s\zeta)\Big|_{s=0} = 0 \qquad \forall \zeta \in T_{x_0}U} \,. \tag{1}$$

There are three main ingredients in this equation: the point $x_0 \in U \subset \mathbb{R}^n$, the function f defined on U and the tangent space $T_{x_0}U$ to which ζ belongs. We will need to identify these ingredients in order to find the critical points of the arclength functional.

2.3 The critical points of the arclength functional

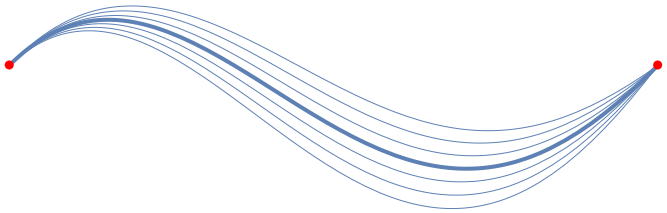
It is clear that the rôle of U is played by the space $\mathcal{C}_{P,Q}$ of paths between P and Q, which is a subset of the (infinite-dimensional, real) vector space of continuously differentiable functions $[0,1] \to \mathbb{R}^2$.

Although all n-dimensional real vector spaces are isomorphic, this is not the case for infinite-dimensional real vector spaces. (This is why several variable calculus is so much easier than functional analysis.) The spaces of functions we deal with in the calculus of variations are infinite dimensional and hence there is some choice as to which function space to take. Usually the problem determines a natural space. In the case of the arclength functional, since we wish to integrate the velocity, it makes sense to take paths which are continuously differentiable (i.e., C^1 paths), but it is possible to work with a more general class of paths with less stringent differentiability requirements. After much soul searching I have decided not to work in that level of generality, at least for now. Those interested in this more analytic treatment can consult [2].

Of course, the rôle of f is played by the arclength functional $S : \mathcal{C}_{P,Q} \to \mathbb{R}$. It remains to identify the tangent space to $\mathcal{C}_{P,Q}$ at a given path x. In our case, they are differences of C^1 curves $x(t)$ and $y(t)$ from P to Q. Let $\varepsilon(t) = y(t) - x(t)$ be one such difference of curves. Then $\varepsilon : [0,1] \to \mathbb{R}^2$ is itself a C^1 function, but now we have that $\varepsilon(0) = \varepsilon(1) = 0 \in \mathbb{R}^2$. Such a ε is called an **(endpoint-fixed) variation**, hence the name of the theory.

Strictly speaking, for every fixed t, $\varepsilon(t) \in T_{x(t)}\mathbb{R}^2$; that is, it is a tangent vector to $\mathbb{R}^2$ at $x(t)$. Moreover the endpoint conditions are $\varepsilon(0) = 0 \in T_P\mathbb{R}^2$ and $\varepsilon(1) = 0 \in T_Q\mathbb{R}^2$. However, we can (and will) identify all the tangent spaces with $\mathbb{R}^2$ by translating them to the origin in $\mathbb{R}^2$ and this is why we have written ε as a map $\varepsilon : [0,1] \to \mathbb{R}^2$ with $\varepsilon(0) = \varepsilon(1) = 0$.

In general, variations are not really differences of curves. They are tangent vectors to the space of curves; that is, they are velocities to a "curve of curves". The idea is to consider families x_s of curves, where s is a small real parameter which labels the curves in the family. In the picture below we show a few of these curves, highlighting the original curve $s = 0$.



We can think of this family as a (twice differentiable) function $x(s,t)$ of two real variables, where for each fixed s, we have a regular curve. We have that $x(0,t)$ is the original curve we are varying, $x(s,0) = \mathrm{P}$ and $x(s,1) = \mathrm{Q}$ for all s. The variation $\varepsilon : [0,1] \to \mathbb{R}^2$ is simply $\varepsilon(t) = \frac{\partial x(s,t)}{\partial s}\Big|_{s=0}$.
The fact that variations are not difference of curves is completely analogous to the fact that tangent vectors (velocities of curves) are not necessarily differences of points. In a vector space, if v is a tangent vector at a point x_0, we can always think of it as the velocity of the curve $x(s) = x_0 + sv$. In this case, $v = x(1) - x(0)$ is a difference of points. The reason we can do this is because we can add points in a vector space. If we were, say, on the sphere, we could not think of a tangent vector to the sphere as the difference between two points on the sphere. We would have to consider a curve $x(s)$ on the sphere and then $v = \dot{x}(0)$.

The condition for a path x being a critical point of the arclength functional S is now given by a formula analogous to (1):

$$\boxed{\frac{d}{ds}\mathrm{S}[x + s\varepsilon]\Big|_{s=0} = 0 \qquad \text{for all endpoint-fixed variations } \varepsilon.}$$

As we now show, this condition translates into a differential equation for the path x. Notice that

$$\begin{aligned}\mathrm{S}[x + s\varepsilon] &= \int_0^1 \|\dot{x}(t) + s\dot{\varepsilon}(t)\| dt \\ &= \int_0^1 \langle \dot{x}(t) + s\dot{\varepsilon}(t), \dot{x}(t) + s\dot{\varepsilon}(t)\rangle^{1/2} dt\,,\end{aligned}$$

whence differentiating inside the integral sign,

$$\begin{aligned}\frac{d}{ds}\mathrm{S}[x + s\varepsilon] &= \int_0^1 \frac{d}{ds}\langle \dot{x} + s\dot{\varepsilon}, \dot{x} + s\dot{\varepsilon}\rangle^{1/2} dt \\ &= \int_0^1 \frac{\langle \dot{x} + s\dot{\varepsilon}, \dot{\varepsilon}\rangle}{\|\dot{x} + s\dot{\varepsilon}\|} dt\,.\end{aligned}$$

Evaluating at $s = 0$, we find

$$\begin{aligned}\frac{d}{ds}\mathrm{S}[x+s\varepsilon]\Big|_{s=0} &= \int_0^1 \frac{\langle \dot{x}, \dot{\varepsilon}\rangle}{\|\dot{x}\|}dt \\ &= \int_0^1 \left\langle \frac{\dot{x}}{\|\dot{x}\|}, \dot{\varepsilon}\right\rangle dt\,.\end{aligned}$$

Integrating by parts and using that $\varepsilon(0) = \varepsilon(1) = 0$, we find that

$$\frac{d}{ds}\mathrm{S}[x+s\varepsilon]\Big|_{s=0} = -\int_0^1 \left\langle \frac{d}{dt}\left(\frac{\dot{x}}{\|\dot{x}\|}\right), \varepsilon\right\rangle dt\,. \tag{2}$$

Therefore a path x is a critical point of the arclength functional S if and only if

$$\int_0^1 \left\langle \frac{d}{dt}\left(\frac{\dot{x}}{\|\dot{x}\|}\right), \varepsilon\right\rangle dt = 0 \tag{3}$$

for all variations ε. We will prove in the next section that this actually implies that

$$\frac{d}{dt}\left(\frac{\dot{x}}{\|\dot{x}\|}\right) = 0\,, \tag{4}$$

which says that the velocity vector $\dot{x}$ has constant direction; i.e., that it is a straight line. There is only one straight line joining P and Q and it is clear from the geometry that this path actually minimises arclength.

Exercise 2.2. *Generalise the preceding discussion to paths in $\mathbb{R}^n$ between any two distinct points.*

Exercise 2.3. *What is the shortest path in the plane from the origin to the line $x = 1$? Solve this problem using the variational calculus, but notice that variations are not necessarily fixed at one endpoint!*

2.4 The fundamental lemma of the calculus of variations

It remains to prove an easy result from analysis which was used above to go from equation (3) to equation (4). This result is fundamental to the calculus of variations, hence the name. There are many versions of this result, so we will state the strongest version we shall need.

Theorem 2.1 (Fundamental Lemma of the Calculus of Variations)**.** *Let $f : [0,1] \to \mathbb{R}^n$ be a continuous function which obeys*

$$\int_0^1 \langle f(t), h(t)\rangle\, dt = 0$$

for all C^∞ functions $h : [0,1] \to \mathbb{R}^n$ with $h(0) = h(1) = 0$. Then $f \equiv 0$.

Proof. Let us first of all prove the case of $n = 1$. Let $f : [0,1] \to \mathbb{R}$ be a continuous function which obeys

$$\int_0^1 f(t)h(t)dt = 0$$

for all C^∞ functions $h : [0,1] \to \mathbb{R}$ with $h(0) = h(1) = 0$. Then we will prove that $f \equiv 0$. Assume for a contradiction that there is a point $t_0 \in (0,1)$ for which $f(t_0) \neq 0$. We will assume without loss of generality that $f(t_0) > 0$, since we can otherwise apply the argument to $-f$ if $f(t_0) < 0$. Because f is continuous, there is some interval $(a,b) \subset (0,1)$ containing t_0 and some $c > 0$ such that $f(t) > c$ for all $t \in (a,b)$.

We will now construct a C^∞ non-negative function $h : [0,1] \to \mathbb{R}$ with the following properties:

(P1) $h(t) = 0$ for all t outside the interval (a,b) (so in particular $h(0) = h(1) = 0$); and

(P2) $\int_0^1 h(t)dt = \int_a^b h(t)dt > 0$.

Postponing for a moment the construction of such a function, let us see how its existence allows us to prove the Lemma. Let us estimate the integral

$$\begin{aligned}
\int_0^1 f(t)h(t)dt &= \int_a^b f(t)h(t)dt && \text{using (P1)} \\
&> c\int_a^b h(t)dt && \text{since } f > c \text{ on } (a,b) \text{ and } h \geq 0 \\
&> 0 && \text{using (P2).}
\end{aligned}$$

This violates the hypothesis of the Lemma, hence we deduce that there is no point $t_0 \in (0,1)$ for which $f(t_0) \neq 0$, whence by continuity, $f(t) = 0$ for all $t \in [0,1]$.

Now let's go back to the case of general n. Let $f : [0,1] \to \mathbb{R}^n$ be a continuous function which obeys

$$\int_0^1 \langle f(t), h(t) \rangle \, dt = 0$$

for all C^∞ functions $h : [0,1] \to \mathbb{R}^n$ with $h(0) = h(1) = 0$. Suppose that there is some $t_0 \in (0,1)$ with $f(t_0) \neq 0$. Then at least one component of the vector $f(t_0)$ is different from zero. For definiteness of exposition, suppose that it is the first component: $f^1(t_0) \neq 0$. By the result above for $n = 1$ there is some smooth function $h^1 : [0,1] \to \mathbb{R}$ with $h^1(0) = h^1(1) = 0$ such that

$$\int_0^1 f^1(t)h^1(t)dt \neq 0.$$

Let $h : [0,1] \to \mathbb{R}^n$ be the function $h(t) = (h^1(t), 0, \ldots, 0)$. It is smooth and $h(0) = h(1) = 0$ and moreover

$$\int_0^1 \langle f(t), h(t) \rangle \, dt = \int_0^1 f^1(t)h^1(t)dt \neq 0.$$

This shows that f is identically zero on $(0,1)$ and by continuity also on $[0,1]$. □

We now come to the construction of the function h in the above proof. Consider the function $\theta : \mathbb{R} \to \mathbb{R}$ defined by

$$\theta(t) = \begin{cases} e^{-1/t} & t > 0 \\ 0 & t \leq 0 \,. \end{cases}$$

This function is clearly smooth (i.e., infinitely differentiable) at every point except, perhaps, at $t = 0$. However it is a standard exercise to prove that θ is smooth there as well.

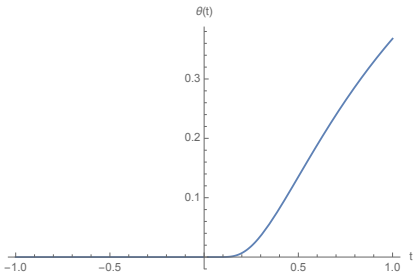


Figure 1: Graph of $\theta(t)$

Exercise 2.4. *Prove that θ so defined is a smooth function.*

Now define the function $\varphi : \mathbb{R} \to \mathbb{R}$ by

$$\varphi(t) = \theta(t)\theta(1-t) \,.$$

Being the product of two smooth functions, it is clearly smooth. Moreover, it vanishes outside the interval $(0, 1)$.

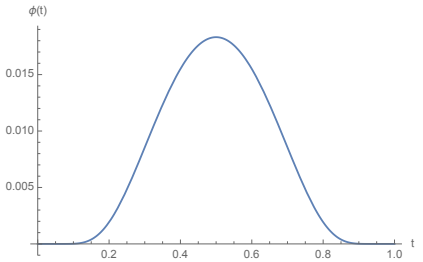


Figure 2: Graph of $\varphi(t)$

By rescaling t, we can make a function $\varphi_{a,b}$ which vanishes outside any chosen interval (a, b):

$$\varphi_{a,b}(t) = \varphi\left(\frac{t-a}{b-a}\right) \,. \tag{5}$$

Furthermore, it is clear from the graph of $\varphi(t)$ that

$$\int_a^b \varphi_{a,b}(t)dt = (b-a)\int_0^1 \varphi(t)dt > 0 \,.$$

The function h in the proof above can be taken to be the restriction to $[0, 1]$ of $\varphi_{a,b}$.

3 The Euler–Lagrange equations

> Cette équation est celle qu'Euler a trouvée le premier.[a]
>
> — Joseph-Louis Lagrange

[a] This equation is the one that Euler was the first to find.

In this section we will discuss the general formalism of the variational calculus and show how to derive a differential equation, known as the Euler–Lagrange equation, for the extremal function.

3.1 Endpoint-fixed variations

As before, let $C_{P,Q}$ be the space of C^1 curves $x : [0,1] \to \mathbb{R}^n$ with $x(0) = P$ and $x(1) = Q$. Let $L : \mathbb{R}^{2n+1} \to \mathbb{R}$, sending $(x, v, t) \mapsto L(x, v, t)$, where $x, v \in \mathbb{R}^n$ and $t \in \mathbb{R}$, be a sufficiently differentiable function (typically smooth in applications) and let us consider the functional $I : C_{P,Q} \to \mathbb{R}$ defined by

$$I[x] = \int_0^1 L(x(t), \dot{x}(t), t)\, dt\ .$$

☞ Please notice that whereas $L : \mathbb{R}^{2n+1} \to \mathbb{R}$ is defined over all $\mathbb{R}^{2n+1}$ or perhaps some open subset $U \subset \mathbb{R}^{2n+1}$, in computing $I[x]$ we are evaluating L on the image of the map $[0,1] \to \mathbb{R}^{2n+1}$ defined by x which sends $t \mapsto (x(t), \dot{x}(t), t)$. In other words, the integrand in the above equation is the composition of two functions: the function $t \mapsto (x(t), \dot{x}(t), t)$ defined by x and the function L.

The function L is called the **lagrangian** and the functional I is called the **action**. Extremising the action will yield a differential equation for x, as we now show. Recall that a path x is a critical point for the action if, for all endpoint-fixed variations ε, we have

$$\frac{d}{ds} I[x + s\varepsilon]\Big|_{s=0} = 0\ .$$

Differentiating under the integral sign, we find

$$\begin{aligned} 0 &= \int_0^1 \frac{d}{ds} L(x + s\varepsilon, \dot{x} + s\dot{\varepsilon}, t)\Big|_{s=0} dt \\ &= \int_0^1 \left(\sum_{i=1}^n \frac{\partial L}{\partial x^i} \varepsilon^i + \sum_{i=1}^n \frac{\partial L}{\partial \dot{x}^i} \dot{\varepsilon}^i \right) dt \\ &= \int_0^1 \sum_{i=1}^n \left(\frac{\partial L}{\partial x^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}^i} \right) \varepsilon^i dt\ , \end{aligned}$$

where in the second line we have used the chain rule and where we have introduced the notation

$$\frac{\partial L}{\partial \dot{x}^i} := \frac{\partial L}{\partial v^i}\Big|_{v=\dot{x}}\ .$$

Finally, in the last line we have integrated by parts and used that $\varepsilon(0) = \varepsilon(1) = 0$. Using the Fundamental Lemma, this is equivalent to

$$\boxed{\frac{\partial \mathrm{L}}{\partial x^i} = \frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{x}^i}}\,, \tag{6}$$

for all $i = 1, 2, \dots, n$. These are the **Euler–Lagrange equations**.

As an example, let us reconsider the lagrangian $\mathrm{L}(x, v, t) = \|v\|$. Then

$$\frac{\partial \mathrm{L}}{\partial x^i} = 0 \quad \text{and} \quad \frac{\partial \mathrm{L}}{\partial \dot{x}^i} = \frac{\dot{x}^i}{\|\dot{x}\|}\,,$$

and the Euler–Lagrange equation simply says that $\frac{\dot{x}^i}{\|\dot{x}\|}$ is constant, as we saw above.

Let us make two remarks:

- If $\frac{\partial \mathrm{L}}{\partial x^i} = 0$ for some i, then the **momentum** $\frac{\partial \mathrm{L}}{\partial \dot{x}^i}$ conjugate to the variable x^i is constant.

  ☞ Notice that the condition $\frac{\partial \mathrm{L}}{\partial x^i} = 0$ is nothing else but the statement that L is invariant under translations $x \mapsto x + a$, where $a \in \mathbb{R}^n$ is a constant. The Euler–Lagrange equations in turn say that the quantity $\frac{\partial \mathrm{L}}{\partial \dot{x}^i} = \frac{\dot{x}^i}{\|\dot{x}\|}$ is constant in time; that is, it is *conserved* along the geodesic: it does not change as we move along the geodesic. This is an instance of *Noether's Theorem*, to which we will return and which is one of the cornerstones of this course. This theorem says that every one-parameter group of symmetries of a functional gives rise to a quantity which is conserved along extremal trajectories.

- If $\frac{\partial \mathrm{L}}{\partial t} = 0$, so that the lagrangian does not depend *explicitly* on t, then the **energy**

  $$\mathrm{E} := \sum_{i=1}^{n} \dot{x}^i \frac{\partial \mathrm{L}}{\partial \dot{x}^i} - \mathrm{L} \tag{7}$$

  is constant along curves obeying the Euler–Lagrange equations. Indeed, let us differentiate the function $\mathrm{L}(x(t), \dot{x}(t), t)$ with respect to t and use that $\frac{\partial \mathrm{L}}{\partial t} = 0$ to obtain

  $$\begin{aligned}\frac{d}{dt}\mathrm{L}(x(t), \dot{x}(t), t) &= \sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial x^i}\dot{x}^i + \sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial \dot{x}^i}\ddot{x}^i \\ &= \sum_{i=1}^{n} \left(\frac{d}{dt}\left(\frac{\partial \mathrm{L}}{\partial \dot{x}^i}\right)\dot{x}^i + \frac{\partial \mathrm{L}}{\partial \dot{x}^i}\ddot{x}^i \right) \\ &= \frac{d}{dt}\left(\sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial \dot{x}^i}\dot{x}^i \right),\end{aligned}$$

  where we have used the Euler–Lagrange equations. Equation (7) is known as **Beltrami's identity**.

Exercise 3.1. *Let $x : [0,1] \to \mathbb{R}$ be a C^3 function. Let the lagrangian L depend also on the second derivative $\ddot{x}$. Derive the Euler–Lagrange equation arising from extremising the action*

$$S[x] = \int_0^1 L(x, \dot{x}, \ddot{x}, t)dt \, .$$

Generalise this to lagrangians depending on the first k derivatives of x, which should now be a C^{k+1} function. Generalise this further to lagrangians depending on the first k derivatives of $x : [0,1] \to \mathbb{R}^n$.
(Hint: For lagrangians depending on the first k derivatives of x, the variations and their first $k-1$ derivatives must vanish at the endpoints.)

3.2 General variations

Now we discuss variational problems where the variations need not be endpoint fixed. This is a large subject, but in this course I will restrict myself to the simplest situations.

As a motivating example, let us consider the problem of finding the shortest path between a point P and a curve C in $\mathbb{R}^2$. For simplicity we will assume that the curve C is described by an equation $g(x) = 0$, where $g : \mathbb{R}^2 \to \mathbb{R}$ is a differentiable function.

We seek minima of the arclength functional

$$S[x] = \int_0^1 \|\dot{x}(t)\|dt \, ,$$

defined for regular curves $x : [0,1] \to \mathbb{R}^2$ with boundary conditions $x(0) = P$ and $x(1) \in C$ or, equivalently, $g(x(1)) = 0$. Again we tackle this by first finding critical points of the functional; that is, admissible curves x such that

$$\frac{d}{ds} S[x + s\varepsilon]\Big|_{s=0} = 0 \, , \tag{8}$$

for all "variations" ε. But what are the admissible variations in this problem?

As usual, $x + s\varepsilon$ should be an admissible curve, at least for small values of s. Alas, this is not usually possible. There is no problem at $t = 0$ since $x(0) = P$, $x(0) + s\varepsilon(0) = P$ just says that $\varepsilon(0) = 0$. However at $t = 1$ we require only that $x(1) + s\varepsilon(1) \in C$ for s small enough. The problem with this is that $x(1) + s\varepsilon(1)$, for s small, is a straight line segment and there is no reason to expect that C contains a straight line segment. This means that we need consider more general variations.

We will consider a one-parameter (s) family of curves $x_s(t)$ defined for s small, where $x_s(0) = P$ and $x_s(1) \in C$ for small s. Moreover we assume that the function $x_s(t)$ of s and t is twice continuously differentiable in both s and t. Let us define

$$\varepsilon(t) := \frac{\partial}{\partial s} x_s(t)\Big|_{s=0} \, .$$

What does ε satisfy? First of all, since $x_s(0) = P$ for all small s, we again have that $\varepsilon(0) = 0$. But what about at $t = 1$? The condition $x_s(1) \in C$ is equivalent

to $g(x_s(1)) = 0$ and this has to hold for s small enough. Differentiating at $s = 0$, we find by the chain rule and the definition of ε,

$$0 = \frac{d}{ds} g(x_s(1))\Big|_{s=0} = Dg(x(1))\varepsilon(1) = \left\langle \nabla g|_{x(1)}, \varepsilon(1) \right\rangle .$$

In other words, $\varepsilon(1)$ is perpendicular to the gradient of g at $x(1)$ or, equivalently, it is tangent to the curve C at $x(1)$.

We will now calculate

$$\frac{d}{ds} S[x_s]\Big|_{s=0} = 0 \tag{9}$$

as we did before in Section 2.3. The calculation is very similar, the only difference is that when we integrate the $\dot{\varepsilon}$ term by parts we do pick a contribution from the boundary at $t = 1$. The final result is that the right-hand side of equation (2) acquires an extra term:

$$\frac{d}{ds} S[x_s]\Big|_{s=0} = -\int_0^1 \left\langle \frac{d}{dt}\left(\frac{\dot{x}}{\|\dot{x}\|}\right), \varepsilon \right\rangle dt + \left\langle \frac{\dot{x}(1)}{\|\dot{x}(1)\|}, \varepsilon(1) \right\rangle .$$

Now, to analyse the vanishing of this expression, we notice that it has to vanish for all admissible variations, and in particular for those which satisfy $\varepsilon(1) = 0$. Then the extra term cancels and we are back in the previous situation, to which the Fundamental Lemma applies. Hence we know that

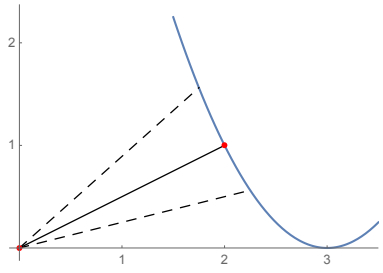
$$\frac{d}{dt}\left(\frac{\dot{x}}{\|\dot{x}\|}\right) = 0 ,$$

so that the geodesics will again be straight lines. The extra term now imposes a boundary condition on the straight line, namely

$$\left\langle \frac{\dot{x}(1)}{\|\dot{x}(1)\|}, \varepsilon(1) \right\rangle = 0 ,$$

which says that the velocity $\dot{x}(1)$ of the line at $x(1)$ has to be perpendicular to $\varepsilon(1)$. Since $\varepsilon(1)$ can be any vector tangent to C at $x(1)$, the velocity $\dot{x}(1)$ has to be normal to C at $x(1)$.

Example 3.1. *Find the shortest path in the plane starting at the origin and ending on the parabola* $y = (x-3)^2$.



Solution: The Euler–Lagrange equation being unchanged, the path will be a straight line from the origin to a point (x_1, y_1) on the parabola. The parabola is the zero locus of the function

$$g(x,y) = y - (x-3)^2$$

whose gradient at (x_1, y_1) is given by

$$\nabla g|_{(x_1,y_1)} = \begin{pmatrix} -2(x_1-3) \\ 1 \end{pmatrix} .$$

The straight line from the origin to (x_1, y_1) has velocity (x_1, y_1), which to intersect the parabola normally must be collinear with the gradient of g at (x_1, y_1). In other words,

$$\begin{Vmatrix} x_1 & -2(x_1-3) \\ y_1 & 1 \end{Vmatrix} = 0 \implies x_1 + 2y_1(x_1-3) = 0 .$$

In addition we have that $y_1 = (x_1-3)^2$. Substituting this equation into the above equation we arrive at a cubic equation for x_1:

$$x_1 + 2(x_1-3)^3 = 0 .$$

This has the obvious solution $x_1 = 2$ and factoring out $(x_1 - 2)$ we see that it's the only real solution. For $x_1 = 2$, we calculate $y_1 = 1$ on the parabola, whence the shortest path is between the origin and the point $(2, 1)$. ◇

This discussion can easily be generalised in several ways. Let me briefly mention each one in turn.

Firstly, we can let the initial condition vary as well. This would correspond, say, to finding the shortest path between two curves C_0 and C_1 in the plane, say. Following the same steps as before we see that the admissible variations are now $\varepsilon : [0,1] \to \mathbb{R}^2$ with $\varepsilon(0)$ tangent to C_0 and $\varepsilon(1)$ tangent to C_1. Varying the arclength functional we would receive now not one but two boundary contributions upon integrating by parts. The end result is that the geodesics are again straight lines with the boundary conditions that $x(0) \in C_0$ with $\dot{x}(0)$ normal to C_0, and also $x(1) \in C_1$ with $\dot{x}(1)$ normal to C_1.

Exercise 3.2. *What is the shortest distance between two non-overlapping circles on the plane?*

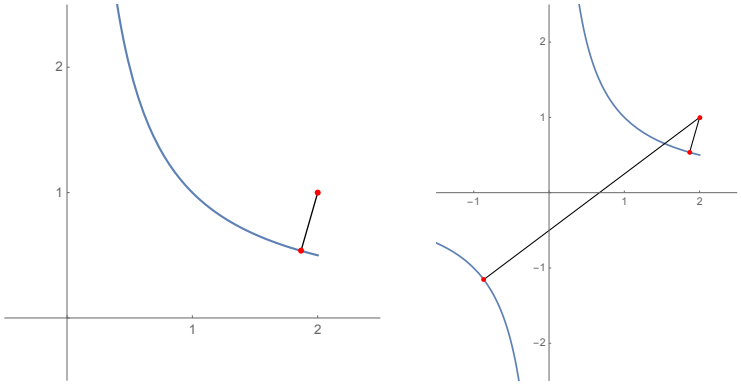
Exercise 3.3. *Let C_1 and C_2 be two closed simple plane curves, with C_1 in the upper half-plane and C_2 in the lower half-plane. Show that there is exists a straight line in the plane which intersects both C_1 and C_2 normally. (The line might intersect each curve at more than one point, but the intersection is normal in at least one of those points.)*

A second generalisation is to consider the shortest path in $\mathbb{R}^n$ between two hypersurfaces (or between a point and a hypersurface), where by a hypersurface I mean the zero set of a differentiable function $g : \mathbb{R}^n \to \mathbb{R}$. The result of the variational analysis is formally identical to the case of $n = 2$: namely, the geodesics are straight lines which are normal to the hypersurface at their intersection.

One important difference between these more general problems and the simpler problem of finding the geodesic between two points in $\mathbb{R}^n$, is that whereas

the latter has a unique solution, the former may have many solutions: there may be both local and global minima (and maxima) and even global minima need not be unique.

Example 3.2. *Find the shortest path in the plane starting at* $\mathrm{P} = (2,1)$ *and ending on the hyperbola* $xy = 1$.



Solution: Let (x_1, y_1) be a point on the hyperbola, so that $x_1 y_1 = 1$, and let $x(t) = (1-t)(2,1) + t(x_1, y_1)$ be the straight line segment from P to the point (x_1, y_1). Its velocity is constant and given by $\dot{x} = (x_1 - 2, y_1 - 1)$. Let $g(x,y) = xy - 1$ be the function whose zero locus defines the hyperbola. Its gradient at (x_1, y_1) is given by $\nabla g = (y_1, x_1)$. The condition that $\dot{x}$ should be collinear with ∇g at (x_1, y_1) is the vanishing of the 2×2 determinant:

$$\begin{Vmatrix} x_1 - 2 & y_1 \\ y_1 - 1 & x_1 \end{Vmatrix} = 0 \implies x_1^4 - 2x_1^3 - (1 - x_1) = 0,$$

where we have used $x_1 y_1 = 1$ to eliminate y_1. In terms of $v = 2x_1 - 1$, the quartic equation becomes a quadratic equation in $u = v^2$:

$$u^2 - 6u - 11 = 0 \implies u = 3 \pm 2\sqrt{5},$$

but only $u = 3 + 2\sqrt{5}$ is positive (after all, $u = v^2$). Solving for v and hence for x_1 we find two real solutions:

$$x_1 = \tfrac{1}{2}\left(1 \pm \sqrt{3 + 2\sqrt{5}}\right).$$

As can be seen in the figures, both correspond to local minima; although one path is definitely shorter than the other. Notice the each solution lies in a separate branch of the hyperbola. In Physics parlance one would say that the two extrema lie in different "topological sectors", each "sector" consisting of curves from P to one of the two connected components of the hyperbola. The lesson here is that the variational calculus, by its very nature, only compares curves which can be continuously "varied" into each other. Two curves from P to the hyperbola whose endpoints lie on different branches cannot be continuously varied into one another and this is why we should expect (at least) two solutions. ◇

Finally, there is no reason we have to consider the arclength functional. We could ask to minimise an action functional

$$\mathrm{I}[x] = \int_0^1 \mathrm{L}(x(t), \dot{x}(t), t)dt\,,$$

associated to some lagrangian function L, over all curves $x : [0,1] \to \mathbb{R}^n$ obeying boundary conditions such as $x(0) \in \mathrm{C}_0$ and $x(1) \in \mathrm{C}_1$, where C_0 and C_1 are

hypersurfaces defined by equations $g_0 : \mathbb{R}^n \to \mathbb{R}$ and $g_1 : \mathbb{R}^n \to \mathbb{R}$. The result of the analysis (which we detail below) is that the critical curves satisfy the Euler–Lagrange equations with the additional requirements that $\frac{\partial \mathrm{L}}{\partial \dot{x}^i}$ should be normal to C_0 at $x(0)$ and to C_1 at $x(1)$. Since $\frac{\partial \mathrm{L}}{\partial \dot{x}^i}$ need not be collinear to $\dot{x}^i$, it is not necessarily the case that the curves are incident to C_i normally.

We vary the action functional $\mathrm{I}[x]$ but we keep the boundary terms:

$$\begin{aligned}\frac{d}{ds}\mathrm{I}[x+s\varepsilon]\Big|_{s=0} &= \int_0^1 \left(\sum_{i=1}^n \frac{\partial \mathrm{L}}{\partial x^i}\varepsilon_i + \sum_{i=1}^n \frac{\partial \mathrm{L}}{\partial \dot{x}^i}\dot{\varepsilon}_i\right) dt \\ &= \int_0^1 \sum_{i=1}^n \left(\frac{\partial \mathrm{L}}{\partial x^i} - \frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{x}^i}\right)\varepsilon_i dt + \sum_{i=1}^n \frac{\partial \mathrm{L}}{\partial \dot{x}^i}\varepsilon_i\Big|_0^1 .\end{aligned}$$

This has to vanish for all admissible variations. In particular, if we take variations which vanish at the endpoints, the Fundamental Lemma implies that the Euler–Lagrange equations (6) have to be obeyed. In addition, the boundary terms must also vanish. If the endpoints are such that $x(0)$ must lie on C_0 and $x(1)$ must lie on C_1, the variations, as shown above, must be tangent to the hypersurfaces, so that $\varepsilon(0)$ is tangent to C_0 and $\varepsilon(1)$ is tangent to C_1. This then imposes that the vector $\frac{\partial \mathrm{L}}{\partial \dot{x}^i}\Big|_0$ is normal to C_0 at $x(0)$ and $\frac{\partial \mathrm{L}}{\partial \dot{x}^i}\Big|_1$ is normal to C_1 at $x(1)$.

This method extends to lagrangians $\mathrm{L}(x, \dot{x}, \ddot{x}, ..., t)$ which depend on derivatives of the dependent variable of higher order. We simply vary the action functional keeping all boundary terms. We then notice that the Euler–Lagrange equations are still the same as for endpoint-fixed variations, but we must also cancel the boundary terms and this depends on the precise nature of the boundary conditions of the problem at hand.

4 Newtonian mechanics

> I think that Isaac Newton is doing most of the driving right now.
> — Bill Anders, Apollo 8 mission

We now review the basic notions of newtonian mechanics, which deals with the motions of (idealised) point particles in space. Being based on empirical evidence, these assumptions have a limited domain of validity and hence the laws derived from them are known to break down in the very small, the very large or the very fast. Nevertheless newtonian mechanics has a remarkably wide domain of applicability, encompassing for instance both apples falling on the surface of the Earth and planets orbiting stars. Historically it was also the first modern physical theory.

4.1 The universe according to Newton

The newtonian universe is $\mathbb{R} \times \mathbb{R}^3$, where $\mathbb{R}$ is **time** and $\mathbb{R}^3$ is a 3-dimensional euclidean **space** with the usual scalar ("dot") product,

$$\langle a, b \rangle = a_1 b_1 + a_2 b_2 + a_3 b_3 = \sum_{i=1}^{3} a_i b_i \,, \tag{10}$$

for $a = (a_1, a_2, a_3)$, $b = (b_1, b_2, b_3)$.

A point (t, a) in the universe is called an **event**. Two events (t, a) and (t', b) are said to be **simultaneous** if and only if $t = t'$. It makes sense to talk about the **distance** between simultaneous events (t, a) and (t, b), and this is given by

$$\|a - b\| = \sqrt{\langle a - b, a - b \rangle} \,. \tag{11}$$

Particle trajectories are given by **worldlines**, which are graphs of functions $x : \mathbb{R} \to \mathbb{R}^3$; that is, subsets of the universe of the form

$$\{(t, x(t)) \mid t \in \mathbb{R}\} \,, \tag{12}$$

where $x(t) = (x^1(t), x^2(t), x^3(t))$ are the components. We will assume that the x^i are (at least) twice continuously differentiable. Figure 3 illustrates the worldlines of two particles.

Let $x : \mathbb{R} \to \mathbb{R}^3$ define the worldline of a particle. The first derivative (with respect to time) $\dot{x}$ is called the **velocity** and the second derivative $\ddot{x}$ the **acceleration**.

We are often interested in mechanical systems consisting of more that one particle. The **configuration space** of an n-particle system is the n-fold cartesian product

$$\underbrace{\mathbb{R}^3 \times \cdots \times \mathbb{R}^3}_{n} = \mathbb{R}^{\mathrm{N}} \,, \qquad \mathrm{N} = 3n \,.$$

Since we do not allow two particles to occupy the same point in space at the same time, the more realistic configuration space of n particles is the complement in $\mathbb{R}^{3n}$ of the subspaces where the coordinates of any two of the particles coincide.

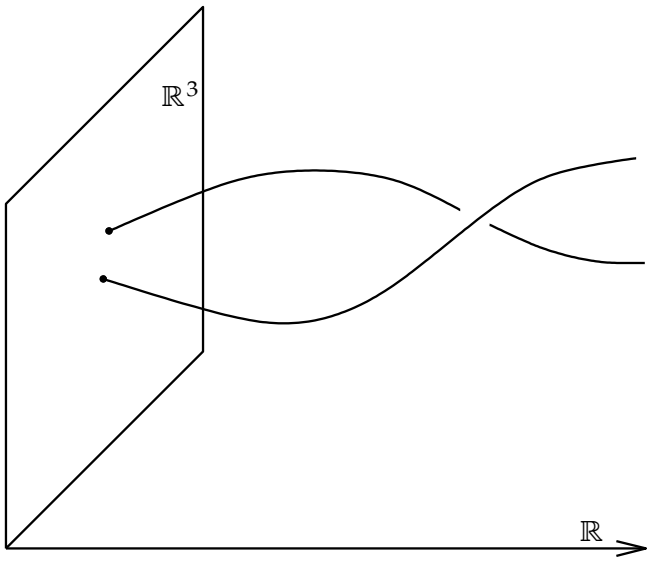


Figure 3: Two worldlines

The worldline of the ith particle is given by $x_i : \mathbb{R} \to \mathbb{R}^3$ and the n worldlines together define a curve $x : \mathbb{R} \to \mathbb{R}^N$ in the configuration space,

$$x(t) = (x_1(t), \ldots, x_n(t)) \ .$$

We are also interested in mechanical systems consisting of "bodies" with structure; that is, not just finite collections of point-like particles. For example, we may want to consider a spinning top or some other solid object. In this case, a configuration of the system might be given by a vector specifying the location of the body and a rotation which says how to body is oriented relative to a fixed set of orthonormal axes. A trajectory of such a *rigid body* is then a map $\mathbb{R} \to \mathbb{R}^3 \times$ SO(3), sending t to $(\mathbf{x}(t), \mathrm{R}(t))$, where SO(3) is the group of unit-determinant orthogonal matrices. This illustrates the fact that the configuration space of a mechanical system need not be $\mathbb{R}^n$, but can in fact be a more general *manifold*. The parameters we use to give coordinates to the configuration space (in the case of SO(3) a convenient choice is given by the *Euler angles*) are usually called **generalised coordinates** and are often written simply as q, to distinguish them from the Cartesian-looking x.

4.2 Newton's equation

The other basic assumption of newtonian mechanics is **determinacy**, which means that the **initial state** of a mechanical system, by which we mean the totality of the positions and velocities of all the particles at a given instant in time, uniquely determines the motion. In other words, for all t_0, $x(t_0)$ and $\dot{x}(t_0)$ determine $x(t)$ for all $t > t_0$.

In particular, the acceleration is determined, so there must exist some rela-

tionship of the form

$$\boxed{\ddot{x}(t) = \Phi(x(t), \dot{x}(t), t)}\ , \tag{13}$$

for some function $\Phi : \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}^N$. This second-order ordinary differential equation (ODE) is called **Newton's equation**. Solving a second-order ODE involves integrating twice, which gives rise to two constants of integration (per degree of freedom). These constants are then fixed by the initial conditions.

We will be dealing almost exclusively with functions Φ depending only on x and possibly on t, but not on $\dot{x}$.

Exercise 4.1. *Let Φ depend only on x. Show that Newton's equation is invariant under time reversal; that is, show that if $x(t)$ solves the equation, so does $\bar{x}(t) := x(-t)$.*

The version of Newton's equation (13) which describes the motion of a particle in the presence of a force field $\mathrm{F} : \mathbb{R}^3 \to \mathbb{R}^3$ is

$$\boxed{\mathrm{F}(x) = m\ddot{x}} \tag{14}$$

where m is the **(inertial) mass** of the particle.

Exercise 4.2. *A particle of mass m is observed moving in a circular trajectory*

$$x(t) = (\mathrm{R}\cos\omega t, \mathrm{R}\sin\omega t, 0)\ , \tag{15}$$

where R, ω are positive constants. What is the force acting on the particle?

Another interesting example is that of the **free particle**. This is a particular case of the previous example, where $\mathrm{F} = 0$. Newton's equation (14) says that there is no acceleration, so that the velocity v is constant. Integrating a second time we obtain

$$x(t) = x_0 + tv\ , \tag{16}$$

where $x_0 = x(0)$ is the initial position. Given x_0 and v there is a unique solution $x(t)$ to Newton's equation with $\mathrm{F} = 0$ with $x(0) = x_0$ and $\dot{x}(0) = v$.

A standard trick allows us to turn Newton's equation (13) into an equivalent first-order ODE. The trick consists in introducing a new function $v : \mathbb{R} \to \mathbb{R}^N$ together with the equation $\dot{x} = v$. Newton's equation is then

$$\ddot{x} = \dot{v} = \Phi(x, v, t)\ .$$

In other words, in terms of the function $(x, v) : \mathbb{R} \to \mathbb{R}^{2N}$, Newton's equation becomes

$$\boxed{(\dot{x}, \dot{v}) = (v, \Phi(x, v, t))\ .} \tag{17}$$

The space of positions and velocities, here $\mathbb{R}^{2N}$, defines the **state space** of the mechanical system. The pair of functions (x, v) defines a curve in the space of

states, which, if it obeys (17), is called a **physical trajectory**. This reformulation of Newton's equation allows us to use the results on uniqueness and existence of solutions to initial value problems, by which if Φ is sufficiently differentiable, equation (17) has a unique solution for specified initial conditions $(x(0), v(0) = \dot{x}(0))$, at least in some time interval. In other words, through every point in state space there passes a unique physical trajectory.

As an example, consider dropping an apple of mass m from the Tower of Pisa. (This is an instance of **galilean gravity**.) Empirical evidence suggests that the force of gravity points downwards, is constant and proportional to the mass. Letting $z(t)$ denote the height at time t, Newton's equation is then

$$m\ddot{z} = -mg \,, \tag{18}$$

where g is a physical constant (known as the *acceleration due to gravity*) and $g \approx 9.8\text{ms}^{-2}$ on the surface of the Earth. We can solve equation (18) by integrating twice

$$z(t) = z_0 + v_0 t - \tfrac{1}{2} g t^2 \,.$$

The relevant space of states is the right half-plane

$$\{(z, v) \mid z \geq 0\} \subset \mathbb{R}^2 \,, \tag{19}$$

and the physical trajectories are the parabolas given by

$$(z(t), v(t)) = \left(z_0 + v_0 t - \tfrac{1}{2} g t^2, v_0 - gt\right) \,. \tag{20}$$

Some of these trajectories are plotted in Figure 4.

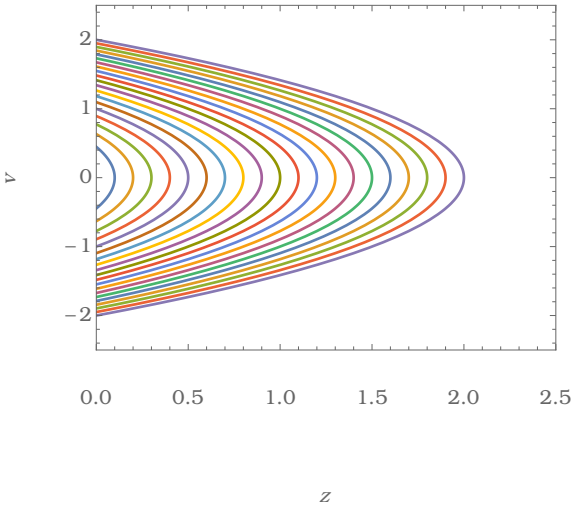


Figure 4: Physical trajectories of equation (18) in units where $g = 1$

Notice that whatever the initial conditions (z_0, v_0) the apple always ends up on the floor. This is contrary to observation (e.g., rockets can break free of Earth's

gravity) and indeed it is known that as the distance from the Earth increases, her gravitational pull weakens. This failure of galilean gravity was corrected by Newton in his universal law of gravitation.

The equivalence principle

The m in the RHS of equation (18) is called the **(gravitational) mass** and it is an empirical fact (famously demonstrated by Galileo and later by Eötvös) that it is equal to the (inertial) mass appearing in the LHS. This equality is called the **equivalence principle**: it hints at a geometric origin of gravity and is a cornerstone of Einstein's general theory of relativity.

4.3 Galilean relativity

This section lies outside the scope of this course and the material in it is certainly non-examinable. I include it because it is part of the logical story and in a way it is essential to understanding the newtonian paradigm.[1]

The first remark is that I oversimplified a little the newtonian universe. The problem with $\mathbb{R}\times\mathbb{R}^3$ is that both the real line and $\mathbb{R}^3$ are vector spaces, and that means that they are "pointed spaces"; that is, they come with a privileged point: namely, the origin. This means that the newtonian universe comes with a "time zero" and also with a centre of the spatial universe. Of course, for the purposes of computations this is almost unavoidable: the moment we give coordinates to a space we are choosing an origin of the coordinate system, but at a conceptual level it is a little unsatisfactory.

The standard way to "forget" the origin in a vector space is to pass to the **affine space**. An affine space for the vector space $\mathbb{R}^n$ is a space A^n on which $\mathbb{R}^n$ acts freely and transitively. In other words, we have a map $\mathbb{R}^n\times\mathrm{A}^n\to\mathrm{A}^n$ such that $(\mathbf{v},a)\mapsto\mathbf{v}+a$, where $\mathbf{v}\in\mathbb{R}^n$, $a,\mathbf{v}+a\in\mathrm{A}^n$ satisfying the properties suggested by the notation: namely,

- $\mathbf{v}+(\mathbf{w}+a)=(\mathbf{v}+\mathbf{w})+a$, and
- $\mathbf{v}+a=a$ if and only if $\mathbf{v}=\mathbf{0}$.

In particular, we cannot add points in A^n, but we can take differences: for all $a,b\in\mathrm{V}$, $b-a\in\mathbb{R}^n$ is the unique vector $\mathbf{v}$ such that $b=\mathbf{v}+a$. Once we declare that a point $a\in\mathrm{A}^n$ is the *origin*, any other point $b\in\mathrm{A}^n$ is labelled by a vector $\mathbf{v}\in\mathbb{R}^n$: namely, the unique vector $b-a$. So we recover the vector space via the choice of an origin.

So the newtonian universe is in fact an affine space A^4, but this is not all. A^4 with a choice of origin becomes $\mathbb{R}^4$, but we said that the newtonian universe is $\mathbb{R}\times\mathbb{R}^3$. This "split" into space and time is tantamount to a *linear* map $\tau:\mathbb{R}^4\to\mathbb{R}$ which measures the **time interval** between two events $a,b\in\mathrm{A}^4$. Indeed, given $a,b\in\mathrm{A}^4$, $b-a\in\mathbb{R}^4$ and hence $\tau(b-a)\in\mathbb{R}$. Events $a,b\in\mathrm{A}^4$ are **simultaneous** if $\tau(b-a)=0$. The kernel of τ is a vector subspace of $\mathbb{R}^4$

[1] There are those who say that declaring lecture material "non-examinable" guarantees that nobody will read it, but I hope this is not universally the case!

isomorphic to $\mathbb{R}^3$. This subspace acts on A^4 in such a way that it takes an event to a simultaneous event. The set of events simultaneous with a given event forms a three-dimensional affine subspace in A^4. Any such **space of simultaneous events** has the structure of a three-dimensional euclidean space, since if $a, b \in A^4$ are simultaneous, we can define their distance by

$$\rho(a,b) = \|a-b\| = \sqrt{\langle a-b, a-b\rangle}\,.$$

So we can rephrase the newtonian universe in the following way: it is a triple (A^4, τ, ρ) consisting of a four-dimensional affine space A^4, a linear function $\tau : \mathbb{R}^4 \to \mathbb{R}$ assigning real numbers to differences of events, and defining affine subspaces of simultaneous events on which we have a distance ρ turning them into euclidean spaces.

Of course, once we choose an origin, we can view A^4 as $\mathbb{R}^4$ with $\tau : \mathbb{R}^4 \to \mathbb{R}$ is projection onto the first entry, so that $\mathbb{R}^4 = \mathbb{R} \times \mathbb{R}^3$ and on $\mathbb{R}^3$ we have the dot product, which gives the euclidean structure. This recovers the definition of the newtonian universe given in the actual lecture.

How about *relativity*? This, along with determinacy and the model of the universe just exposed, is one of the conceptual cornerstones of newtonian mechanics.

Every model of the universe has a **relativity group**; that is, a group of transformations on the universe preserving whichever structure we've endowed the universe with. The newtonian universe is an affine space with additional structure, hence its relativity group will be the subgroup of affine transformations which leaves the additional structure invariant.

Rather than working with abstract affine transformations, let us revert to the first version of the newtonian universe $\mathbb{R} \times \mathbb{R}^3$.

An affine transformation of $\mathbb{R} \times \mathbb{R}^3$ is given by a general linear transformation followed by a translation, so that affine transformations are given by pairs consisting of an invertible 4×4 real matrix and a vector in $\mathbb{R}^4$. Since we are splitting $\mathbb{R}^4$ as $\mathbb{R} \times \mathbb{R}^3$, let us write the matrix as

$$\begin{pmatrix} \alpha & \mathbf{w}^t \\ \mathbf{v} & L \end{pmatrix}$$

where $\alpha \in \mathbb{R}$, $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$, and L is a 3×3 matrix. Therefore a general affine transformation on $\mathbb{R} \times \mathbb{R}^3$ can be written as

$$\begin{pmatrix} t \\ \mathbf{x} \end{pmatrix} \mapsto \begin{pmatrix} t' \\ \mathbf{x}' \end{pmatrix} = \begin{pmatrix} \alpha & \mathbf{w}^t \\ \mathbf{v} & L \end{pmatrix} \begin{pmatrix} t \\ \mathbf{x} \end{pmatrix} + \begin{pmatrix} s \\ \mathbf{a} \end{pmatrix} = \begin{pmatrix} \alpha t + \mathbf{w}^t\mathbf{x} + s \\ \mathbf{v}t + L\mathbf{x} + \mathbf{a} \end{pmatrix},$$

where $s \in \mathbb{R}$ and $\mathbf{a} \in \mathbb{R}^3$. The subgroup of affine transformations of $\mathbb{R} \times \mathbb{R}^3$ which leave invariant the time interval between events and the distance between simultaneous events is called the **Galilean group**.

Exercise 4.3. *Show that the Galilean group consists of affine transformations of the form*

$$\begin{pmatrix} t \\ \mathbf{x} \end{pmatrix} \mapsto \begin{pmatrix} t' \\ \mathbf{x}' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \mathbf{v} & \mathrm{R} \end{pmatrix} \begin{pmatrix} t \\ \mathbf{x} \end{pmatrix} + \begin{pmatrix} s \\ \mathbf{a} \end{pmatrix} = \begin{pmatrix} t+s \\ \mathbf{v}t + \mathrm{R}\mathbf{x} + \mathbf{a} \end{pmatrix},$$

where $\mathrm{R} \in \mathrm{O}(3)$ *is an orthogonal matrix; that is,* $\mathrm{R}^t = \mathrm{R}^{-1}$*. In addition, show that every galilean transformation can be written uniquely as the composition of three elementary galilean transformations:*

1. *translations in space and time:*

$$\begin{pmatrix} t \\ \mathbf{x} \end{pmatrix} \mapsto \begin{pmatrix} t+s \\ \mathbf{x}+\mathbf{a} \end{pmatrix};$$

2. *orthogonal transformations in space:*

$$\begin{pmatrix} t \\ \mathbf{x} \end{pmatrix} \mapsto \begin{pmatrix} t \\ \mathrm{R}\mathbf{x} \end{pmatrix} \qquad \mathrm{R} \in \mathrm{O}(3);$$

3. *and* ***galilean boosts***

$$\begin{pmatrix} t \\ \mathbf{x} \end{pmatrix} \mapsto \begin{pmatrix} t \\ \mathbf{x}+t\mathbf{v} \end{pmatrix}.$$

5 Conservation laws

Nature uses as little as possible of anything.
— Johannes Kepler

As a mechanical system evolves in time it will change its state (x, v) according to Newton's equation (17). However there are functions of (x, v) which remain constant. Such functions are called **integrals** of the motion. Among them there are some which are of particular importance in mechanics. At a fundamental level they are related to symmetries of the physical system: homogeneity of space and time (the fact that there is no preferred origin or initial time) and isotropy of space (the fact that there is no preferred direction), for example. Such an integral of the motion is called a **conserved quantity** due to the fact that it is additive in the sense that, if a mechanical system is composed of two non-interacting parts, then its value for the system is the sum of its values for each of the parts.

For the free particle of the previous section, the **momentum** $p : \mathbb{R}^6 \to \mathbb{R}^3$, defined by $p(x, v) = mv$ is conserved. Of course, v is also conserved, but it is the momentum which is additive. Indeed, if we now consider a system of two non-interacting free particles, with momenta $p_1 = m_1 v_1$ and $p_2 = m_2 v_2$, the momentum of the system will be the sum $p = p_1 + p_2$.

For the falling apple in galilean gravity, the **energy**

$$\mathrm{E}(z, v) = \tfrac{1}{2} m v^2 + mgz$$

is conserved. Since $z \geq 0$, the energy is non-negative. In this case, the physical trajectories are the parabolas $\frac{1}{2}v^2 + gz = \mathrm{E}/m$, just as we had found by integrating Newton's equation.

5.1 Conservative forces: potentials

A force field $\mathrm{F} : \mathbb{R}^3 \to \mathbb{R}^3$ is said to be **conservative** if it can be expressed as (minus) the gradient of a function $\mathrm{V} : \mathbb{R}^3 \to \mathbb{R}$, called the **potential**: $\mathrm{F} = -\nabla \mathrm{V}$. The potential is only defined up to a constant and the minus sign is conventional.

Exercise 5.1. *Show that the gravitational potential in galilean gravity is given by* $\mathrm{V} = mgz$.

More generally, for an n-particle mechanical system with configuration space $\mathbb{R}^N$ and state space $\mathbb{R}^{2N}$, a potential is a function $\mathrm{U} : \mathbb{R}^N \to \mathbb{R}$ such that Newton's equation (13) can be written as

$$\ddot{x} = -\nabla \mathrm{U} \, .$$

A common feature of conservative force fields is that energy is conserved along physical trajectories. Indeed, Newton's equation (14) for a conservative force field is (after bringing the force term to the LHS):

$$m\ddot{x} + \nabla \mathrm{V} = 0 \, .$$

Taking the inner product with $\dot{x}$ we obtain

$$m\,\langle \ddot{x}, \dot{x}\rangle + \langle \nabla \mathrm{V}, \dot{x}\rangle = 0\,,$$

which we recognise as the constancy along physical trajectories of the **energy**

$$\boxed{\mathrm{E}(x,v) = \tfrac{1}{2}m\|v\|^2 + \mathrm{V}\,.} \tag{21}$$

Indeed, a version of the product (Leibniz) rule says that

$$\frac{d}{dt}\|v\|^2 = 2\,\langle v, \dot{v}\rangle\ ,$$

and the chain rule says that

$$\frac{d}{dt}\mathrm{V} = \langle \dot{x}, \nabla \mathrm{V}\rangle\ ,$$

whence along physical trajectories, where $v = \dot{x}$ and $\ddot{x} = -\nabla \mathrm{V}$, we find

$$\frac{d}{dt}\mathrm{E}(x,v) = m\,\langle \ddot{x}, \dot{x}\rangle + \langle \nabla \mathrm{V}, \dot{x}\rangle = 0\,.$$

The first term in the RHS of the expression (21) for the energy is called the **kinetic energy** and depends on the motion of the particle, whereas the second term is the **potential energy** and depends on the position. Physical trajectories lie on "constant energy surfaces" in the space of states, defined by

$$\{(x,v) \mid \mathrm{E}(x,v) = \mathrm{E}_0\}\ .$$

In the case of one-dimensional motion energy conservation alone suffices to determine the physical trajectories, as we saw already in the case of galilean gravity.

5.2 Equilibria and critical points

Consider a particle moving in a force field $\mathrm{F} : \mathbb{R}^3 \to \mathbb{R}^3$. A point $x_0 \in \mathbb{R}^3$ where $\mathrm{F}(x_0) = 0$ is said to be a point of **equilibrium**, since a particle sitting at x_0 feels no force.

Exercise 5.2. *Let $x(t)$ be a physical trajectory for which $x(0) = x_0$ is a point of equilibrium. Show that if $\dot{x}(0) = 0$ then $x(t) = x_0$ for all t.*

Everyday experience provides a certain intuitive notion of 'stability' of an equilibrium point. We will now try to make this intuition precise. We will restrict ourselves for simplicity to the case of conservative force fields. An equilibrium point is then a **critical point** of the potential; that is, a point x_0 where $\nabla \mathrm{V}(x_0)$ vanishes.

Let us consider a one-dimensional mechanical system; that is, a particle moving in one dimension. Let $x(t)$ denote the position of the particle at time t and let $\mathrm{V}(x)$ denote the potential. Newton's equation is then simply

$$m\ddot{x} = -\frac{d\mathrm{V}}{dx} \,. \tag{22}$$

Without loss of generality let us assume that $x = 0$ is a critical point of the potential. Since the potential is defined up to an additive constant, let us choose that constant so that $\mathrm{V}(0) = 0$ without loss of generality. Expanding V around 0 we find that

$$\mathrm{V}(x) = \tfrac{1}{2}kx^2 + \mathrm{O}(x^3) \,, \tag{23}$$

where $k := \mathrm{V}''(0)$ is the second derivative of V with respect to x evaluated at $x = 0$. For small deviations from equilibrium we may truncate this expansion and assume that $\mathrm{V}(x) = \frac{1}{2}kx^2$. Newton's equation is then

$$m\ddot{x} = -kx \,. \tag{24}$$

There are three cases depending on whether k is positive, zero or negative. If $k = 0$ we say that the critical point is **degenerate** and we cannot say anything about its stability. Otherwise we have a **non-degenerate** critical point: a local minimum if $k > 0$ and a local maximum if $k < 0$. In the former case the force tends to restore the system to equilibrium and we say that the equilibrium point is **stable**. In the latter case the force tends to push away from equilibrium and we say that the equilibrium point is **unstable**. Stable equilibria give rise to simple harmonic motion, as we now discuss.

5.3 Simple harmonic motion

Consider a point-particle of mass m moving on the x-axis subject to a conservative force with potential $\mathrm{V}(x) = \frac{1}{2}kx^2$, with $k > 0$. The resulting force is $\mathrm{F} = -kx$, which is a good approximation to the restoring force of a spring, an empirical law due to Hooke. Of course, from the previous discussion we see that this is not a particular property of springs, but a universal property of small displacements about stable equilibria!

Hooke's law leads to simple harmonic motion. Indeed, Newton's equation in this case reads

$$m\ddot{x} = -kx \,, \tag{25}$$

whose solutions are

$$x(t) = x_0 \cos \omega t + \frac{v_0}{\omega} \sin \omega t \,,$$

where $\omega^2 = k/m$ and $x_0 = x(0)$ and $v_0 = \dot{x}(0)$ are the initial position and velocity of the particle, respectively.

Exercise 5.3. *Show that the solution can also be written as*

$$x(t) = \sqrt{\frac{2\mathrm{E}}{k}}\,\sin(\omega t + \varphi)\,, \qquad \textit{with} \quad \tan\varphi = \frac{\omega x_0}{v_0}\,,$$

where E *is the energy. In particular, the amplitude of oscillation goes like* $\mathrm{E}^{1/2}$.

The physical trajectories in the space of states $\mathbb{R}^2$ are ellipses corresponding to the constant energy curves

$$\tfrac{1}{2}mv^2 + \tfrac{1}{2}kx^2 = \mathrm{E} \geq 0\,,$$

and some of these curves are plotted in Figure 5. In particular, the physical trajectories are closed and the motion is therefore periodic, with period $2\pi/\omega$.

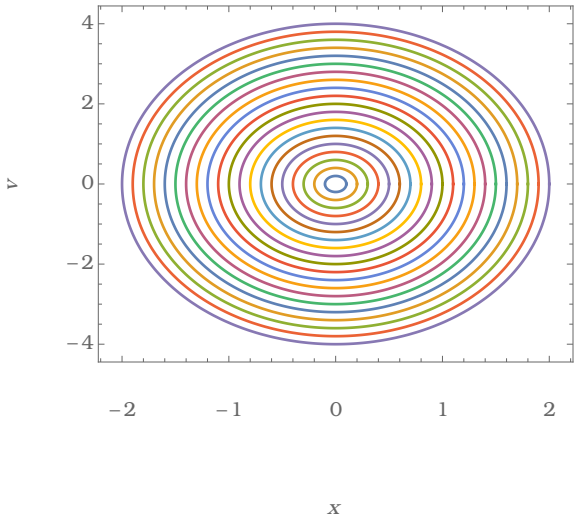


Figure 5: Physical trajectories of equation (25) for $\omega = 2$

5.4 Applications of energy conservation

In this section we show how energy conservation determines the physical trajectories of a one-dimensional conservative mechanical system; that is, one described by equation (22). As remarked above, the energy

$$\mathrm{E} = \tfrac{1}{2}mv^2 + \mathrm{V}(x) \tag{26}$$

is a constant of the motion. Notice that the kinetic energy term ($\frac{1}{2}mv^2$) is always non-negative, therefore $\mathrm{E} \geq \mathrm{V}(x)$ and equality holds if and only if the velocity vanishes; that is, at a turning point. Configurations with potential energy greater than the energy of a particle are inaccessible. In particular, classical

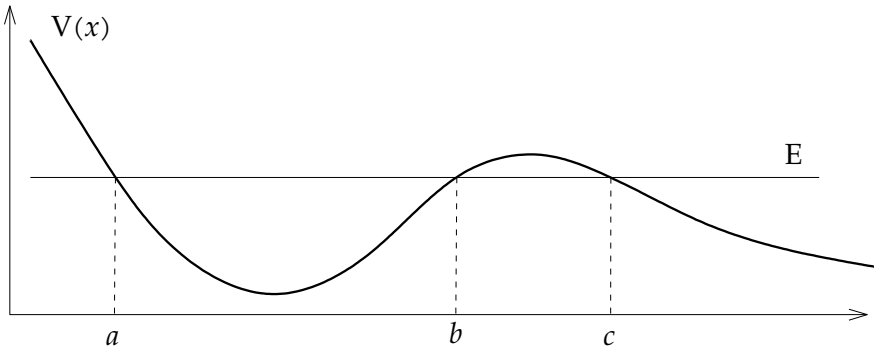


Figure 6: One-dimensional potential motion

particles cannot penetrate potential barriers unless they have sufficient energy. (These statements need to be revised in quantum mechanics.)

Figure 6 illustrates this discussion. It shows the graph of a potential function $\mathrm{V}(x)$ and three turning points: $x = a$, $x = b$ and $x = c$, where $\mathrm{V}(x) = \mathrm{E}$, a fixed value of the energy. Energy conservation means that there are two accessible regions: either the finite interval $[a, b]$ or the semi-infinite interval $[c, \infty)$. If $a \leq x(t) \leq b$ the motion will be oscillatory and if $x(t) \geq c$ then there are two possibilities: either $\dot{x}(0) < 0$, whence $x(0) > c$ and it will move towards to c and then turn and move away forever, or else $\dot{x}(0) \geq 0$ in which case it will move away from c forever.

In the case of oscillatory motion, we can actually prove that the motion is periodic. This follows from uniqueness of the solution of the initial value problem. Let a be a turning point for a given fixed energy. Then there exists a unique solution $x(t)$ with $x(0) = a$, and hence $\dot{x}(0) = 0$. Now suppose that a certain time T later, $x(\mathrm{T}) = a$ again. The function $x_{\mathrm{T}}(t) := x(t + \mathrm{T})$ solves the same differential equation as x, and $x_{\mathrm{T}}(0) = a$, whence $\dot{x}_{\mathrm{T}}(0) = 0$. By uniqueness, $x_{\mathrm{T}} = x$ and we see that $x(t) = x(t + \mathrm{T})$; that is, x is periodic.

Furthermore, one can use energy conservation to derive an expression for the period. Indeed, from

$$\tfrac{1}{2}m\dot{x}^2 + \mathrm{V}(x) = \mathrm{E}$$

we solve for $\dot{x}$ to obtain

$$\dot{x} = \pm\sqrt{\frac{2(\mathrm{E} - \mathrm{V}(x))}{m}}\,.$$

Integrating, we find that the time taken from a to b is

$$\mathrm{T}(a \to b) = \sqrt{\frac{m}{2}}\int_a^b \frac{dx}{\sqrt{\mathrm{E} - \mathrm{V}(x)}}\,.$$

Because of the invariance of Newton's equation (22) under time reversal, this is also the time taken from b to a, whence the period of oscillation is given by

$$\boxed{\mathrm{T} = \sqrt{2m}\int_a^b \frac{dx}{\sqrt{\mathrm{E} - \mathrm{V}(x)}}\,.} \tag{27}$$

Exercise 5.4. *Show that for* $\mathrm{V}(x) = \frac{1}{2}kx^2$*, with* $k > 0$*, the period indeed agrees with* $2\pi/\omega$ *with* $\omega^2 = k/m$*.*

The integral in equation (27) has to be treated with care, because the integrand is singular at the limits of integration, since a and b are zeros of $\mathrm{E} - \mathrm{V}(x)$. In fact, it is not hard to show that the integral converges if and only if a and b are *not* critical points of the potential. To illustrate this, let us suppose that we increase the energy of the particle so that it coincides with a maximum value of the potential, as shown in Figure 7.

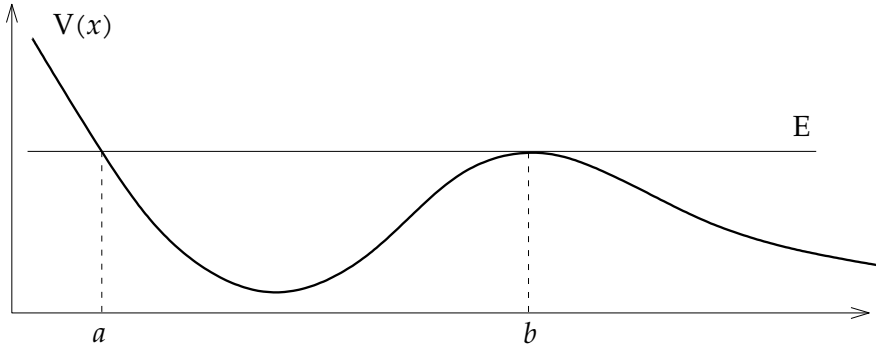


Figure 7: One-dimensional potential motion (cont'd)

Suppose the particle starts from rest at $x(0) = a$. One might be tempted to think that it will move towards b and, upon reaching b, it will turn and come back to a; however this cannot happen, because b is an equilibrium point: if the particle reaches b and turns, it means that it would have zero velocity there, whence it will remain in b forever. What happens in this idealised newtonian universe is that the particle never reaches b! This can be demonstrated by analysing the convergence of the integral which computes the time taken for the particle from a to b, which we now properly write as a limit:

$$\mathrm{T}(a \to b) = \sqrt{\frac{m}{2}} \lim_{\epsilon \to 0^+} \int_a^{b-\epsilon} \frac{dx}{\sqrt{\mathrm{E} - \mathrm{V}(x)}} \,.$$

Exercise 5.5. *Expand* $\mathrm{E} - \mathrm{V}(x)$ *around* b *and show that if* $\mathrm{V}'(b) = 0$ *the integral does not converge.*

In summary, if b is a critical point of the potential, it takes an "infinite" time for the particle to reach b.

6 Hamilton's principle of least action

Physics is where the action is.
— "Popular" saying

In this section we would like to elevate Newton's equations to a variational principle, whenever possible. Let us start with the simplest case of a particle of mass m moving in $\mathbb{R}^3$ under the influence of a conservative force field with potential $\mathrm{V} : \mathbb{R}^3 \to \mathbb{R}$. Then Newton's equation is given by

$$m\ddot{x}^i = -\frac{\partial \mathrm{V}}{\partial x^i} \,. \tag{28}$$

We would like to exhibit this as the Euler–Lagrange equation of some functional.

☞ This problem is known as the *inverse problem* in the calculus of variations. Given a differential equation, determine when it is the Euler–Lagrange equation of a functional and, if so, determine the functional. The modern approach to this problem is *homological* in nature and goes by the name of the *variational bicomplex*.

Let us first consider $\mathrm{V} = 0$ (or indeed any constant). Then Newton's equation is $m\ddot{x} = 0$ and we saw that this is the Euler–Lagrange equation for the action

$$\mathrm{I}_{\text{free}}[x] = \int_0^1 \tfrac{1}{2} m\|\dot{x}\|^2 dt \,.$$

So we expect that for nonzero force, the lagrangian for equation (28) might start like

$$\mathrm{L}(x, v, t) = \tfrac{1}{2} m\|v\|^2 + \cdots \,.$$

We now recall that conservative force fields have a first integral, namely the total energy

$$\mathrm{E} = \tfrac{1}{2} m\|\dot{x}\|^2 + \mathrm{V}(x) \,, \tag{29}$$

and that so do the Euler–Lagrange equation coming from a lagrangian which does not depend explicitly on t; that is, $\frac{\partial \mathrm{L}}{\partial t} = 0$, where now

$$\mathrm{E}' = \sum_i \dot{x}^i \frac{\partial \mathrm{L}}{\partial \dot{x}^i} - \mathrm{L} \,.$$

Write $\mathrm{L} = \frac{1}{2} m\|v\|^2 + \mathrm{U}(x, v)$, whence

$$\begin{aligned} \mathrm{E}' &= m\|\dot{x}\|^2 + \sum_i \dot{x}^i \frac{\partial \mathrm{U}}{\partial \dot{x}^i} - \tfrac{1}{2} m\|\dot{x}\|^2 - \mathrm{U}(x, \dot{x}) \\ &= \tfrac{1}{2} m\|\dot{x}\|^2 + \sum_i \dot{x}^i \frac{\partial \mathrm{U}}{\partial \dot{x}^i} - \mathrm{U}(x, \dot{x}) \,. \end{aligned}$$

Comparing with E in equation (29), we see that choosing $\mathrm{U}(x, v) = -\mathrm{V}(x)$, the two expressions agree.

This is not the only choice for $U(x,v)$, though. For example, we could add to $-V(x)$ a function linear in v, i.e., $U(x,v) = \langle A(x), v\rangle - V(x)$, where $A : \mathbb{R}^3 \to \mathbb{R}^3$ is a vector field. The Euler–Lagrange equation for this more general lagrangian

$$L = \tfrac{1}{2}m\|v\|^2 - V(x) + \langle A(x), v\rangle$$

agrees with Newton's equation if $\nabla \times A = 0$, which in $\mathbb{R}^3$ implies that $A = \nabla\theta$, so that

$$L(x,\dot{x}) = \tfrac{1}{2}m\|\dot{x}\|^2 - V(x) + \langle \nabla\theta, \dot{x}\rangle = \tfrac{1}{2}m\|\dot{x}\|^2 - V(x) + \frac{d\theta}{dt}.$$

For general A, the RHS of the Euler–Lagrange equation defines the Lorentz force a charged particle feels in the presence of an electromagnetic field.

In summary, the Euler–Lagrange equation for the lagrangian $L(x,v) = \frac{1}{2}m\|v\|^2 - V(x)$ is Newton's equation (28). In other words, the physical trajectories $(x(t), v(t))$ in the space of states are precisely those which extremise the **action functional**

$$\boxed{I[x] = \int_0^1 \left(\tfrac{1}{2}m\|\dot{x}\|^2 - V(x)\right) dt}\,.$$

This is Hamilton's **Principle of Least Action**.

Notice that the lagrangian $L = T - V$ is the difference between the kinetic energy $T = \frac{1}{2}m\|\dot{x}\|^2$ and the potential energy $V(x(t))$ along the path.

More generally, for any lagrangian (not necessarily of the form $T - V$) one calls the quantity $\frac{\partial L}{\partial x^i}$ the **force**, the quantity $\frac{\partial L}{\partial \dot{x}^i}$ the **momentum**, and the quantity $\sum_{i=1}^n \dot{x}^i \frac{\partial L}{\partial \dot{x}^i} - L$ the **energy**. For the above lagrangian $L = T - V$, the energy is $T + V$.

In the next section we will see that these and other conservation laws result from *symmetries* of the lagrangian.

The principle of least action changes the language of Newtonian mechanics: force, momentum, energy are now derived concepts from the lagrangian. The name of the game is then the construction and study of lagrangians.

To develop some intuition about "action-building" let us consider some examples.

6.1 The planar pendulum

Figure 8 shows a pendulum of length ℓ with a bob of mass m which is free to move in a plane under the effect of galilean gravity. As usual, we idealise the system by assuming that there is no friction of any kind, that the bob is point-like and that the rod to which the bob is attached is massless.

The bob moves in a circle of radius ℓ centred at the pivot. An arc of this circle is shown by the dotted red line in Figure 8. In other words, the configuration space of this mechanical system is the circle.

The kinetic energy is given by $\frac{1}{2}m\|\dot{x}\|^2$, where $x(t)$ is the position of the bob at time t. From the picture and letting θ denote the angle of the pendulum with

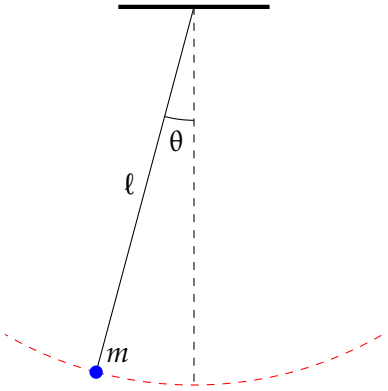


Figure 8: A planar pendulum

respect to the vertical, we have that

$$x(t) = \begin{pmatrix} -\ell \sin\theta(t) \\ \ell(1-\cos\theta(t)) \end{pmatrix}$$

where we take $\theta = 0$ as the origin of coordinates in the plane where the motion takes place. Differentiating,

$$\dot{x} = \begin{pmatrix} -\ell\cos\theta\,\dot{\theta} \\ \ell\sin\theta\,\dot{\theta} \end{pmatrix} \implies \|\dot{x}\|^2 = \ell^2\dot{\theta}^2 \,.$$

The kinetic energy is therefore $\mathrm{T} = \frac{1}{2}m\ell^2\dot{\theta}^2$. The potential energy is given by mg times the height of the bob, whence $\mathrm{V} = mg\ell(1-\cos\theta)$. The lagrangian is then

$$\mathrm{L}(\theta,\dot{\theta}) = \tfrac{1}{2}m\ell^2\dot{\theta}^2 - mg\ell(1-\cos\theta)\,.$$

Since it does not depend manifestly on t, the energy is conserved, whence the possible trajectories in state space are such that

$$\mathrm{E} = \tfrac{1}{2}m\ell^2 v^2 + 2mg\ell\sin^2\tfrac{\theta}{2}\,,$$

where we have used the identity $1-\cos\theta = 2\sin^2\frac{\theta}{2}$, which shows that E is manifestly non-negative. Some of the classical trajectories are depicted in Figure 9.

6.2 Planar pendulum with moving pivot

Consider now a pendulum whose pivot is now a bead of mass M which is free to move horizontally without friction, as shown in Figure 10. Let x denote the position of the pivot measured from the origin O and we now let x_2 denote the position of the bob of the pendulum.

The kinetic energy is the sum of the kinetic energies of the pivot and the bob:

$$\mathrm{T} = \tfrac{1}{2}\mathrm{M}\dot{x}^2 + \tfrac{1}{2}m\|\dot{x}_2\|^2\,,$$

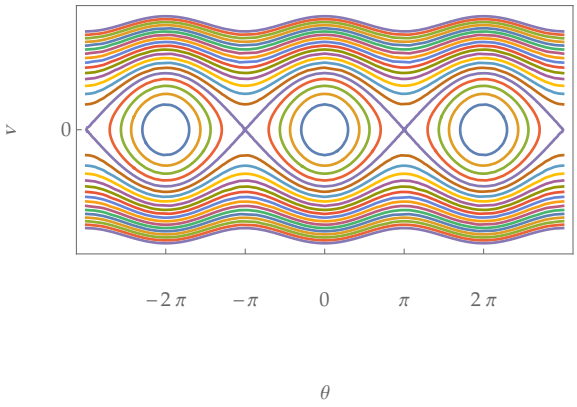


Figure 9: Physical trajectories of pendulum

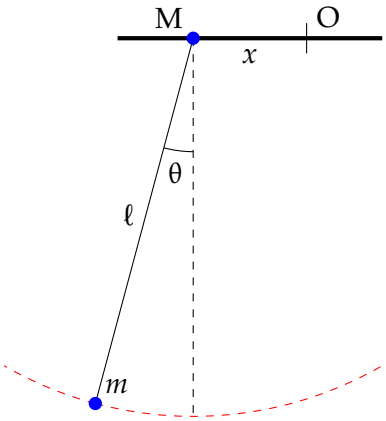


Figure 10: A planar pendulum with moving pivot

where

$$x_2 = \begin{pmatrix} -x - \ell \sin\theta \\ -\ell\cos\theta \end{pmatrix} \implies \dot{x}_2 = \begin{pmatrix} -\dot{x} - \ell\cos\theta\,\dot{\theta} \\ \ell\sin\theta\,\dot{\theta} \end{pmatrix},$$

whence

$$\|\dot{x}_2\|^2 = \dot{x}^2 + \ell^2\dot{\theta}^2 + 2\ell\cos\theta\,\dot{x}\dot{\theta}\,.$$

The kinetic energy is finally given by

$$\mathrm{T} = \tfrac{1}{2}(\mathrm{M}+m)\dot{x}^2 + \tfrac{1}{2}m\ell^2\dot{\theta}^2 + m\ell\cos\theta\,\dot{x}\dot{\theta}\,.$$

The potential energy is again $\mathrm{V} = mg\ell(1-\cos\theta)$, so that the lagrangian is

$$\mathrm{L}(\theta,\dot{x},\dot{\theta}) = \mathrm{T} - \mathrm{V} = \tfrac{1}{2}(\mathrm{M}+m)\dot{x}^2 + \tfrac{1}{2}m\ell^2\dot{\theta}^2 + m\ell\cos\theta\,\dot{x}\dot{\theta} - mg\ell(1-\cos\theta)\,.$$

Since L does not depend explicitly on t, the total energy T + V is conserved:

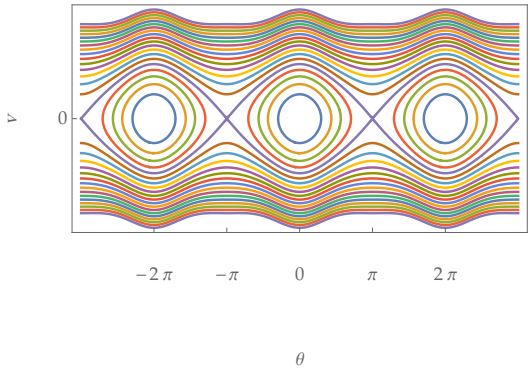
$$\mathrm{E} = \tfrac{1}{2}(\mathrm{M}+m)\dot{x}^2 + \tfrac{1}{2}m\ell^2\dot{\theta}^2 + m\ell\cos\theta\,\dot{x}\dot{\theta} + mg\ell(1-\cos\theta)\,,$$

and similarly, since L does not depend explicitly on x, the momentum along x is also conserved

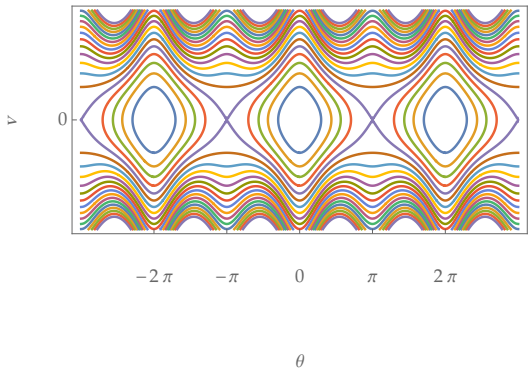
$$\mathrm{P} = \frac{\partial \mathrm{L}}{\partial \dot{x}} = (\mathrm{M} + m)\dot{x} + m\ell \cos\theta\, \dot{\theta} = \frac{d}{dt}\left((\mathrm{M} + m)x + m\ell \sin\theta\right) \ .$$

One can use this to solve for $\dot{x}$ and inserting back into the equation for the energy, gives a family of curves in the $(\theta, \dot{\theta})$ plane for different values of E, P and different ratios of M/m. We expect that if $\mathrm{M} \gg m$ then we are basically back to the previous case where the pivot was fixed, whereas if $\mathrm{M} \approx m$ or $\mathrm{M} \ll m$ we get markedly different behaviour.

(a) M $\gg m$



(b) M $\approx m$



(c) M $\ll m$

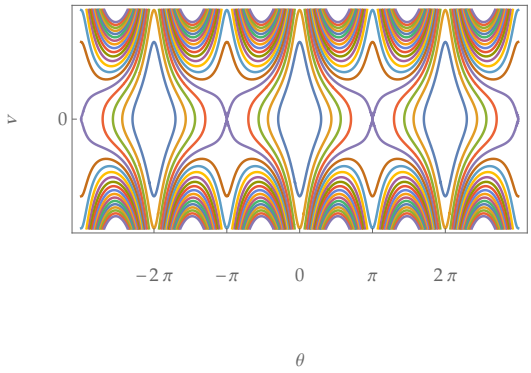


Figure 11: Physical trajectories for planar pendulum with moving pivot

7 Noether's theorem and conservation laws

> Meine Herren, der Senat ist doch keine Badeanstalt.[a] — David Hilbert

[a] Gentlemen, but the Senate is not a bathhouse!

We have seen already several examples of conservation laws. One of the examples arises when the Lagrangian does not depend explicitly on one of the coordinates, say, z; that is, when $\frac{\partial L}{\partial z} = 0$. In this case the conjugate momentum $\frac{\partial L}{\partial \dot{z}}$ is constant along physical trajectories. An equivalent rephrasing of the condition $\frac{\partial L}{\partial z} = 0$ is to say that L is invariant under constant shifts in z; that is, $L(z, \dot{z}, \dots) = L(z + s, \dot{z}, \dots)$ for all $s \in \mathbb{R}$. This is an example of a *continuous symmetry* of the lagrangian and Noether's celebrated theorem says that associated to every continuous symmetry (which need not necessarily be a constant shift of one of the variables) there is associated a conserved quantity. In this section we will prove this theorem as well as a slight generalisation.

7.1 Mappings between functionals

To set the context, let $I[x] = \int_0^1 L(x, \dot{x}, t)dt$ be an action for C^1 curves $x : [0, 1] \to \mathbb{R}^n$ with $x(0) = P$ and $x(1) = Q$, for some points $P, Q \in \mathbb{R}^n$.

Let $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ be a twice continuously differentiable (i.e., C^2) function. Then the composition $y = \varphi \circ x : [0, 1] \to \mathbb{R}^n$ is again a C^1 curve in $\mathbb{R}^n$ from the point $\varphi(P)$ to the point $\varphi(Q)$. If x extremises I, there is no reason to believe that y should too. Indeed, suppose that I is the arclength functional, which is extremised by straight lines, and let x be the straight line segment from P to Q. But then for most φ, y is not going to be the straight line segment from $\varphi(P)$ to $\varphi(Q)$. Hence y will not extremise the arclength functional.

But we can understand the relation between the action extremised by x and that extremised by y. Let us assume that φ is a C^2 **diffeomorphism**; that is, not just is φ C^2, but it has an inverse φ^{-1} which is also C^2. Then let $L(y, \dot{y}, t)$ be a lagrangian for $y(t) = \varphi(x(t))$ and let

$$K(x, \dot{x}, t) = L(\varphi(x(t)), D\varphi_{x(t)}\dot{x}(t), t) \ .$$

Notice that we have used the chain rule to write

$$\dot{y}(t) = \frac{d}{dt}\varphi(x(t)) = D\varphi_{x(t)}\dot{x}(t) \ ,$$

where $D\varphi_{x(t)}$ is the derivative matrix of φ evaluated at $x(t)$. In this case it is an $n \times n$ matrix.

Lemma 7.1. *Let* $I[x] = \int_0^1 K(x, \dot{x}, t)dt$ *and* $J[y] = \int_0^1 L(y, \dot{y}, t)$ *where* K *and* L *are defined as above. Then* x *extremises* I *if and only if* $y = \varphi \circ x$ *extremises* J. *In other words,* φ *sets up a bijective correspondence between extremals of* I *and extremals of* J.

Proof. We have that

$$\mathrm{K}(x, \dot{x}, t) = \mathrm{L}(\varphi(x(t)), \mathrm{D}\varphi_{x(t)}\dot{x}(t), t) \ ,$$

whence, using the Einstein summation convention,[2]

$$\frac{\partial \mathrm{K}}{\partial x^i} = \frac{\partial \mathrm{L}}{\partial y^j}\frac{\partial \varphi^j}{\partial x^i} + \frac{\partial \mathrm{L}}{\partial \dot{y}^j}\frac{\partial^2 \varphi^j}{\partial x^i \partial x^k}\dot{x}^k \qquad \text{and} \qquad \frac{\partial \mathrm{K}}{\partial \dot{x}^i} = \frac{\partial \mathrm{L}}{\partial \dot{y}^j}\frac{\partial \varphi^j}{\partial x^i} \ .$$

The Euler–Lagrange equation for K is then

$$\begin{aligned}
0 &= \frac{d}{dt}\frac{\partial \mathrm{K}}{\partial \dot{x}^i} - \frac{\partial \mathrm{K}}{\partial x^i} \\
&= \frac{d}{dt}\left(\frac{\partial \mathrm{L}}{\partial \dot{y}^j}\frac{\partial \varphi^j}{\partial x^i}\right) - \frac{\partial \mathrm{L}}{\partial y^j}\frac{\partial \varphi^j}{\partial x^i} - \frac{\partial \mathrm{L}}{\partial \dot{y}^j}\frac{\partial^2 \varphi^j}{\partial x^i \partial x^k}\dot{x}^k \\
&= \left(\frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{y}^j} - \frac{\partial \mathrm{L}}{\partial y^j}\right)\frac{\partial \varphi^j}{\partial x^i} + \frac{\partial \mathrm{L}}{\partial \dot{y}^j}\frac{d}{dt}\frac{\partial \varphi^j}{\partial x^i} - \frac{\partial \mathrm{L}}{\partial \dot{y}^j}\frac{\partial^2 \varphi^j}{\partial x^i \partial x^k}\dot{x}^k \\
&= \left(\frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{y}^j} - \frac{\partial \mathrm{L}}{\partial y^j}\right)\frac{\partial \varphi^j}{\partial x^i} + \frac{\partial \mathrm{L}}{\partial \dot{y}^j}\left(\frac{\partial^2 \varphi^j}{\partial x^k \partial x^i} - \frac{\partial^2 \varphi^j}{\partial x^i \partial x^k}\right)\dot{x}^k \ .
\end{aligned}$$

But the last term in the right-hand side vanishes because φ is twice continuously differentiable, leaving us with

$$\frac{d}{dt}\frac{\partial \mathrm{K}}{\partial \dot{x}^i} - \frac{\partial \mathrm{K}}{\partial x^i} = \left(\frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{y}^j} - \frac{\partial \mathrm{L}}{\partial y^j}\right)\frac{\partial \varphi^j}{\partial x^i} \ ,$$

whence using that φ is invertible and hence so is the derivative matrix $\frac{\partial \varphi^j}{\partial x^i}$, we see that x solves the Euler–Lagrange equation for K if and only if y solves the Euler–Lagrange equation for L. □

We say that φ is a **symmetry** of the lagrangian L if

$$\mathrm{L}(x, \dot{x}, t) = \mathrm{L}(y, \dot{y}, t) \ ;$$

that is, if $\mathrm{K} = \mathrm{L}$. Equivalently one says that L is **invariant** under φ. Then it follows from the lemma that φ does take extremals to extremals.

The converse of this lemma is *not* true. There are transformations taking extremals to extremals which are not symmetries of the lagrangian. Indeed, it is now a good time to point out that the lagrangian admits some ambiguity. For mechanical systems with $\mathrm{L} = \mathrm{T} - \mathrm{V}$, we already saw that V is defined up to an additive constant, since it is only its gradient which appears in Newton's equations. Therefore we can always add a constant to L without altering the resulting Euler–Lagrange equation. But more is true. Suppose that we add to L a total t-derivative, say,

$$\mathrm{L} \mapsto \mathrm{L} + \frac{d\mathrm{K}}{dt} \ .$$

[2] This convention states that one is to sum over repeated indices.

Then the action changes by a constant which is independent of the path: namely,

$$\mathrm{I}[x] = \int_0^1 \mathrm{L}dt \mapsto \mathrm{I}[x] + \mathrm{K}\Big|_0^1 ,$$

and therefore it does not alter the Euler–Lagrange equations. Hence if a transformation changes the lagrangian by a total t-derivative, it will again send solutions of the Euler–Lagrange equation to solutions.

7.2 Noether's Theorem (first version)

Going back to our discussion of symmetries, now consider not one function φ but a one-parameter family $\varphi_s : \mathbb{R}^n \to \mathbb{R}^n$ of C^2 functions, defined for all $s \in \mathbb{R}$ and depending differentiably on s. Assume moreover that this family satisfies the following two properties:

(D1) $\varphi_0(x) = x$ for all $x \in \mathbb{R}^n$; and

(D2) $\varphi_s \circ \varphi_t = \varphi_{s+t}$, for all $s, t \in \mathbb{R}$.

Exercise 7.1. *Show that these properties imply that φ_s is invertible, with inverse $\varphi_s^{-1} = \varphi_{-s}$. So in particular, φ_s is a C^2 diffeomorphism of $\mathbb{R}^n$. Show that the family $\{\varphi_s\}$ defines a group of transformations isomorphic to $(\mathbb{R}, +)$.*

The family $\{\varphi_s\}$ is called a **one-parameter subgroup of C^2 diffeomorphisms** of $\mathbb{R}^n$.

The following theorem tells us what happens when a lagrangian is invariant under a one-parameter group of diffeomorphisms; that is, when it is invariant under φ_s for all s.

Theorem 7.2 (Noether's Theorem)**.** *Let $\mathrm{I}[x] = \int_0^1 \mathrm{L}(x, \dot{x}, t)dt$ be an action for curves $x : [0, 1] \to \mathbb{R}^n$, and let L be invariant under a one-parameter group of diffeomorphisms $\{\varphi_s\}$. Then the* ***Noether charge*** N*, defined by*

$$\boxed{\mathrm{N}(x, \dot{x}, t) = \sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial \dot{x}^i} \frac{\partial \varphi_s^i(x)}{\partial s}\Big|_{s=0}} ,$$

is conserved; that is, $d\mathrm{N}/dt = 0$ along physical trajectories.

Proof. Let $x(t)$ be a solution of the Euler–Lagrange equation for L. Then by the above lemma, so is $y(s, t) := \varphi_s(x(t))$ for every s; in other words,

$$\frac{\partial \mathrm{L}}{\partial y^i} = \frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{y}^i} .$$

Because L is invariant under φ_s for every s, $\mathrm{L}(y, \dot{y}, t)$ does not depend on s. Taking the derivative with respect to s, we obtain

$$0 = \frac{d}{ds}\mathrm{L}(y, \dot{y}, t) = \sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial y^i}\frac{\partial y^i}{\partial s} + \sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial \dot{y}^i}\frac{\partial \dot{y}^i}{\partial s} .$$

Using the Euler–Lagrange equation, we can rewrite this as

$$\begin{aligned} 0 &= \sum_{i=1}^{n} \frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{y}^i}\frac{\partial y^i}{\partial s} + \sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial \dot{y}^i}\frac{\partial \dot{y}^i}{\partial s} \\ &= \frac{d}{dt}\sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial \dot{y}^i}\frac{\partial y^i}{\partial s} \,. \end{aligned}$$

Finally we evaluate at $s = 0$, using that $y(0,t) = x(t)$ by property (D1) above, to arrive at

$$\frac{d}{dt}\sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial \dot{x}^i}\frac{\partial y^i}{\partial s}\bigg|_{s=0} = 0 \,,$$

which completes the proof. □

As an example, consider a lagrangian $\mathrm{L} = \mathrm{L}(\dot{x})$, where $x : [0,1] \to \mathbb{R}$, which does not depend explicitly on x. This means that L is invariant under the one-parameter group of diffeomorphisms $\varphi_s(x) = x + s$. According to Noether's theorem the momentum $\partial \mathrm{L}/\partial \dot{x}$ is conserved.

Exercise 7.2 (Conservation of angular momentum). *Let* $\mathrm{L} = \frac{1}{2}m\|\dot{x}\|^2 - \mathrm{V}(x)$ *be a lagrangian for plane curves* $x : [0,1] \to \mathbb{R}^2$. *Assume that* V *only depends on* $\|x\|$. *Show that* L *is invariant under the one-parameter symmetry group* $\varphi_s : \mathbb{R}^2 \to \mathbb{R}^2$ *defined by*

$$\varphi_s(x) = (x^1 \cos s - x^2 \sin s, x^1 \sin s + x^2 \cos s) \,.$$

Find the expression for the Noether charge associated to this symmetry.
(Answer: You should find $\mathrm{N} = m(x^1\dot{x}^2 - x^2\dot{x}^1)$.)

Notice that Noether's theorem is even more general. Suppose that the one parameter family of diffeomorphisms $\{\varphi_s\}$ does not leave the lagrangian invariant, but that instead of vanishing,

$$\frac{d}{ds}\mathrm{L}(y,\dot{y},t) = \frac{d}{dt}\mathrm{K}_s(y,\dot{y},t) \,.$$

Then the same proof *mutatis mutandis* of Noether's Theorem given above shows now that

$$\mathrm{N}(x,\dot{x},t) = \sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial \dot{x}^i}\frac{\partial y^i}{\partial s}\bigg|_{s=0} - \mathrm{K}_0(x,\dot{x},t) \,,$$

is constant along physical trajectories.

Before we tackle a more general still version of Noether's Theorem, it is perhaps useful to observe that whereas the hypothesis of the theorem require the action functional to be invariant under a continuous one-parameter family of symmetries, the conclusion of the theorem results in a conserved charge which only depends on $\frac{\partial y^i}{\partial s}\Big|_{s=0}$ which would be the first-order term in the Taylor expansion of $\varphi_s(x(t))$ about $s = 0$. This suggests that perhaps we should be able to weaken the hypothesis by considering only how the action changes to first

order in s. This is indeed the case. We claim that it is enough to check that the action is infinitesimally invariant. This follows from existence and uniqueness of initial value problems; although it is special case of the *Lie correspondence*, a cornerstone in the theory of Lie groups.

Let us sketch a proof for the case at hand. Suppose that $\{\varphi_s\}$ is a one-parameter group of C^2 diffeomorphisms of $\mathbb{R}^n$. Then φ_s will act on the space $\mathcal{F}$ of functionals of C^1 curves $x : [0,1] \to \mathbb{R}^n$ as follows. A typical functional in $\mathcal{F}$ is given by integrating a lagrangian function:

$$I = \int_0^1 L(t, x(t), \dot{x}(t))dt$$

and the functional $\varphi_s I$ is given by

$$\varphi_s I = \int_0^1 L(t, y(t), \dot{y}(t))dt,$$

where $y(t) = \varphi_s(x(t))$. Fix some functional $I \in \mathcal{F}$ and consider $\varphi_s I$. This defines a curve in $\mathcal{F}$ via I. The velocity of this curve at the point $\varphi_s I$ on the curve is given by

$$V(\varphi_s I) = \frac{d}{ds}(\varphi_s I),$$

which defines a vector field along the curve. Under suitable conditions, the vector field V will be Lipschitz continuous in s and therefore the initial value problem:

$$\frac{d}{ds} I(s) = V(I(s)) \qquad \text{and} \qquad I(0) = I$$

will have a unique solution. Assume that is the case and suppose that $V(I) = 0$. Then the constant curve $I(s) = I$ for all s solves the initial value problem and hence is the unique solution. In other words,

$$\frac{d}{ds}(\varphi_s I)\Big|_{s=0} = 0 \quad \Longrightarrow \quad \varphi_s I = I \quad \text{for all } s.$$

This observation allows us to strengthen the statement of Noether's Theorem. Let $\zeta = \frac{\partial \varphi_s(x)}{\partial s}\Big|_{s=0}$. Then Noether's theorem says that if

$$\sum_i \frac{\partial L}{\partial x^i}\zeta^i + \sum_i \frac{\partial L}{\partial \dot{x}^i}\dot{\zeta}^i = \frac{dK}{dt},$$

then the Noether charge

$$N = \sum_i \frac{\partial L}{\partial \dot{x}^i}\zeta^i - K$$

is conserved.

7.3 A more general Noether's Theorem

Noether's theorem admits another generalisation where we also transform the independent variable.

We consider functionals which are defined by lagrangians $\mathrm{L}(x, \dot{x}, t)$ and such that they are invariant under one-parameter family of diffeomorphisms of $\mathbb{R}^n \times \mathbb{R}$ sending (x, t) to $(\bar{x}(x, t, s), \bar{t}(x, t, s))$. This means, in particular, that

$$\bar{x}^j = x^j + \zeta^j(t, x)s + \mathrm{O}(s^2) \qquad \text{and} \qquad \bar{t} = t + \tau(t, x)s + \mathrm{O}(s^2)\,,$$

so that $\zeta^j = \left.\frac{\partial \bar{x}^j}{\partial s}\right|_{s=0}$ and $\tau = \left.\frac{\partial \bar{t}}{\partial s}\right|_{s=0}$. Other partial derivatives we can read off from the above are

$$\left.\frac{\partial \bar{x}^j}{\partial x^k}\right|_{s=0} = \delta^j_k \qquad \left.\frac{\partial \bar{x}^j}{\partial t}\right|_{s=0} = 0 \qquad \left.\frac{\partial \bar{t}}{\partial x^k}\right|_{s=0} = 0 \qquad \left.\frac{\partial \bar{t}}{\partial t}\right|_{s=0} = 1\,,$$

where δ is the Kronecker δ, defined by

$$\delta^j_k = \begin{cases} 1, & j = k \\ 0, & j \neq k\,. \end{cases}$$

We can also read off the following second derivatives:

$$\left.\frac{\partial^2 \bar{x}^j}{\partial s \partial x^k}\right|_{s=0} = \frac{\partial \zeta^j}{\partial x^k} \qquad \left.\frac{\partial^2 \bar{t}}{\partial s \partial x^k}\right|_{s=0} = \frac{\partial \tau}{\partial x^k}$$
$$\left.\frac{\partial^2 \bar{x}^j}{\partial s \partial t}\right|_{s=0} = \frac{\partial \zeta^j}{\partial t}\,, \qquad \left.\frac{\partial^2 \bar{t}}{\partial s \partial t}\right|_{s=0} = \frac{\partial \tau}{\partial t}\,.$$

In this more general case, we say that $(x, t) \mapsto (\bar{x}, \bar{t})$ is a **symmetry** if

$$\mathrm{L}(\bar{x}, \frac{d\bar{x}}{d\bar{t}}, \bar{t})\frac{d\bar{t}}{dt} = \mathrm{L}(x, \frac{dx}{dt}, t)\,. \tag{30}$$

If this is case, notice that the right-hand side does not depend on s, hence neither does the left-hand side. So we take the derivative with respect to s of both sides of the equation and arrive at

$$\left(\frac{\partial \mathrm{L}}{\partial \bar{t}}\frac{\partial \bar{t}}{\partial s} + \sum_k \frac{\partial \mathrm{L}}{\partial \bar{x}^k}\frac{\partial \bar{x}^k}{\partial s} + \sum_k \frac{\partial \mathrm{L}}{\partial \dot{\bar{x}}^k}\frac{\partial}{\partial s}\frac{d\bar{x}^k}{d\bar{t}}\right)\frac{d\bar{t}}{dt} + \mathrm{L}\frac{\partial}{\partial s}\frac{d\bar{t}}{dt} = 0\,,$$

where we have used the shorthand $\dot{\bar{x}}^k = \frac{d\bar{x}^k}{d\bar{t}}$. We now evaluate at $s = 0$ to arrive at

$$\left(\frac{\partial \mathrm{L}}{\partial t}\tau + \sum_k \frac{\partial \mathrm{L}}{\partial x^k}\zeta^k + \sum_k \frac{\partial \mathrm{L}}{\partial \dot{x}^k}\left.\frac{\partial}{\partial s}\frac{d\bar{x}^k}{d\bar{t}}\right|_{s=0}\right)\left.\frac{d\bar{t}}{dt}\right|_{s=0} + \mathrm{L}\left.\frac{\partial}{\partial s}\frac{d\bar{t}}{dt}\right|_{s=0} = 0\,. \tag{31}$$

To make this expression useful, we need to derive expressions for

$$\left.\frac{\partial}{\partial s}\frac{d\bar{x}^k}{d\bar{t}}\right|_{s=0}\,, \qquad \left.\frac{d\bar{t}}{dt}\right|_{s=0} \qquad \text{and} \qquad \left.\frac{\partial}{\partial s}\frac{d\bar{t}}{dt}\right|_{s=0}\,.$$

Using the chain rule, we find

$$\frac{d\bar{t}}{dt} = \frac{\partial \bar{t}}{\partial t} + \sum_j \frac{\partial \bar{t}}{\partial x^j}\frac{dx^j}{dt} \,,$$

from where it follows that

$$\left.\frac{d\bar{t}}{dt}\right|_{s=0} = 1 \tag{32}$$

and

$$\begin{aligned}\left.\frac{\partial}{\partial s}\frac{d\bar{t}}{dt}\right|_{s=0} &= \left.\frac{\partial}{\partial s}\left(\frac{\partial \bar{t}}{\partial t} + \sum_j \frac{\partial \bar{t}}{\partial x^j}\frac{dx^j}{dt}\right)\right|_{s=0} \\ &= \left.\frac{\partial^2 \bar{t}}{\partial s \partial t}\right|_{s=0} + \sum_j \left.\frac{\partial^2 \bar{t}}{\partial s \partial x^j}\right|_{s=0}\frac{dx^j}{dt}\end{aligned}$$

whence

$$\left.\frac{\partial}{\partial s}\frac{d\bar{t}}{dt}\right|_{s=0} = \frac{\partial \tau}{\partial t} + \sum_j \frac{\partial \tau}{\partial x^j}\dot{x}^j = \frac{d\tau}{dt} \,. \tag{33}$$

To find an expression for $\left.\frac{\partial}{\partial s}\frac{d\bar{x}^k}{d\bar{t}}\right|_{s=0}$, we use the chain rule again in the form

$$\frac{d\bar{x}^k}{dt} = \frac{d\bar{x}^k}{d\bar{t}}\frac{d\bar{t}}{dt} \,.$$

Evaluating at $s = 0$, we have

$$\frac{dx^k}{dt} = \left.\frac{d\bar{x}^k}{d\bar{t}}\right|_{s=0} \,,$$

and, taking its s derivative at $s = 0$, we have

$$\left.\frac{\partial}{\partial s}\frac{d\bar{x}^k}{dt}\right|_{s=0} = \left.\frac{\partial}{\partial s}\frac{d\bar{x}^k}{d\bar{t}}\right|_{s=0}\left.\frac{d\bar{t}}{dt}\right|_{s=0} + \left.\frac{d\bar{x}^k}{d\bar{t}}\right|_{s=0}\left.\frac{\partial}{\partial s}\frac{d\bar{t}}{dt}\right|_{s=0} \,.$$

The RHS becomes, after using the above results,

$$\text{RHS} = \left.\frac{\partial}{\partial s}\frac{d\bar{x}^k}{d\bar{t}}\right|_{s=0} + \frac{dx^k}{dt}\left.\frac{\partial}{\partial s}\frac{d\bar{t}}{dt}\right|_{s=0} = \left.\frac{\partial}{\partial s}\frac{d\bar{x}^k}{d\bar{t}}\right|_{s=0} + \dot{x}^k\frac{d\tau}{dt} \,,$$

whereas to evaluate the LHS, we use the chain rule in the form

$$\frac{d\bar{x}^k}{dt} = \frac{\partial \bar{x}^k}{\partial t} + \sum_j \frac{\partial \bar{x}^k}{\partial x^j}\frac{dx^j}{dt} \,,$$

so that

$$\text{LHS} = \left.\frac{\partial}{\partial s}\left(\frac{\partial \bar{x}^k}{\partial t} + \sum_j \frac{\partial \bar{x}^k}{\partial x^j}\frac{dx^j}{dt}\right)\right|_{s=0} = \frac{\partial \zeta^k}{\partial t} + \sum_j \frac{\partial \zeta^k}{\partial x^j}\frac{dx^j}{dt} = \frac{d\zeta^k}{dt} \,.$$

Finally, we can solve for $\frac{\partial}{\partial s}\left(\frac{d\bar{x}^k}{d\bar{t}}\right)\Big|_{s=0}$ and obtain

$$\frac{\partial}{\partial s}\left(\frac{d\bar{x}^k}{d\bar{t}}\right)\Big|_{s=0} = \frac{d\zeta^k}{dt} - \dot{x}^k\frac{d\tau}{dt}\,. \tag{34}$$

Finally, we use equations (32), (33) and (34) in equation (31) to arrive at the infinitesimal version of the invariance of the functional under the symmetry:

$$\boxed{\frac{\partial \mathrm{L}}{\partial t}\tau + \sum_k \frac{\partial \mathrm{L}}{\partial x^k}\zeta^k + \sum_k \frac{\partial \mathrm{L}}{\partial \dot{x}^k}\left(\frac{d\zeta^k}{dt} - \dot{x}^k\frac{d\tau}{dt}\right) + \mathrm{L}\frac{d\tau}{dt} = 0}\,. \tag{35}$$

The ζ^k terms in equation (35) can be rewritten as

$$\sum_k \frac{\partial \mathrm{L}}{\partial x^k}\zeta^k + \sum_k \frac{\partial \mathrm{L}}{\partial \dot{x}^k}\frac{d\zeta^k}{dt} = \frac{d}{dt}\sum_k \frac{\partial \mathrm{L}}{\partial \dot{x}^k}\zeta^k + \sum_k \left(\frac{\partial \mathrm{L}}{\partial x^k} - \frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{x}^k}\right)\zeta^k\,,$$

whereas the τ terms can be rewritten using that the total t derivative of the lagrangian is given by

$$\frac{d\mathrm{L}}{dt} = \frac{\partial \mathrm{L}}{\partial t} + \sum_k \frac{\partial \mathrm{L}}{\partial x^k}\dot{x}^k + \sum_k \frac{\partial \mathrm{L}}{\partial \dot{x}^k}\dot{x}^k\,,$$

whence

$$\frac{\partial \mathrm{L}}{\partial t}\tau = \left(\frac{d\mathrm{L}}{dt} - \sum_k \frac{\partial \mathrm{L}}{\partial x^k}\dot{x}^k - \sum_k \frac{\partial \mathrm{L}}{\partial \dot{x}^k}\dot{x}^k\right)\tau\,.$$

Using this, the τ terms in equation (35) are given by

$$\frac{d}{dt}\left(\mathrm{L}\tau - \sum_k \frac{\partial \mathrm{L}}{\partial \dot{x}^k}\dot{x}^k\tau\right) - \sum_k \left(\frac{\partial \mathrm{L}}{\partial x^k} - \frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{x}^k}\right)\dot{x}^k\tau\,.$$

In summary, we can rewrite equation (35) as

$$\frac{d}{dt}\left(\mathrm{L}\tau + \sum_k \frac{\partial \mathrm{L}}{\partial \dot{x}^k}(\zeta^k - \dot{x}^k\tau)\right) + \sum_k \left(\frac{\partial \mathrm{L}}{\partial x^k} - \frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{x}^k}\right)(\zeta^k - \dot{x}^k\tau) = 0\,.$$

In other words, the **Noether charge**

$$\boxed{\mathrm{N} := \mathrm{L}\tau + \sum_k \frac{\partial \mathrm{L}}{\partial \dot{x}^k}(\zeta^k - \dot{x}^k\tau)} \tag{36}$$

is conserved along extremals; that is, along curves which obey the Euler–Lagrange equation.

Exercise 7.3 (Conservation of energy from Noether's theorem). *Derive conservation of energy from this more general Noether theorem as a consequence of the invariance of the lagrangian under* $(x,t) \mapsto (x,t+s)$.

☞ Notice that again this more general Noether's Theorem can be further generalised to allow for invariance up to total derivatives. In this case, the first-order (or infinitesimal) invariance condition (35) is modified to

$$\frac{\partial \mathrm{L}}{\partial t}\tau + \sum_k \frac{\partial \mathrm{L}}{\partial x^k}\zeta^k + \sum_k \frac{\partial \mathrm{L}}{\partial \dot{x}^k}\left(\frac{d\zeta^k}{dt} - \dot{x}^k\frac{d\tau}{dt}\right) + \mathrm{L}\frac{d\tau}{dt} = \frac{d\mathrm{K}}{dt}\,,$$

and the conserved Noether charge (36) also changes:

$$\mathrm{N} := \mathrm{L}\tau + \sum_k \frac{\partial \mathrm{L}}{\partial \dot{x}^k}(\zeta^k - \dot{x}^k\tau) - \mathrm{K}\,.$$

8 Hamilton's canonical formalism

> Hamiltonian mechanics is geometry in phase space.
>
> — Vladimir Arnol'd

Noether's Theorem says that every continuous symmetry of the lagrangian (or more generally, the action) gives rise to a conserved quantity: the associated Noether charge. One might wonder whether there is a converse to this theorem. Namely, if we manage to find, by whatever means, a conserved quantity, does it come from a continuous symmetry? As it turns out this is indeed the case, but to explain it we need to rephrase the Euler–Lagrange equation in terms of an equivalent first-order differential equation on *phase space*. In the case of one-dimensional dynamical systems, the phase space is nothing but the phase plane familiar from **HDEq**.

8.1 The canonical form of the Euler–Lagrange equation

To set the stage, let us consider a functional $\mathrm{I}[x] = \int_0^1 \mathrm{L}(x, \dot{x}, t)dt$ for C^1 curves $x : [0,1] \to \mathbb{R}^n$. Then as we know already well, x is an extremal for I if it obeys the Euler–Lagrange equation:

$$\frac{\partial \mathrm{L}}{\partial x^i} = \frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{x}^i} \,.$$

This is a second-order ODE for $x(t)$ and in **HDEq** we learn how to turn this into an equivalent first-order ODE

$$\frac{dx^i}{dt} = v^i \qquad \text{and} \qquad \frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{x}^i} = \frac{\partial \mathrm{L}}{\partial x^i} \tag{37}$$

for the variables $(x(t), v(t))$. One striking feature of this set of equations is their lack of symmetry: one equation looks much simpler than the other. The idea of the canonical (or hamiltonian) formalism is to restore some symmetry in the equations by changing variables from (x, v) to another set of (so-called *canonical*) variables.

To understand how one might go about this, let us consider the case of one-dimensional motion $x : [0,1] \to \mathbb{R}$ and $\mathrm{L}(x, v) = \frac{1}{2}mv^2 - \mathrm{V}(x)$. We would like the second equation to resemble the first, namely: $\dot{x} = v$. This suggests introducing the momentum p conjugate to x, defined by

$$p = \frac{\partial \mathrm{L}}{\partial v} = mv \,,$$

so that the second equation becomes $\dot{p} = -\frac{d\mathrm{V}}{dx}$. We now solve for v as a function of p: $v = p/m$ and then define the function

$$\mathrm{H}(x, p) = pv - \mathrm{L}(x, v) \,,$$

but where v is now a function of p; that is,

$$\mathrm{H}(x,p) = \frac{p^2}{m} - \left(\frac{p^2}{2m} - \mathrm{V}(x)\right) = \frac{p^2}{2m} + \mathrm{V}(x)\,.$$

Notice that

$$\frac{\partial \mathrm{H}}{\partial x} = \frac{d\mathrm{V}}{dx} \qquad \text{and} \qquad \frac{\partial \mathrm{H}}{\partial p} = \frac{p}{m}\,,$$

whence the first-order system (37) can be written in a much more symmetrical form:

$$\frac{dx}{dt} = \frac{\partial \mathrm{H}}{\partial p} \qquad \text{and} \qquad \frac{dp}{dt} = -\frac{\partial \mathrm{H}}{\partial x}\,,$$

which are Newton's equations in **Hamilton form**. Often they are referred to as **Hamilton's equations** and the function H is called the **hamiltonian**.

☞ Notice that the relation between the lagrangian $\mathrm{L}(x,v,t)$ and the hamiltonian $\mathrm{H}(x,p,t)$ is very symmetrical. Indeed, notice the following parallel:

$\mathrm{L}(x,v,t)$	$\mathrm{H}(x,p,t)$
$p = \dfrac{\partial \mathrm{L}}{\partial v}$	$v = \dfrac{\partial \mathrm{H}}{\partial p}$
$\mathrm{H} = pv - \mathrm{L}$	$\mathrm{L} = vp - \mathrm{H}$

The two sides are related by the simultaneous substitutions $\mathrm{L} \leftrightarrow \mathrm{H}$ and $v \leftrightarrow p$. And hence one says that H is the **Legendre transform** of L, and viceversa! In other words, the Legendre transform is involutive.

At this moment it may not be entirely clear what, other than the aesthetic pleasure of the equations themselves, this reformulation has gained us, but before we go into that, let us make one remark and let us plough ahead with the general formalism.

The remark is that you may recognise these equations as reminiscent of a **gradient flow**, which would take the form

$$\frac{dx}{dt} = \frac{\partial \mathrm{U}}{\partial x} \qquad \text{and} \qquad \frac{dp}{dt} = \frac{\partial \mathrm{U}}{\partial p}\,,$$

for some function $\mathrm{U}(x,p)$. They are indeed similar—both being given by the derivative of a function—but they are *crucially* different: whereas solutions of the gradient flow are perpendicular to the level sets of the function U, solutions of Hamilton's equations are tangent to the level sets of the function H. To see this notice that we can rewrite Hamilton's equations as

$$\frac{d}{dt}\begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial \mathrm{H}}{\partial x} \\ \frac{\partial \mathrm{H}}{\partial p} \end{pmatrix}\,.$$

So it would be a gradient flow but for the presence of the matrix $\mathrm{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. This matrix can be interpreted as a rotation by $\pi/2$, which is why the hamilto-

nian flow is tangent and not normal to the level sets of the hamiltonian function H.

☞ The matrix J, when thought of as a bilinear form on $\mathbb{R}^2$, defines a **symplectic structure** on $\mathbb{R}^2$. Hamilton's equations are thus said to correspond to a *symplectic* gradient flow.

Let us now return to the case of a general lagrangian $L(x,v,t)$ for $x,v \in \mathbb{R}^n$. We can again define the conjugate momentum p_i to x^i by

$$p_i = \frac{\partial L}{\partial v^i} ,$$

but we need to ensure that we can invert this equation to solve for v^i in terms of the $2n+1$ variables $x^1, x^2, \dots, x^n, p_1, p_2, \dots, p_n, t$. The *Implicit Function Theorem* guarantees that the solution set of the n equations $p_i = \frac{\partial L}{\partial v^i}$ is the graph of a function $v^i = v^i(t,x,p)$ around any point where the matrix with entries $\frac{\partial^2 L}{\partial v^i \partial v^j}$ is invertible. A lagrangian L satisfying this property for all (x,v,t) is said to be **regular** (or **nondegenerate**).

☞ To use the Implicit Function Theorem as in the **SVC notes**, consider the function $F : \mathbb{R}^{3n+1} \to \mathbb{R}^n$ defined by

$$F_i(t,x,v,p) = \frac{\partial L}{\partial v^i} - p_i .$$

The Implicit Function Theorem says that the subset of $\mathbb{R}^{3n+1}$ consisting of the zero locus of the function F:

$$Z = \left\{ (t,x,v,p) \in \mathbb{R}^{3n+1} \;\middle|\; \frac{\partial L}{\partial v^i} - p_i = 0 \right\}$$

can be described as the graph

$$G = \left\{ (t,x,v(t,x,p),p) \;\middle|\; (t,x,p) \in \mathbb{R}^{2n+1} \right\}$$

of a function $v : \mathbb{R}^{2n+1} \to \mathbb{R}^n$ provided that the derivative matrix

$$\frac{\partial F_i}{\partial v^j} = \frac{\partial^2 L}{\partial v^j \partial v^i}$$

is invertible. More precisely, we can describe Z locally as such a graph in in the neighbourhood of any point where the above matrix is invertible.

We shall assume in what follows that L is regular. In practice, of course, when working with an explicit lagrangian, one can check whether this is the case. For a regular lagrangian, the Implicit Function Theorem guarantees that we can invert $p = \frac{\partial L}{\partial v}$ to write $v^i = v^i(x,p,t)$ and thus we can define the **hamiltonian** $H(x,p,t)$ by

$$H(x,p,t) = \sum_{i=1}^{n} v^i p_i - L(x,v,t) ,$$

but where $v^i = v^i(x, p, t)$.

We will now translate the Euler–Lagrange equation in first-order form to Hamilton form. This requires rewriting the equation in terms of the canonical variables (x, p). We could do this by writing L in terms of H and plugging the result into the Euler–Lagrange equation, but that would require calculating

$$\frac{\partial \mathrm{L}}{\partial x^i} \qquad \text{and} \qquad \frac{\partial \mathrm{L}}{\partial v^i} \,,$$

where by these expressions we mean the partial derivatives of L with respect to x^i and v^i keeping all the v^i and all the x^i fixed, respectively. (The standard notation for partial derivatives suffers from the drawback that it does not specify which variables are being kept fixed.) Experience tells us that in performing calculations of this sort we are bound to make errors, whence we prefer to work instead with total derivatives, which are independent of the coordinates used.


Students of **Geom** might recognise this as the simplest case of the formalism of *differential forms*.

Let us calculate the total derivative of H using its definition:

$$\begin{aligned} d\mathrm{H} &= d\left(\sum_{i=1}^{n} p_i v^i - \mathrm{L}\right) \\ &= \sum_{i=1}^{n} (dp_i v^i + p_i dv^i) - \left(\frac{\partial \mathrm{L}}{\partial t} dt + \sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial x^i} dx^i + \sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial v^i} dv^i\right) \\ &= \sum_{i=1}^{n} dp_i v^i - \frac{\partial \mathrm{L}}{\partial t} dt - \sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial x^i} dx^i \,, \end{aligned}$$

where we have used that $p_i = \frac{\partial \mathrm{L}}{\partial v^i}$, allowing us (thankfully!) to cancel two terms. We can then read off the following:

$$\frac{\partial \mathrm{H}}{\partial t} = -\frac{\partial \mathrm{L}}{\partial t} \,, \qquad \frac{\partial \mathrm{H}}{\partial p_i} = v^i \qquad \text{and} \qquad \frac{\partial \mathrm{H}}{\partial x^i} = -\frac{\partial \mathrm{L}}{\partial x^i} \,.$$

This then allows us to write the first-order version (37) of the Euler–Lagrange equations in **canonical** (or **hamiltonian**) form:

$$\boxed{\frac{dx^i}{dt} = \frac{\partial \mathrm{H}}{\partial p_i} \qquad \text{and} \qquad \frac{dp_i}{dt} = -\frac{\partial \mathrm{H}}{\partial x^i}} \tag{38}$$

We call these **Hamilton's equations**.

8.2 Conserved quantities and the Poisson bracket

We are now ready to reap the benefits of translating the Euler–Lagrange equations to canonical form.

Suppose that L does not depend explicitly on t; that is, $\frac{\partial \mathrm{L}}{\partial t} = 0$. Then as we saw above, $\frac{\partial \mathrm{H}}{\partial t} = 0$ and hence

$$\begin{aligned}\frac{d\mathrm{H}}{dt} &= \sum_{i=1}^{n} \frac{\partial \mathrm{H}}{\partial x^i}\frac{dx^i}{dt} + \sum_{i=1}^{n} \frac{\partial \mathrm{H}}{\partial p_i}\frac{dp_i}{dt} \\ &= \sum_{i=1}^{n} \left(\frac{\partial \mathrm{H}}{\partial x^i}\frac{\partial \mathrm{H}}{\partial p_i} + \frac{\partial \mathrm{H}}{\partial p_i}\left(-\frac{\partial \mathrm{H}}{\partial x^i}\right)\right) = 0\,,\end{aligned}$$

where we have used Hamilton's equations (38). Of course we knew this: the energy is conserved if the lagrangian is independent of time, but this provides a check of the formalism.

Now let us consider some differentiable function $\Phi(x,p)$ and ask how it evolves in time when x and p evolve according to (38). We calculate as we did for H above,

$$\begin{aligned}\frac{d\Phi}{dt} &= \sum_{i=1}^{n} \left(\frac{\partial \Phi}{\partial x^i}\frac{dx^i}{dt} + \frac{\partial \Phi}{\partial p_i}\frac{dp_i}{dt}\right) \\ &= \sum_{i=1}^{n} \left(\frac{\partial \Phi}{\partial x^i}\frac{\partial \mathrm{H}}{\partial p_i} - \frac{\partial \Phi}{\partial p_i}\frac{\partial \mathrm{H}}{\partial x^i}\right) = [\Phi, \mathrm{H}]\,,\end{aligned}$$

where we have introduced the **Poisson bracket**

$$\boxed{[f,g] := \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial x^i}\right)} \tag{39}$$

of any two differentiable functions $f, g : \mathbb{R}^{2n} \to \mathbb{R}$ of (x,p).

Please note that in the literature one often sees the notation $\{f,g\}$ for the Poisson bracket. This is particularly the case in the quantum mechanical literature in order to distinguish it from the commutator of operators.

Therefore we have shown that

$$\frac{d\Phi}{dt} = 0 \iff [\Phi, \mathrm{H}] = 0\,.$$

The condition on the right-hand side is often spoken as Φ *(Poisson-)commutes with* H.

If Φ were to depend explicitly on t, then the same calculation as above would show that

$$\frac{d\Phi}{dt} = \frac{\partial \Phi}{\partial t} + [\Phi, \mathrm{H}]\,.$$

Notice that it is possible for a function which depends explicitly on t to be conserved. For such a function Φ, $\frac{\partial \Phi}{\partial t} = -[\Phi, \mathrm{H}]$, so that it does not Poisson-commute with H.

Proposition 8.1. *The Poisson bracket obeys the following properties:*

(P0) $[-,-]$ *is bilinear,*

(P1) $[f,g] = -[g,f]$,

(P2) $[f,gh] = [f,g]h + g[f,h]$, *and*

(P3) $[f,[g,h]] = [[f,g],h] + [g,[f,h]]$,

for all $f,g,h : \mathbb{R}^{2n} \to \mathbb{R}$ *(at least twice) differentiable functions of* (x,p).

Proof. The first property (P1) is obvious from the definition, whereas (P2) and one half of (P0) follows from properties of the first-order differential operator

$$[f,-] = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x^i} \frac{\partial}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial}{\partial x^i} \right) .$$

The other half of (P0) follows from the first half and from (P1). Finally, property (P3) is a straight-forward, if somewhat lengthy, calculation using the definition. □

Exercise 8.1. *Supply the missing details to the proof of the proposition.*

This proposition is very important. It says that the smooth (i.e., infinitely differentiable) functions of (x,p), form a *Lie algebra* under the Poisson bracket. (The reason we require infinite differentiability is that every time we take the Poisson bracket of two functions, we lower the order of differentiability. So only infinite differentiability guarantees that no matter how many times we take Poisson brackets, we never leave the space of functions.)
Property (P3) is called the **Jacobi identity** of the Poisson bracket and historically made its first appearance precisely in this context. Moreover, together with the commutative multiplication of functions, the smooth functions form a *Poisson algebra*. Such algebras appear in many guises in Mathematics, particularly in the context of deformations of algebraic structures, of which quantum mechanics can be viewed as one instance.

The above proposition has a wonderful consequence. (Pay attention!) Suppose that Φ_1, Φ_2 are two functions of (x,p) which Poisson-commute with the hamiltonian H; that is, $[\Phi_i, \mathrm{H}] = 0$ for $i = 1,2$. Suppose that they do not depend explicitly on t, so that, as seen above, they are conserved along physical trajectories. The Jacobi identity of the Poisson bracket then says that

$$[[\Phi_1, \Phi_2], \mathrm{H}] = [\Phi_1, [\Phi_2, \mathrm{H}]] - [\Phi_2, [\Phi_1, \mathrm{H}]] = 0 ,$$

whence their Poisson bracket $[\Phi_1, \Phi_2]$ is also conserved. In other words, the Poisson bracket gives us a way to generate new conserved quantities from old.

Exercise 8.2. *Show that if* Φ_1, Φ_2 *are conserved, so is their Poisson bracket* $[\Phi_1, \Phi_2]$ even *if they depend explicitly on* t.

We can now finally explain why this formulation provides a sort of converse to Noether's Theorem. Remember that Noether's Theorem says that a continuous symmetry of the lagrangian gives rise to a conserved quantity. Now let us start with a conserved quantity, namely a function $\Phi(x,p)$ which Poisson-commutes with the hamiltonian, and consider the Poisson bracket of Φ with x^i and p_i:

$$[x^i,\Phi] = \frac{\partial\Phi}{\partial p_i} \qquad \text{and} \qquad [p_i,\Phi] = -\frac{\partial\Phi}{\partial x^i}\,.$$

This defines a *vector field* on the **phase space** $\mathbb{R}^{2n}$ which we may integrate to find through every point a solution of the differential equation

$$\frac{dx^i}{ds} = [x^i,\Phi] = \frac{\partial\Phi}{\partial p_i} \qquad \text{and} \qquad \frac{dp_i}{ds} = [p_i,\Phi] = -\frac{\partial\Phi}{\partial x^i}\,.$$

In other words, the theorem of existence and uniqueness of initial value problems guarantees the existence of a curve $(x(s),p(s))$ defined for $s\in(-\varepsilon,\varepsilon)$ such that $(x(0),p(0))$ is any prescribed point and such that

$$\frac{dx^i}{ds} = \left.\frac{\partial\Phi}{\partial p_i}\right|_{(x(s),p(s))} \qquad \text{and} \qquad \frac{dp_i}{ds} = -\left.\frac{\partial\Phi}{\partial x^i}\right|_{(x(s),p(s))}\,.$$

Along such curves, the hamiltonian does not change, since

$$\begin{aligned}\frac{d}{ds}\mathrm{H}(x(s),p(s)) &= \sum_{i=1}^{n}\left(\frac{\partial\mathrm{H}}{\partial x^i}\frac{dx^i}{ds} + \frac{\partial\mathrm{H}}{\partial p_i}\frac{dp_i}{ds}\right)\\ &= \sum_{i=1}^{n}\left(\frac{\partial\mathrm{H}}{\partial x^i}\frac{\partial\Phi}{\partial p_i} + \frac{\partial\mathrm{H}}{\partial p_i}\left(-\frac{\partial\Phi}{\partial x^i}\right)\right)\\ &= [\mathrm{H},\Phi] = -[\Phi,\mathrm{H}] = 0\,.\end{aligned}$$

So we have a continuous symmetry of Hamilton's equations, which then takes solutions to solutions.

The caveat is that this may not actually extend to a one-parameter family of diffeomorphisms of the phase space; although in practice for most "reasonable" functions Φ, $[-,\Phi]$ will generate a one-parameter family of diffeomorphisms. This is the question of completeness of the "hamiltonian vector field" $[-,\Phi]$. Sufficient conditions guaranteeing the completeness of hamiltonian vector fields are known: see, e.g., [3, 4].

As an example, let us consider the case of a particle of mass m moving in the plane subject to a (conservative) central force field:

$$\mathrm{L}(x,v) = \tfrac{1}{2}m\|v\|^2 - \mathrm{V}(\|x\|)\,,$$

where $x,v\in\mathbb{R}^2$. The conjugate momenta are given by

$$p_i = \frac{\partial\mathrm{L}}{\partial v^i} = mv^i \implies v^i = \frac{p_i}{m}\,,$$

and hence the hamiltonian $\mathrm{H}(x,p)$ is given by

$$\mathrm{H}(x,p) = \frac{\|p\|^2}{2m} + \mathrm{V}(\|x\|) \ .$$

We found earlier that the angular momentum

$$\mathrm{J} = x^1 p_2 - x^2 p_1$$

is conserved.

Exercise 8.3. *Show again that* J *is conserved, but this time by showing that* $[\mathrm{J},\mathrm{H}] = 0$.

Let us compute the Poisson bracket of J with the canonical variables:

$$[x^1,\mathrm{J}] = -x^2 \qquad [x^2,\mathrm{J}] = x^1 \qquad [p_1,\mathrm{J}] = -p_2 \qquad [p_2,\mathrm{J}] = p_1 \ ,$$

which trivially integrates (being a linear vector field) to

$$\begin{aligned} x^1(s) &= x^1(0)\cos s - x^2(0)\sin s & \qquad p_1(s) &= p_1(0)\cos s - p_2(0)\sin s \\ x^2(s) &= x^2(0)\cos s + x^1(0)\sin s & \qquad p_2(s) &= p_2(0)\cos s + p_1(0)\sin s \ , \end{aligned}$$

which agrees with the continuous symmetry found earlier in Exercise 7.2.

Exercise 8.4. *Prove that if* $\mathrm{V}(\|x\|) = \frac{1}{2}k\|x\|^2$, *then the following functions, for* $i = 1,2$, *are also conserved:*

$$\mathrm{J}_{ij} = \frac{1}{m}p_i p_j + k x^i x^j \ ;$$

i.e., show that $[\mathrm{J}_{ij},\mathrm{H}] = 0$. *Notice that* $\mathrm{J}_{12} = \mathrm{J}_{21}$ *and that* $\mathrm{J}_{11} + \mathrm{J}_{22} = 2\mathrm{H}$, *whence* $[\mathrm{J}_{11},\mathrm{H}] = 0$ *already shows that* $[\mathrm{J}_{22},\mathrm{H}] = 0$. *This means that we actually only have two new conserved quantities:* J_{12} *and* $\mathrm{J}_{11} - \mathrm{J}_{22}$, *say, in addition to* J *and* H. *Compute all the Poisson brackets between the three functions:* J, J_{12} *and* $\mathrm{J}_{11} - \mathrm{J}_{22}$. *Do you recognise the resulting Lie algebra?*

8.3 Canonical transformations

It follows from the definition (39) of the Poisson bracket that the $2n$ variables (x,p) themselves, when thought of as functions, satisfy:

$$[x^i,x^j] = 0 \qquad [p_i,p_j] = 0 \qquad \text{and} \qquad [x^i,p_j] = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \ . \tag{40}$$

Any $2n$ variables (x,p) with the above Poisson brackets are said to be **canonical conjugate variables** or **Darboux coordinates**.

Suppose we change variables to (X,P) where

$$\mathrm{X}^i = \mathrm{X}^i(x,p) \qquad \text{and} \qquad \mathrm{P}_i = \mathrm{P}_i(x,p) \ .$$

The change of variables $(x,p) \mapsto (\mathrm{X},\mathrm{P})$ is said to be a **canonical transformation** or a **symplectomorphism** if (X,P) are canonical conjugate variables; that is, if

their Poisson brackets (calculated as functions of (x,p)) obey the properties (40), or in other words,

$$[X^i, X^j] = 0 \qquad [P_i, P_j] = 0 \qquad \text{and} \qquad [X^i, P_j] = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} .$$

(It follows, in particular, that we can invert and write (x,p) as functions of (X,P).) Explicitly, using the definition (39) of the Poisson bracket, these conditions are

$$\begin{aligned} \sum_{k=1}^{n} \left(\frac{\partial X^i}{\partial x^k}\frac{\partial X^j}{\partial p_k} - \frac{\partial X^j}{\partial x^k}\frac{\partial X^i}{\partial p_k} \right) &= 0 \\ \sum_{k=1}^{n} \left(\frac{\partial P_i}{\partial x^k}\frac{\partial P_j}{\partial p_k} - \frac{\partial P_j}{\partial x^k}\frac{\partial P_i}{\partial p_k} \right) &= 0 \\ \sum_{k=1}^{n} \left(\frac{\partial X^i}{\partial x^k}\frac{\partial P_j}{\partial p_k} - \frac{\partial P_j}{\partial x^k}\frac{\partial X^i}{\partial p_k} \right) &= \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} . \end{aligned}$$


Students of **Geom** might recognise these equations as the following equality between *symplectic* 2-forms in $\mathbb{R}^{2n}$:

$$\sum_{k=1}^{n} dx^k \wedge dp_k = \sum_{k=1}^{n} dX^k \wedge dP_k .$$

Canonical transformations are important because they preserve the form of Hamilton's equations. Indeed, suppose that $K(X,P)$ is a hamiltonian for X,P and let us define $H(x,p) := K(X(x,p),P(x,p))$. Then we will prove below that

$$\dot{X}^i = [X^i, H] = \frac{\partial K}{\partial P_i} \qquad \text{and} \qquad \dot{P}_i = [P_i, H] = -\frac{\partial K}{\partial X^i} , \tag{41}$$

where the Poisson brackets are calculated as functions of (x,p). But now notice that this is nothing else but Hamilton's equations for K:

$$\dot{X}^i = [X^i, K] \qquad \text{and} \qquad \dot{P}_i = [P_i, K] ,$$

where the Poisson brackets are now calculated as functions of (X,P). We shall prove equation (41) for the special case of $n = 1$, leaving the general proof as an (easy) exercise for the reader.

So let $(x,p) \mapsto (X(x,p),P(x,p))$ be a canonical transformation and let $H(x,p) := K(X(x,p),P(x,p))$. Then using the chain rule,

$$\begin{aligned} \dot{X} = [X,H] &= \frac{\partial X}{\partial x}\frac{\partial H}{\partial p} - \frac{\partial H}{\partial x}\frac{\partial X}{\partial p} \\ &= \frac{\partial X}{\partial x}\left(\frac{\partial K}{\partial X}\frac{\partial X}{\partial p} + \frac{\partial K}{\partial P}\frac{\partial P}{\partial p} \right) - \left(\frac{\partial K}{\partial X}\frac{\partial X}{\partial x} + \frac{\partial K}{\partial P}\frac{\partial P}{\partial x} \right)\frac{\partial X}{\partial p} \\ &= \frac{\partial K}{\partial P}\left(\frac{\partial X}{\partial x}\frac{\partial P}{\partial p} - \frac{\partial P}{\partial x}\frac{\partial X}{\partial p} \right) \\ &= \frac{\partial K}{\partial P}[X,P] = \frac{\partial K}{\partial P} , \end{aligned}$$

where we have used that (X, P) are canonical conjugate variables. Similarly, one computes that

$$\begin{aligned}\dot{\mathrm{P}} = [\mathrm{P}, \mathrm{H}] &= \frac{\partial \mathrm{P}}{\partial x}\frac{\partial \mathrm{H}}{\partial p} - \frac{\partial \mathrm{H}}{\partial x}\frac{\partial \mathrm{P}}{\partial p} \\ &= \frac{\partial \mathrm{P}}{\partial x}\left(\frac{\partial \mathrm{K}}{\partial \mathrm{X}}\frac{\partial \mathrm{X}}{\partial p} + \frac{\partial \mathrm{K}}{\partial \mathrm{P}}\frac{\partial \mathrm{P}}{\partial p}\right) - \left(\frac{\partial \mathrm{K}}{\partial \mathrm{X}}\frac{\partial \mathrm{X}}{\partial x} + \frac{\partial \mathrm{K}}{\partial \mathrm{P}}\frac{\partial \mathrm{P}}{\partial x}\right)\frac{\partial \mathrm{P}}{\partial p} \\ &= \frac{\partial \mathrm{K}}{\partial \mathrm{X}}\left(\frac{\partial \mathrm{P}}{\partial x}\frac{\partial \mathrm{X}}{\partial p} - \frac{\partial \mathrm{X}}{\partial x}\frac{\partial \mathrm{P}}{\partial p}\right) \\ &= \frac{\partial \mathrm{K}}{\partial \mathrm{X}}[\mathrm{P}, \mathrm{X}] = -\frac{\partial \mathrm{K}}{\partial \mathrm{X}}\,.\end{aligned}$$

 **Exercise 8.5.** *Prove equation* (41) *for general* n.

8.4 Integrability

We now come to a very rich topic in mathematics deriving from Hamilton's canonical formalism: *integrability*. We will only be able to scratch the surface of this fascinating topic. We have seen in the previous section that the same hamiltonian system can be expressed in terms of different sets of canonical variables related via canonical transformations. Therefore the question arises whether given a hamiltonian system described by $\mathrm{H}(x, p)$, with $x, p \in \mathbb{R}^n$, we can find a different set (X, P), say, of canonical variables such that when we express (x, p) in terms of (X, P), the hamiltonian $\mathrm{K}(\mathrm{X}, \mathrm{P}) := \mathrm{H}(x(\mathrm{X}, \mathrm{P}), p(\mathrm{X}, \mathrm{P}))$ is *easy to solve™*.

What does *easy to solve™* mean? Perhaps the easiest hamiltonian system to solve is one with hamiltonian

$$\mathrm{H} = \frac{\|\mathrm{P}\|^2}{2m}\,,$$

whose Hamilton's equations simply say that $\dot{\mathrm{P}} = 0$ and hence that $\mathrm{X}(t) = \mathrm{X}(0) + \frac{t}{m}\mathrm{P}(0)$, for some $\mathrm{X}(0), \mathrm{P}(0) \in \mathbb{R}^n$. So let us take a system of this type as a working definition for *easy to solve™*.

The first observation is that the components P_i of the momentum in the *easy to solve™* system do not just Poisson-commute with the hamiltonian, but actually Poisson-commute among themselves. And this in turns explains why they Poisson-commute with the hamiltonian, since the hamiltonian is a function of the momentum alone. A set of functions which Poisson-commute among themselves are said to be **in involution**. At any rate, *easy to solve™* systems have lots of conserved quantities, and this explains why conserved quantities (and hence, via Noether's Theorem, continuous symmetries) play such a crucial rôle in mechanics.

So if we are given a hamiltonian $\mathrm{H}(x, p)$, one question we might like to explore is whether it has enough conserved quantities in involution. These quantities should also be suitably independent: clearly we are not going to get a lot of mileage out of taking the hamiltonian H and its powers H^2, H^3,... for example.

By contrast, the momenta P_i in our *easy to solve™* system are such that their gradients (equivalently $d\mathrm{P}_i$) are pointwise linearly independent.

The wonderful (but hard) **Liouville's theorem** says that if a hamiltonian $\mathrm{H}(x,p)$ on $\mathbb{R}^{2n}$ admits n independent conserved quantities in involution then there is a canonical transformation to so-called **action/angle variables** (X,P) such that Hamilton's equations say that P is constant and hence $\mathrm{X}(t) = \mathrm{X}(0) + t\mathrm{P}(0)$. Such a system is said to be **integrable**.

Let us consider a deceptively simple example: the one-dimensional harmonic oscillator, with hamiltonian

$$\mathrm{H}(x,p) = \frac{p^2}{2m} + \frac{kx^2}{2} \,.$$

Let $\omega^2 = k/m$, so that

$$\mathrm{H} = \frac{1}{2m}\left(p^2 + m^2\omega^2 x^2\right) \,.$$

Since H is a sum of squares, we seek a canonical transformation of the form

$$\begin{aligned} p(\mathrm{X},\mathrm{P}) &= f(\mathrm{P})\cos\mathrm{X} \\ x(\mathrm{X},\mathrm{P}) &= \frac{f(\mathrm{P})}{m\omega}\sin\mathrm{X} \,, \end{aligned}$$

for some function $f(\mathrm{P})$, since then the induced hamiltonian in the new variables is simply

$$\mathrm{K}(\mathrm{X},\mathrm{P}) = \mathrm{H}(x(\mathrm{X},\mathrm{P}),p(\mathrm{X},\mathrm{P})) = \frac{1}{2m}f(\mathrm{P})^2 \,.$$

To determine $f(\mathrm{P})$ we demand that the transformation be canonical, which translates into

$$\frac{\partial x}{\partial \mathrm{X}}\frac{\partial p}{\partial \mathrm{P}} - \frac{\partial p}{\partial \mathrm{X}}\frac{\partial x}{\partial \mathrm{P}} = 1 \,,$$

or in the language of **Geom**, $dp \wedge dx = d\mathrm{P} \wedge d\mathrm{X}$. Either way, we arrive at the equation

$$\frac{f(\mathrm{P})f'(\mathrm{P})}{m\omega} = 1 \implies (f(\mathrm{P})^2)' = 2m\omega \,,$$

which has as a possible solution (we don't need to find the most general solution)

$$f(\mathrm{P}) = \sqrt{2m\omega\mathrm{P}} \,.$$

Therefore the hamiltonian becomes

$$\mathrm{H} = \omega\mathrm{P} \,,$$

whence the conserved P is related to the conserved energy by $\mathrm{P} = \mathrm{E}/\omega$. Finally,

$$\dot{\mathrm{X}} = \frac{\partial \mathrm{K}}{\partial \mathrm{P}} = \omega \implies \mathrm{X}(t) = \omega t + \alpha \,,$$

whence

$$x(t) = \sqrt{\frac{2\mathrm{E}}{m\omega^2}}\sin(\omega t + \alpha)$$
$$p(t) = \sqrt{2m\mathrm{E}}\cos(\omega t + \alpha)\ .$$

Integrable systems are rare by any measure and even if one can prove (via Liouville's theorem) that a system is integrable, it is a whole other matter to find the canonical transformation which renders the hamiltonian *easy to solve™*.

The study of integrable systems branches into many areas of mathematics: mathematical physics, PDEs, algebraic and differential geometry, representation theory,... and it is well represented in the School!

9 Isoperimetric problems

What strange arts necessity finds out.
— Christopher Marlowe (*Dido*)

In this section we begin to consider the kind of problems where we are asked to extremise a functional subject to a functional constraint. The prototypical example is the original *isoperimetric problem*, which has in fact lent its name to all such problems.

The original isoperimetric problem was known to the ancient Greeks[3]: *to find the closed plane curve of a given length that encloses the largest area*. They even managed to convince themselves that the intuitive answer (the circle) was correct. The reason this problem is called isoperimetric is that one is maximising the area inside the curve while keeping the perimeter fixed. More generally, an isoperimetric problem is one where one is trying to extremise a functional subject to a (functional) constraint. In this section we will learn how to deal with such constrained extremisation in the context of the variational calculus. Let us start by setting up the classical isoperimetric problem in this context.

☞ Let $\mathrm{C} \subset \mathbb{R}^2$ be a positively-oriented, (piecewise) smooth, simple closed curve on the plane and let D be the region it bounds. Let $\mathrm{F}, \mathrm{G} : \mathbb{R}^2 \to \mathbb{R}$ be differentiable functions. Then **Green's Theorem** says that

$$\int_{\mathrm{C}} (\mathrm{F}(x_1, x_2)dx_1 + \mathrm{G}(x_1, x_2)dx_2) = \iint_{\mathrm{D}} \left(\frac{\partial \mathrm{G}}{\partial x_1} - \frac{\partial \mathrm{F}}{\partial x_2} \right) d^2x \, .$$

Parametrising C as $(x_1(t), x_2(t))$ for $t \in [0, 1]$, we can rephrase the theorem as

$$\int_0^1 (\mathrm{F}(x_1(t), x_2(t))\dot{x}_1 + \mathrm{G}(x_1(t), x_2(t))\dot{x}_2)\, dt = \iint_{\mathrm{D}} \left(\frac{\partial \mathrm{G}}{\partial x_1} - \frac{\partial \mathrm{F}}{\partial x_2} \right) d^2x \, .$$

Let $x : [0, 1] \to \mathbb{R}^2$ be a C^1 curve which is closed: $x(0) = x(1)$. It follows from Green's Theorem that the area enclosed by the curve can be calculated from the following functional:

$$\mathrm{A}[x] = \tfrac{1}{2} \int_0^1 (x_1\dot{x}_2 - x_2\dot{x}_1)dt \, ,$$

whereas the perimeter of the curve is given by the following functional:

$$\mathrm{S}[x] = \int_0^1 \sqrt{\dot{x}_1^2 + \dot{x}_2^2}dt \, .$$

The isoperimetric problem is the following: extremise $\mathrm{A}[x]$ subject to $\mathrm{S}[x] = \ell$.

[3] This problem is also called *Dido's problem*, in memory of the legendary founder and first queen of Carthage. Legend has it that when she arrived to present-day Tunisia, she tricked the Berber king into giving her some land, by asking only for a plot of land which could be encompassed by an oxhide. When the king granted her request, she cut the oxhide into thin strips and managed to enclose a nearby hill.

9.1 Constrained optimization in several variable calculus

Surely you recognise the finite-dimensional analogue to this problem. Let $f, g : \mathrm{U} \subset \mathbb{R}^n \to \mathbb{R}$ be functions of n variables. One can then extremise f subject to $g = 0$, which means to find $x \in \mathrm{U}$ for which $f(x)$ is an extremum among all nearby y such that $g(y) = 0$. As in **SVCDE**, one can use the method of Lagrange multipliers. This method is predicated on the following result.

Lemma 9.1. *Let $x \in \mathbb{R}^n$ be an extremum of f subject to $g = 0$. Then if $\nabla g(x) \neq 0$, there exists a **Lagrange multiplier** $\lambda \in \mathbb{R}$ such that (x, λ) is a critical point of the function $\mathrm{F} : \mathrm{U} \times \mathbb{R} \to \mathbb{R}$ defined by $\mathrm{F}(y, \lambda) = f(y) - \lambda g(y)$.*

The method of Lagrange multipliers consists in finding the critical points of the function $\mathrm{F} = f - \lambda g$; that is, to solve the equations $\nabla \mathrm{F}(x, \lambda) = 0$. The resulting equations are

$$\frac{\partial \mathrm{F}}{\partial \lambda} = 0 \implies g(x) = 0 \qquad \text{and} \qquad \frac{\partial \mathrm{F}}{\partial x^i} = 0 \implies \frac{\partial f}{\partial x^i} = \lambda \frac{\partial g}{\partial x^i} \,.$$

9.2 The method of Lagrange multipliers

The method of Lagrange multipliers extends to the calculus of variations. We are going to change the notation slightly from previous lectures in that the independent variable will be called x and the dependent variable y. The reason is that isoperimetric problems are generally geometric as opposed to dynamical and the independent variable is not interpretable as time.

Suppose that we wish to extremise the functional

$$\mathrm{J}[y] = \int_0^1 \mathrm{L}(y, y', x) dx$$

on C^1 functions $y : [0, 1] \to \mathbb{R}$, subject to the constraint

$$\mathrm{I}[y] = \int_0^1 \mathrm{K}(y, y', x) dx = 0 \,,$$

where K, L are smooth functions.

Without loss of generality we have taken the constraint to be $\mathrm{I}[y] = 0$ as opposed to $\mathrm{I}[y] = \ell$ for some constant ℓ. Clearly if $\mathrm{I}[y] = \ell$, $\tilde{\mathrm{I}}[y] = \mathrm{I}[y] - \ell = 0$.

Suppose that $y(x)$ is such an extremal. This means that

$$\frac{d}{ds} \mathrm{J}[y + s\varepsilon] \Big|_{s=0} = 0$$

for all variations such that the isoperimetric condition $\mathrm{I}[y + s\varepsilon] = 0$ is satisfied for all s in some neighbourhood of 0. The problem with this formulation is that not every variation ε will satisfy the isoperimetric condition — there may in fact

be no such nonzero variations! — and hence we would not be able to use the Fundamental Lemma to arrive at some differential equation.

One way out is to consider a two-parameter family of nearby curves:

$$\hat{y}(x) = y(x) + s\varepsilon(x) + r\eta(x) \,,$$

where ε is an arbitrary variation and and the "correction term" $r\eta$ is chosen so that for a fixed but arbitrary ε and s in a neighbourhood of 0, the isoperimetric condition $\mathrm{I}[\hat{y}] = 0$ is satisfied. For fixed ε and η, $\mathrm{I}[\hat{y}]$ defines a function $g(r,s)$ of two variables, which is differentiable because we assume that K is. Since y obeys the isoperimetric condition, it follows that $g(0,0) = 0$. If we *assume* that $\nabla g(0,0) \neq 0$, then the Implicit Function Theorem guarantees the existence of a function[4] $r = r(s)$ such that for s in a neighbourhood of 0, $g(r(s),s) = 0$. Let $f(r,s) = \mathrm{J}[\hat{y}]$. Then the fact that $y(x)$ is an extremum of J subject to $\mathrm{I}[y] = 0$, says that $(0,0)$ is an extremum of f subject to $g = 0$. Since $\nabla g(0,0) \neq 0$, it follows from Lemma 9.1 above that there is some $\lambda \in \mathbb{R}$ such that $(0,0,\lambda)$ is a critical point of the function $\mathrm{F}(r,s,\lambda) = f(r,s) - \lambda g(r,s)$. In other words, there is $\lambda \in \mathbb{R}$ satisfying $\nabla \mathrm{F}(0,0,\lambda) = 0$.

These equations are

$$\left.\frac{\partial \mathrm{F}}{\partial s}\right|_{(0,0,\lambda)} = 0 \,, \qquad \left.\frac{\partial \mathrm{F}}{\partial r}\right|_{(0,0,\lambda)} = 0 \,, \qquad \text{and} \qquad \left.\frac{\partial \mathrm{F}}{\partial \lambda}\right|_{(0,0,\lambda)} = 0 \,.$$

The last equation simply says that $g(0,0) = 0$, which is just $\mathrm{I}[y] = 0$. The first two equations can be rewritten as follows:

$$\left.\frac{\partial}{\partial s}(\mathrm{J}[\hat{y}] - \lambda \mathrm{I}[\hat{y}])\right|_{s=r=0} = 0 \qquad \text{and} \qquad \left.\frac{\partial}{\partial r}(\mathrm{J}[\hat{y}] - \lambda \mathrm{I}[\hat{y}])\right|_{s=r=0} = 0 \,. \tag{42}$$

Let us see what these equations mean. In the first equation,

$$\begin{aligned}
\left.\frac{\partial}{\partial s}\mathrm{J}[\hat{y}]\right|_{s=r=0} &= \int_0^1 \left.\frac{\partial}{\partial s}\mathrm{L}(y,y',x)\right|_{s=r=0} dx \\
&= \int_0^1 \left(\frac{\partial \mathrm{L}}{\partial y}\varepsilon + \frac{\partial \mathrm{L}}{\partial y'}\varepsilon'\right) dx \\
&= \int_0^1 \varepsilon\left(\frac{\partial \mathrm{L}}{\partial y} - \frac{d}{dx}\frac{\partial \mathrm{L}}{\partial y'}\right) dx \,,
\end{aligned}$$

where we have integrated by parts and dropped the boundary terms. Similarly,

$$\left.\frac{\partial}{\partial s}\mathrm{I}[\hat{y}]\right|_{s=r=0} = \int_0^1 \varepsilon\left(\frac{\partial \mathrm{K}}{\partial y} - \frac{d}{dx}\frac{\partial \mathrm{K}}{\partial y'}\right) dx \,,$$

so that putting it all together

$$\left.\frac{\partial}{\partial s}(\mathrm{J}[\hat{y}] - \lambda \mathrm{I}[\hat{y}])\right|_{s=r=0} = \int_0^1 \varepsilon\left(\frac{\partial \mathrm{L}}{\partial y} - \frac{d}{dx}\frac{\partial \mathrm{L}}{\partial y'} - \lambda\left(\frac{\partial \mathrm{K}}{\partial y} - \frac{d}{dx}\frac{\partial \mathrm{K}}{\partial y'}\right)\right) dx \,.$$

[4] It may happen that we have $s = s(r)$ instead.

Since ε is arbitrary, the Fundamental Lemma says that

$$\frac{\partial \mathrm{L}}{\partial y} - \frac{d}{dx}\frac{\partial \mathrm{L}}{\partial y'} = \lambda\left(\frac{\partial \mathrm{K}}{\partial y} - \frac{d}{dx}\frac{\partial \mathrm{K}}{\partial y'}\right) ,$$

which is the Euler–Lagrange equation for $\mathrm{M} := \mathrm{L} - \lambda \mathrm{K}$:

$$\frac{\partial \mathrm{M}}{\partial y} = \frac{d}{dx}\frac{\partial \mathrm{M}}{\partial y'} .$$

The second equation in (42) does not give any extra information, since η is taken from a smaller class of variations than ε.

What about the assumption $\nabla g(0,0) \neq 0$? This can be understood variationally. In fact, if $\nabla g(0,0) = 0$, then it follows from the calculation we have just made that y satisfies the Euler–Lagrange equation for K. Hence $\nabla g(0,0) \neq 0$ is the same as requiring y not to be an extremal of $\mathrm{I}[y]$ (with $\mathrm{I}[y] = 0$). In summary, we have proved the following:

Theorem 9.2. *Let* $\mathrm{J}[y] = \int_0^1 \mathrm{L}(y, y', x)dx$ *and* $\mathrm{I}[y] = \int_0^1 \mathrm{K}(y, y', x)dx$ *be functionals for* C^1 *functions* $y : [0,1] \to \mathbb{R}$ *subject to the boundary conditions* $y(0) = y_0$ *and* $y(1) = y_1$. *Suppose that* $y(x)$ *is an extremal of* J *subject to the isoperimetric constraint* $\mathrm{I}[y] = 0$. *Then if* y *is* not *an extremal of* $\mathrm{I}[y]$, *there is a constant* $\lambda \in \mathbb{R}$ *so that* y *is an extremal of* $\mathrm{J} - \lambda \mathrm{I}$.

This theorem justifies the **method of Lagrange multipliers**. Suppose that we wish to extremise a functional

$$\mathrm{J}[y] = \int_0^1 \mathrm{L}(y, y', x)dx$$

for $y : [0,1] \to \mathbb{R}^n$ subject to the boundary conditions $y(0) = y_0$ and $y(1) = y_1$ and the isoperimetric condition $\mathrm{I}[y] = 0$, where

$$\mathrm{I}[y] = \int_0^1 \mathrm{K}(y, y', x)dx .$$

Then the method consists of the following steps:

1. We ensure that $\mathrm{I}[y]$ has no extremals satisfying $\mathrm{I}[y] = 0$
2. We solve the Euler–Lagrange equation for $\mathrm{M}(y, y', x, \lambda) = \mathrm{L}(y, y', x) - \lambda \mathrm{K}(y, y', x)$, which is a second-order ODE for $y(x)$
3. We fix the constants of integration from the boundary conditions
4. We fix the value of λ from $\mathrm{I}[y] = 0$.

☞ If y is an extremal of I, so that $\nabla g(0,0) = 0$, the problem is said to be **abnormal**. In that case, the theorem is modified: there now exist $\lambda, \mu \in \mathbb{R}$, not both zero, such that y is an extremal of $\mu \mathrm{J} - \lambda \mathrm{I}$.

Finally, let us point out that if there are more than one constraint, one must introduce an equal number of Lagrange multipliers.

9.3 Classical isoperimetric problems

Let us now do an example. This is a variant of the original isoperimetric problem. Consider a C^1 function $y : [0,1] \to \mathbb{R}$ with the property that $y(0) = y(1) = 0$ and $y(x) > 0$ everywhere else. Its graph $(x, y(x))$ is a curve from the origin to the point $(1,0)$ and lying in the upper half-plane. We would like to maximise the area under the curve provided that the arclength is fixed to some number $\ell > 1$. In other words, we want to maximise the area functional

$$A[y] = \int_0^1 y(x)\, dx$$

subject to the constraint

$$S[y] = \int_0^1 \sqrt{1 + y'(x)^2}\, dx = \ell\ .$$

According to the method of Lagrange multipliers, we must extremise the modified functional

$$\tilde{A}[y, \lambda] = \int_0^1 \left(y(x) - \lambda \sqrt{1 + y'(x)^2} \right) dx\ .$$

The Euler–Lagrange equation resulting from this functional is

$$1 + \frac{d}{dx}\left(\lambda \frac{y'(x)}{\sqrt{1 + y'(x)^2}} \right) = 0\ .$$

Integrating once we find

$$x + \lambda \frac{y'(x)}{\sqrt{1 + y'(x)^2}} = c_1\ ,$$

for some constant c_1. From this equation we can solve for $y'(x)$ and then integrate again to solve for $y(x)$.

Exercise 9.1. *Complete the above analysis and prove that the graph $(x, y(x))$ traces a circle of radius λ passing through $(0,0)$ and $(1,0)$ and with centre at the point $(\frac{1}{2}, -\sqrt{\lambda^2 - \frac{1}{4}})$. Finally show that λ is determined from the arclength ℓ by the transcendental equation $2\lambda \sin(\ell/2\lambda) = 1$.*

Exercise 9.2. *Solve the original isoperimetric problem stated at the start of this section. Deduce that for any closed plane curve the area A of the enclosed region and the perimeter ℓ of the curve satisfy the following* ***isoperimetric inequality*** *$A \leq \ell^2/4\pi$, with equality if and only if the curve is a circle.*
(Hint: Extremise the modified action

$$\tilde{A}[x] = \int_0^1 \left(\tfrac{1}{2}(x_1 \dot{x}_2 - x_2 \dot{x}_1) - \lambda \sqrt{(\dot{x}_1)^2 + (\dot{x}_2)^2} \right) dt\ ,$$

and deduce that the resulting curve is a circle of radius λ.)

☞ A similar "isoperimetric" inequality exists between the surface area S of a closed surface in $\mathbb{R}^3$ and the volume V it encloses. In this case, one can prove the bound $V^2 \leq S^3/36\pi$, which is now saturated by the sphere.

9.4 The catenary

We do a second classical example: the catenary. Suppose that a uniform chain (or rope) of length ℓ hangs under its own weight from two poles of height h, a distance $2\ell_0 < \ell$ apart. What is the curve that the chain will describe?

Physics (i.e., Torricelli's Principle, although presumably already known to Galileo) tells us that the chain will try to lower its centre of gravity or, equivalently, minimise its gravitational potential energy. If we let $H(s)$ denote the height as a function of the arclength s, then the potential energy is given (in some units) by $\int_0^\ell H(s)ds$. Parametrising the chain by $x \in [-\ell_0, \ell_0]$ with $y(x) = H(s(x))$ the height at x, we can rewrite the potential energy as

$$\int_{-\ell_0}^{\ell_0} y(x)\sqrt{1+y'(x)^2}dx\,,$$

subject to the boundary conditions $y(\pm\ell_0) = h$ and the isoperimetric constraint

$$\int_{-\ell_0}^{\ell_0} \sqrt{1+y'(x)^2}dx = \ell\,.$$

Notice that all extremals of the arclength functional are straight lines and hence have length $2\ell_0$, so there are no extremals with length ℓ, since by assumption $\ell > 2\ell_0$. So we can use the method of Lagrange multipliers.

We look for extremals of the functional defined by the lagrangian

$$L(y, y', \lambda) = (y-\lambda)\sqrt{1+y'^2}\,.$$

Since the lagrangian does not depend explicitly on x, Beltrami's identity gives us a first integral:

$$\frac{\partial L}{\partial y'}y' - L = \text{constant} \implies \frac{y-\lambda}{\sqrt{1+y'^2}} = c\,,$$

which we can turn into a differential equation for y:

$$y'^2 = ((y-\lambda)/c)^2 - 1\,.$$

✍ **Exercise 9.3.** *Solve the above differential equation subject to the boundary conditions: $y(\pm\ell_0) = h$ to arrive at*

$$y(x) = c\cosh(x/c) - c\cosh(\ell_0/c) + h\,.$$

Finally we impose the isoperimetric condition:

$$\begin{aligned}\int_{-\ell_0}^{\ell_0} \sqrt{1+y'(x)^2}dx &= \int_{-\ell_0}^{\ell_0} \sqrt{1+\sinh(x/c)^2}dx \\ &= \int_{-\ell_0}^{\ell_0} \cosh(x/c)dx \\ \therefore \qquad \ell &= 2c\sinh(\ell_0/c)\end{aligned}$$

Introducing $\zeta = 1/(2c)$, we find the following transcendental equation:

$$\ell\zeta = \sinh(2\ell_0\zeta)\ .$$

This equation has a trivial solution $\zeta = 0$. For $\zeta > 0$ small, the condition $\ell > 2\ell_0$ ensures that $\ell\zeta > \sinh(2\ell_0\zeta)$, but clearly for $\zeta \gg 1$ the exponential term dominates and $\ell\zeta < \sinh(2\ell_0\zeta)$. Therefore by continuity, there is some $\zeta_0 > 0$ where $\ell\zeta_0 = \sinh(2\ell_0\zeta_0)$.

I claim that there is precisely one such ζ. To see this, consider the function $f(\zeta) := \sinh(2\ell_0\zeta) - \ell\zeta$. As we observed above, $f(0) = 0$ and for $\zeta > 0$ small, $f(\zeta) < 0$. Differentiating $f'(\zeta) = 2\ell_0\cosh(2\ell_0\zeta) - \ell$ and $f''(\zeta) = 4\ell_0^2\sinh(2\ell_0\zeta)$. Since $f(\zeta_0) = 0$, f has a minimum somewhere in $(0, \zeta_0)$. If there were a second positive zero, say at ζ_1, then there would be a maximum somewhere in (ζ_0, ζ_1), but in fact for all positive ζ, $f''(\zeta) > 0$, so no such maximum can exist.

Exercise 9.4. *Consider the same problem but where the height of the second pole is not fixed. In other words, minimise the gravitational potential energy subject to the same isoperimetric constraint, but now subject to the boundary conditions $y(-\ell_0) = h$ and $y(\ell_0)$ free.*

10 Holonomic and nonholonomic constraints

The method of Lagrange multipliers in the calculus of variations extends to other types of constrained extremisation, where the subsidiary condition is not a functional but actually a function; that is, rather than a constraint of the form $\int_0^1 \mathrm{K}(x, \dot{x}, t)dt = 0$, we have one of the form $\mathrm{G}(x, \dot{x}, t) = 0$.

Such constraints come in several flavours, with rather baroque names. A constraint which does not depend explicitly on t is called **scleronomic**, whereas one which does is called **rheonomic**. If the constraint does not depend explicitly on the velocities, it is said to be **holonomic**, otherwise it is **nonholonomic**.

This latter nomenclature is not standard: some people restrict *nonholonomic* to constraints which cannot be integrated.

The types of problems where these constraints appear naturally include the following:

- Finding geodesics on a surface defined as the zero locus of a function, say, $\mathrm{G}(x) = 0$, which is an example of a scleronomic, holonomic constraint;
- Reducing higher-order lagrangians to first-order lagrangians. For example, given a lagrangian $\mathrm{L}(x, \dot{x}, \ddot{x}, t)$ depending on the second derivative of the function x, it can be replaced by a first-order lagrangian $\mathrm{L}(x, \dot{x}, \dot{y}, t)$ with the subsidiary condition $y = \dot{x}$, which is a scleronomic, nonholonomic constraint;
- Mechanical problems of the type "rolling without slipping", which are typically nonholonomic.

10.1 Holonomic constraints

Let $x : [0, 1] \to \mathbb{R}^2$ with $x(0) = x_0$ and $x(1) = x_1$ be an extremal for a functional

$$\mathrm{J}[x] = \int_0^1 \mathrm{L}(x, \dot{x}, t)dt$$

with L smooth and subject to $g(x(t), t) = 0$ for all $t \in [0, 1]$, where $g : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$ is a smooth function. In particular, it follows that $g(x_0, 0) = 0$ and $g(x_1, 1) = 0$. We will also assume that $\nabla g(x(t), t) \neq 0$ for all $t \in [0, 1]$. This last condition allows us to use the Implicit Function Theorem to guarantee that we can solve (at least in principle) $g(x, t) = 0$ for one of the components of x. This approach, even if were practical to solve $g(x, t) = 0$, is not optimal as it singles out one of the components of x, thus potentially destroying an underlying symmetry. Luckily, the method of Lagrange multipliers extends to this situation.

We seek a necessary condition on x for $\mathrm{J}[x]$ to be an extremum. We say that $\hat{x} := x + s\varepsilon$ is an **admissible variation** if $\hat{x}$ is C^2, $\hat{x}(0) = x_0$, $\hat{x}(1) = x_1$ and

$g(\hat{x}(t), t) = 0$ for all $t \in [0,1]$.[5] In particular, it follows that ε is C^2 and that $\varepsilon(0) = \varepsilon(1) = 0$ and that at least for s near 0,

$$g(x + s\varepsilon, t) = 0 \,.$$

For simplicity, let us assume that the condition $\nabla g(x(t), t) \neq 0$ is satisfied due to $\frac{\partial g}{\partial x_2}\Big|_{(x(t),t)} \neq 0$ for all $t \in [0,1]$. Then the Implicit Function Theorem says that $g(x + s\varepsilon, t) = 0$ can be solved for ε_2 in terms of ε_1 and s for s near 0. This allows to regard ε_1 as an arbitrary function and ε_2 as the unique (again by the Implicit Function Theorem) solution of $g(x + s\varepsilon, t) = 0$. Notice that since $(x_0, 0)$ solves $g(x_0 + s\varepsilon, 0) = 0$ and since $\varepsilon_1(0) = 0$, it also follows that $\varepsilon_2(0) = 0$ and similarly $\varepsilon_2(1) = 0$.

Now suppose that $\hat{x}$ is admissible. Since x is an extremal of J,

$$\int_0^1 \left(\left(\frac{\partial \mathrm{L}}{\partial x_1} - \frac{d}{dt} \frac{\partial \mathrm{L}}{\partial \dot{x}_1} \right) \varepsilon_1 + \left(\frac{\partial \mathrm{L}}{\partial x_2} - \frac{d}{dt} \frac{\partial \mathrm{L}}{\partial \dot{x}_2} \right) \varepsilon_2 \right) dt = 0 \,. \tag{43}$$

We cannot deduce the Euler–Lagrange equation because ε_2 is not arbitrary: it is linked to ε_1 by $g(\hat{x}, t) = 0$. Differentiating this condition with respect to s at $s = 0$, we obtain

$$0 = \frac{d}{ds} g(\hat{x}, t)\Big|_{s=0} = \frac{\partial g}{\partial x_1} \varepsilon_1 + \frac{\partial g}{\partial x_2} \varepsilon_2 \,.$$

(There is also in principle a term in $\frac{d\varepsilon_2}{ds}$, but this vanishes at $s = 0$.) Since by assumption $\frac{\partial g}{\partial x_2} \neq 0$, we can solve for ε_2:

$$\varepsilon_2 = -\frac{\frac{\partial g}{\partial x_1}}{\frac{\partial g}{\partial x_2}} \varepsilon_1 \,. \tag{44}$$

Since L is smooth, for any smooth x,

$$\frac{d}{dt} \frac{\partial \mathrm{L}}{\partial \dot{x}_2} - \frac{\partial \mathrm{L}}{\partial x_2}$$

is a continuous function of t. Since $\frac{\partial g}{\partial x_2}$ is also continuous and moreover nonzero, there is some continuous function $\lambda(t)$ so that

$$\frac{d}{dt} \frac{\partial \mathrm{L}}{\partial \dot{x}_2} - \frac{\partial \mathrm{L}}{\partial x_2} = \lambda(t) \frac{\partial g}{\partial x_2} \,. \tag{45}$$

[5] As in the discussion of general variations in Section 3.2, the expression for $\hat{x}$ represents an abuse of notation. It would be more correct to consider a C^2 family x_s of curves, with $\varepsilon := \frac{\partial x_s}{\partial s}\Big|_{s=0}$, et cetera.

(Of course, the function λ depends on x.) Therefore, inserting this into equation (43),

$$\begin{aligned} 0 &= \int_0^1 \left(\left(\frac{\partial \mathrm{L}}{\partial x_1} - \frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{x}_1} \right) \varepsilon_1 + \lambda(t) \frac{\partial g}{\partial x_2} \varepsilon_2 \right) dt \\ &= \int_0^1 \left(\left(\frac{\partial \mathrm{L}}{\partial x_1} - \frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{x}_1} \right) \varepsilon_1 - \lambda(t) \frac{\partial g}{\partial x_1} \varepsilon_1 \right) dt \\ &= \int_0^1 \left(\frac{\partial \mathrm{L}}{\partial x_1} - \frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{x}_1} - \lambda(t) \frac{\partial g}{\partial x_1} \right) \varepsilon_1 dt \,, \end{aligned}$$

where we have used equation (44). Since ε_1 is arbitrary, the Fundamental Lemma says that

$$\frac{\partial \mathrm{L}}{\partial x_1} - \frac{d}{dt}\frac{\partial \mathrm{L}}{\partial \dot{x}_1} = \lambda(t) \frac{\partial g}{\partial x_1} \,,$$

which together with equation (45), says that x satisfies the Euler–Lagrange equations for the lagrangian $\mathrm{L}(x, \dot{x}, t) - \lambda(t) g(x, t)$.

In summary we have proved the following:

Theorem 10.1. *Let $x : [0,1] \to \mathbb{R}^2$ be a smooth extremal for $\mathrm{J}[x] = \int_0^1 \mathrm{L}(x, \dot{x}, t) dt$ subject to $g(x,t) = 0$ and with $\nabla g(x(t), t) \neq 0$ for all $t \in [0,1]$. Then there exists $\lambda : [0,1] \to \mathbb{R}$ such that x obeys the Euler–Lagrange equations of $\mathrm{L} - \lambda g$.*

This theorem extends to $x : [0,1] \to \mathbb{R}^n$ with a similar proof. Another easy extension is to the case of more than one constraint: one simply introduces as many Lagrange multipliers as constraints.

Therefore we can extend the method of Lagrange multipliers to the case of holonomic constraints. To find extrema of $\mathrm{J}[x] = \int_0^1 \mathrm{L}(x, \dot{x}, t) dt$ for $x : [0,1] \to \mathbb{R}^n$ subject to $g(x,t) = 0$ and boundary conditions $x(0) = x_0$ and $x(1) = x_1$, we solve the Euler–Lagrange equations for the lagrangian $\mathrm{K}(x, \dot{x}, t, \lambda) = \mathrm{L}(x, \dot{x}, t) - \lambda g(x,t)$.

A typical problem is that of geodesics on a regular surface Σ in $\mathbb{R}^3$ defined implicitly by $g(x,y,z) = 0$, where $\nabla g(x,y,z) \neq 0$ on for $(x,y,z) \in \Sigma$. The modified lagrangian is now

$$\mathrm{L}(x, y, z, \dot{x}, \dot{y}, \dot{z}, \lambda) = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} - \lambda g(x,y,z) \,,$$

whose Euler–Lagrange equations are

$$\begin{aligned} \frac{d}{dt} \frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} &= -\lambda \frac{\partial g}{\partial x} \\ \frac{d}{dt} \frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} &= -\lambda \frac{\partial g}{\partial y} \\ \frac{d}{dt} \frac{\dot{z}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} &= -\lambda \frac{\partial g}{\partial z} \,, \end{aligned}$$

in addition to the constraint $g(x,y,z) = 0$ which is the Euler–Lagrange equation for λ.

Exercise 10.1 (Geodesics on the sphere revisited). *Using the method of Lagrange multipliers prove that the geodesics on a sphere are given by great circles.*
(Hint: Extremise $S[x] = \int_0^1 \|\dot{x}\|\, dt$ for $x : [0,1] \to \mathbb{R}^3$ subject to $\|x\|^2 = 1$.)

10.2 Nonholonomic constraints

Although we will not give a proof, the method of Lagrange multipliers extends as well to the case of nonholonomic constraints $g(x,\dot{x},t) = 0$ *which are at most linear in the velocities*, the so-called **pfaffian constraints**.

To illustrate the method, let us derive the Euler–Lagrange equation for a lagrangian $L(x,\dot{x},\ddot{x},t)$ depending on the second derivative of a C^2 function $x : [0,1] \to \mathbb{R}$. The method of Lagrange multipliers says this problem is the same as extremising the action with lagrangian $L(x,\dot{x},\dot{y},t)$ subject to the constraint $y = \dot{x}$. As suggested above, we construct the modified action

$$J[x,y,\lambda] = \int_0^1 \left(L(x,\dot{x},\dot{y},t) - \lambda(t)(y-\dot{x})\right)\, dt\,.$$

The Euler–Lagrange equations are

$$\frac{\partial L}{\partial x} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}} + \lambda\right)\,, \qquad -\lambda = \frac{d}{dt}\frac{\partial L}{\partial \dot{y}} \quad \text{and} \quad y = \dot{x}\,.$$

Solving for λ from the second equation and inserting the result in the second, we obtain

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} + \frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{x}} = 0\,,$$

where we have used that $y = \dot{x}$. This result should be compared with your result for the first part in Exercise 3.1.

We can now also explain the Hint to Exercise 3.1 concerning the fact that the variations and their first derivatives should both vanish at the endpoints. In the above formulation there are variations for both x and y. Let's call them ε and η respectively. We have that $\varepsilon(0) = \varepsilon(1) = \eta(0) = \eta(1) = 0$. If the variations are to preserve the constraint $y = \dot{x}$, then we must have that $\eta = \dot{\varepsilon}$; whence $\dot{\varepsilon}$ should vanish at the endpoints.

10.3 Some examples of holonomic and nonholonomic constraints

It may be instructive to see a couple of examples of holonomic and non-holonomic constraints to get a feeling for the difference. Holonomic constraints restrict the "position", e.g., if we impose $g(x) = 0$, then we are declaring that $x(t)$ lies on the surface defined by $g(x) = 0$ for all t. Nonholonomic constraints typically do not restrict the position, but the "velocity". I say "typically", because there are times when nonholonomic constraints can be integrated to give rise to constraints which do not involve the velocities. (Some people reserve the name *nonholonomic* for those constraints which cannot be so integrated.)

10.3.1 A unicycle

A unicycle is described by three coordinates: (x, y, θ), where (x, y) is the point of contact of the wheel with the ground and θ is the angle of orientation of the wheel relative to the x-axis, say, as illustrated in Figure 12.

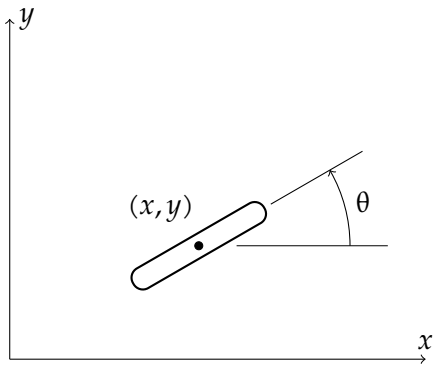


Figure 12: A unicycle

The only constraint is that the velocity $(\dot{x}, \dot{y})$ lie in the direction to where the unicycle is pointing. This can be paraphrased as saying that the unicycle cannot move sideways. In other words,

$$\dot{x} \sin\theta - \dot{y} \cos\theta = 0 \,,$$

since that is the inner product of the velocity vector on the ground $(\dot{x}, \dot{y})$ with the unit vector perpendicular to the unicycle $(\sin\theta, -\cos\theta)$. This is a non-holonomic constraint and cannot be integrated to a holonomic constraint. To see this, simply observe that the unicycle can go anywhere and point in any direction, so its "position" is not constrained.

For the benefit of students of **Geom**, another way to write the above constraint is in terms of differential forms. The velocity vector must be annihilated by the 1-form

$$\alpha = \sin\theta dx - \cos\theta dy \,.$$

The reason this constraint is nonholonomic is that α is not exact. Indeed, it is not even closed:

$$d\alpha = \cos\theta d\theta \wedge dx + \sin\theta d\theta \wedge dy \neq 0 \,.$$

10.3.2 A cylindrical hoop rolling off of a static cylinder

This second example is one my favourite applications of the method of Lagrange multipliers.

Consider a cylindrical hoop of mass m and radius r which is rolling on a static cylinder of radius R, as shown in Figure 13. The hoop is rolling without slipping under the effect of gravity. If the hoop sits at rest at the top of the cylinder it will remain there, since that is a point of (unstable) equilibrium. However if we perturb it slightly, it will start rolling down. Imagine we set the hoop in motion

from the top (with a very small velocity just so that it can leave that point of unstable equilibrium). *When will the hoop fall off the cylinder?*

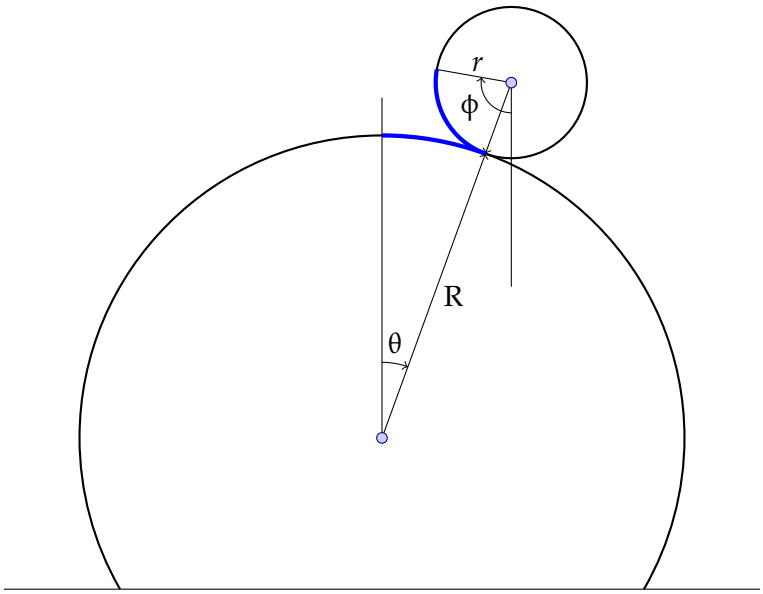


Figure 13: A hoop on a cylinder

Let ρ denote the distance between the centres of the two cylinders. While the hoop is rolling on the cylinder, $\rho = R + r$, which can think of as a holonomic constraint. The other constraint comes from the fact that the hoop rolls without slipping, which can be written as

$$r(\dot{\phi} - \dot{\theta}) = R\dot{\theta} \, .$$

This looks like a nonholonomic constraint, but it is easily integrated to give a second holonomic constraint:

$$r(\phi - \theta) = R\theta \, ,$$

where the integration constant vanishes since the initial condition is such that the hoop is at the top $\theta = 0$ when it starts moving.

The lagrangian for this system consists of several terms. First we have the kinetic energy, which consists of two parts: the motion of the centre of mass of the hoop, which we write using polar coordinates with origin at the centre of the static cylinder:

$$\tfrac{1}{2}m\left(\dot{\rho}^2 + \rho^2\dot{\theta}^2\right)$$

and the rotation of the hoop about its axis of symmetry. We have not discussed rigid bodies in this course, and hence we will have to take it on faith that this rotation contributes the following term to the lagrangian:

$$\tfrac{1}{2}mr^2\dot{\phi}^2 \, .$$

We then have the potential energy term, which is the gravitational potential energy of the hoop, as if all its mass were concentrated at its centre. Hence, $V = mg\rho\cos\theta$. Finally, we subtract the Lagrange multiplier terms

$$\lambda(\rho - r - R) + \mu(r\phi - (R + r)\theta)\,,$$

where we have introduced one multiplier per constraint. In summary, the lagrangian is given by

$$L = \tfrac{1}{2}m\left(\dot\rho^2 + \rho^2\dot\theta^2\right) + \tfrac{1}{2}mr^2\dot\phi^2 - mg\rho\cos\theta - \lambda(\rho - r - R) - \mu(r\phi - (R + r)\theta)\,.$$

We find

$$\begin{aligned}
\frac{\partial L}{\partial\dot\rho} &= m\dot\rho & \frac{\partial L}{\partial\rho} &= m\rho\dot\theta^2 - mg\cos\theta - \lambda\\
\frac{\partial L}{\partial\dot\theta} &= m\rho^2\dot\theta & \frac{\partial L}{\partial\theta} &= mg\rho\sin\theta + \mu(R + r)\\
\frac{\partial L}{\partial\dot\phi} &= mr^2\dot\phi & \frac{\partial L}{\partial\phi} &= -\mu r\,.
\end{aligned}$$

We can now write down five Euler–Lagrange equations for the five dependent variables $(\rho, \theta, \phi, \lambda, \mu)$:

$$\begin{aligned}
m\ddot\rho - m\rho\dot\theta^2 + mg\cos\theta + \lambda &= 0 & mr^2\ddot\phi + \mu r &= 0\\
m\frac{d}{dt}(\rho^2\dot\theta) - mg\rho\sin\theta - \mu(R + r) &= 0 & \rho - r - R &= 0\\
& & r\phi - (R + r)\theta &= 0\,.
\end{aligned}$$

The constraint $\rho = r + R$ implies that $\dot\rho = 0$ and $\ddot\rho = 0$, whence inserting into the Euler–Lagrange equations we arrive at

$$\begin{aligned}
-m(R + r)\dot\theta^2 + mg\cos\theta + \lambda &= 0 & mr^2\ddot\phi + \mu r &= 0\\
m(R + r)\ddot\theta - mg\sin\theta - \mu &= 0 & r\phi - (R + r)\theta &= 0\,.
\end{aligned}$$

Solving for μ from the ϕ-equation, we obtain $\mu = -mr\ddot\phi = -m(R + r)\ddot\theta$, where we have used (the second derivative of) the no-slip constraint. Inserting this into the θ-equation, we arrive at

$$2m(R + r)\ddot\theta - mg\sin\theta = 0\,.$$

Multiplying by $\dot\theta$ we find can integrate this equation to

$$m(R + r)\dot\theta^2 = -mg\cos\theta + C\,,$$

where C is a constant of integration fixed to be mg by the requirement that at $\theta = 0$, $\dot\theta = 0$.

This is a convenient idealisation: we have that at $\theta = 0$ there is a small velocity just to get the hoop rolling, so we would take $\dot\theta = \epsilon$ and then C would have a small correction of order ϵ^2. This will decrease the angle at which the hoop leaves the cylinder by terms of order ϵ^2.

In summary,

$$m(\mathrm{R}+r)\dot{\theta}^2 = mg(1-\cos\theta)\ .$$

Finally, we plug this back in the ρ-equation to arrive at

$$-mg(1-\cos\theta) + mg\cos\theta + \lambda = 0\ ,$$

from which we solve for λ:

$$\lambda = mg(1-2\cos\theta)\ ,$$

which vanishes when $\cos\theta = \frac{1}{2}$ or $\theta = \frac{\pi}{3}$. The Lagrange multiplier has the interpretation as the force required to impose the constraint, so in this case λ is the force required to keep the hoop on the cylinder. When $\lambda = 0$, that force vanishes and the hoop is free to fly off. This happens when $\theta = \frac{\pi}{3}$.

One can generalise this problem by considering other shapes: say a coin rolling on a cylinder or a sphere, or even a ball rolling on a cylinder or on a larger ball. The calculation is very similar, except that the term that the rotation of the smaller body contributes to the lagrangian will change. For instance, if instead of a hoop we had a solid cylinder, the rotational term in the kinetic energy would be

$$\tfrac{1}{4}mr^2\dot{\phi}^2\ .$$

Exercise 10.2. *At what angle does the solid cylinder fly off the larger cylinder?*

11 Some variational PDEs

> In contrast to ordinary differential equations, there is no unified theory of partial differential equations. Some equations have their own theories, while others have no theory at all. — Vladimir Arnol'd

Thus far we have considered functionals defined on curves; that is, on functions of one variable. The Euler–Lagrange equations obtained in this way are always ordinary differential equations. In the same way, one can obtain partial differential equations by varying functionals of functions of several variables. In fact, many of the interesting partial differential equations arise in this way.

11.1 Extremal surfaces

By way of introduction let us consider the problem of extremising functionals defined on surfaces as opposed to curves. Let $y : \mathrm{D} \subset \mathbb{R}^2 \to \mathbb{R}^n$ be a C^2 function defined on a bounded set D in the plane. Let $\partial\mathrm{D}$ denote the boundary of the set D, which we will assume to be smooth or at least piecewise smooth. We will let u, v denote the coordinates on the plane, so that $y(u, v) \in \mathbb{R}^n$ for every $(u, v) \in \mathrm{D}$. Consider a lagrangian $\mathrm{L} : \mathbb{R}^n \times \mathbb{R}^{2n} \times \mathbb{R}^2 \to \mathbb{R}$, and the corresponding action

$$\mathrm{S}[y] = \int_{\mathrm{D}} \mathrm{L}(y, y_u, y_v, u, v)\, du\, dv\,,$$

where y_u and y_v denote collectively the $2n$ partial derivatives $\partial y^i/\partial u$ and $\partial y^i/\partial v$ for $i = 1, \dots, n$. The boundary conditions are specified by asking that $y(x) = \phi(x)$ for $x \in \partial\mathrm{D}$, where $\phi : \partial\mathrm{D} \to \mathbb{R}^n$ is a given function. This may seem odd, but notice that this is precisely what we did earlier in the one-dimensional case. Indeed, in that case, $\mathrm{D} = [0, 1] \subset \mathbb{R}$ and $\partial\mathrm{D} = \{0\} \cup \{1\}$ consists of two points.

The variations for this problem are now C^1 functions $\varepsilon : \mathrm{D} \to \mathbb{R}^n$ such that ε vanishes on the boundary: $\varepsilon(x) = 0$ for $x \in \partial\mathrm{D}$. The condition that a function $y : \mathrm{D} \to \mathbb{R}^n$ be a critical point of the action $\mathrm{S}[y]$ is then that

$$\frac{d}{ds}\mathrm{S}[y + s\varepsilon]\Big|_{s=0} = 0 \qquad \text{for all variations } \varepsilon \text{ with } \varepsilon|_{\partial\mathrm{D}} = 0.$$

To derive the Euler–Lagrange equation, let us differentiate under the integral

sign to find

$$\begin{aligned}
0 &= \frac{d}{ds}\mathrm{S}[y+s\varepsilon]\Big|_{s=0} \\
&= \int_{\mathrm{D}} \frac{d}{ds}\mathrm{L}(y+s\varepsilon, y_u + s\varepsilon_u, y_v + s\varepsilon_v, u, v)\Big|_{s=0} du\,dv \\
&= \int_{\mathrm{D}} \left(\sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial y^i}\varepsilon^i + \sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial y^i_u}\varepsilon^i_u + \sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial y^i_v}\varepsilon^i_v \right) du\,dv \\
&= \int_{\mathrm{D}} \sum_{i=1}^{n} \left(\frac{\partial \mathrm{L}}{\partial y^i} - \frac{\partial}{\partial u}\frac{\partial \mathrm{L}}{\partial y^i_u} - \frac{\partial}{\partial v}\frac{\partial \mathrm{L}}{\partial y^i_v} \right) \varepsilon^i du\,dv \\
&\quad + \int_{\mathrm{D}} \sum_{i=1}^{n} \left(\frac{\partial}{\partial u}\left(\frac{\partial \mathrm{L}}{\partial y^i_u}\varepsilon^i \right) + \frac{\partial}{\partial v}\left(\frac{\partial \mathrm{L}}{\partial y^i_v}\varepsilon^i \right) \right) du\,dv\,.
\end{aligned}$$

The Divergence Theorem (see below) allows us to rewrite the last integral as an integral over the boundary $\partial\mathrm{D}$, which is then seen to vanish since $\varepsilon|_{\partial\mathrm{D}} = 0$. The generalisation of the Fundamental Lemma to functions of more than one variable (see below) then allows us to deduce that the first term above vanishes for all variations ε if and only if the Euler–Lagrange equations

$$\frac{\partial \mathrm{L}}{\partial y^i} = \frac{\partial}{\partial u}\frac{\partial \mathrm{L}}{\partial y^i_u} + \frac{\partial}{\partial v}\frac{\partial \mathrm{L}}{\partial y^i_v}$$

are satisfied.

11.2 The multidimensional Euler–Lagrange equations

There is no reason why we must restrict ourselves to functions of only two variables, but to state the general result it will prove crucial to introduce some notation. Let $\mathrm{D} \subset \mathbb{R}^m$ be a bounded region with (piecewise) smooth boundary. We will let $x^\mu = (x^1, \ldots, x^m)$ denote coordinates for D, which we will often just denote collectively as x. Let $\mathrm{L}(y, \nabla y, x)$ be a lagrangian for maps $y : \mathrm{D} \to \mathbb{R}^n$, where ∇y denotes collectively the mn partial derivatives $y^i_\mu := \frac{\partial y^i}{\partial x^\mu}$ for $i = 1, \ldots n$ and $\mu = 1, \ldots, m$. Then the very same manipulations as in the case of $m = 2$ above would yield the general multidimensional Euler–Lagrange equations:

$$\boxed{\frac{\partial \mathrm{L}}{\partial y^i} = \sum_{\mu=1}^{m} \frac{\partial}{\partial x^\mu}\frac{\partial \mathrm{L}}{\partial y^i_\mu}}\,. \tag{46}$$

Notice that the partial derivatives $\frac{\partial}{\partial x^\mu}$ in the above expression are not just differentiating any explicit dependence L might have on x but *also* the function y. In other words, we treat $\frac{\partial \mathrm{L}}{\partial y^i_\mu}$ as a function of the x's and differentiate with respect to x^μ keeping all other x's fixed.

It remains to discuss the Divergence Theorem and the generalisation of the Fundamental Lemma. We start with the the Divergence Theorem. It states the following.

Theorem 11.1 (Divergence Theorem). *Let* $D \subset \mathbb{R}^m$ *be a bounded open set with (piecewise) smooth boundary* ∂D. *Let* $X = (X^1, \dots X^m)$ *be a smooth vector field defined on* $D \cup \partial D$. *Let* N *be the unit outward-pointing normal of* ∂D. *Then*

$$\int_D \sum_{\mu=1}^{m} \partial_\mu X^\mu dV = \int_{\partial D} \langle X, N \rangle \, dA \,,$$

where dV *is the volume element in* $\mathbb{R}^m$ *and* dA *is the area element in* ∂D, *and* $\langle -, - \rangle$ *denotes the euclidean inner product in* $\mathbb{R}^m$.

The divergence theorem used to be taught in a third-year course on PDEs. Michael Singer (who used to teach that course) wrote an excellent set of notes which are available from the following URL:

`http://goo.gl/781489`

We will not say more about this theorem here, but we will use it freely.

Finally we turn to the generalisation of the Fundamental Lemma to the case of multidimensional integrals. The lemma now says the following.

Theorem 11.2 (Multidimensional version of the Fundamental Lemma). *Let* $D \subset \mathbb{R}^m$ *be a bounded open set with (piecewise) smooth boundary* ∂D. *Let* $f : D \to \mathbb{R}^n$ *be a continuous function which obeys*

$$\int_D \langle f(x), h(x) \rangle \, d^m x = 0$$

for all C^∞ *functions* $h : D \to \mathbb{R}^n$ *vanishing on the boundary* ∂D. *Then* $f \equiv 0$.

Proof for $n = 1$. As before we will prove the case $n = 1$ and leave the trivial extension to general n as an exercise. Mutatis mutandis, the proof is the same as if $m = 1$, which was done in Section 2.4, so we will be brief.

Assume for a contradiction that there exists a point $x_0 \in D$ where $f(x_0) \neq 0$. Without loss of generality we will assume that $f(x_0) > 0$. Then there is an open ball B centred at x_0 contained in D with the property that, for all $x \in B$, $f(x) > c > 0$ for some constant c. We will now construct a non-negative smooth function $h : D \to \mathbb{R}$ with the usual properties: it vanishes outside B and it has positive integral $\int_D h dV = \int_B h dV > 0$. Then, as in the proof of the one-dimensional Fundamental Lemma, it follows that $\int_D f h dV > 0$, violating the hypothesis.

The construction of the function h is again very similar to what was done before. Inside the open ball B there is a hypercube centred at x_0 with sides of length 2δ for some $\delta > 0$. Explicitly, the hypercube is the cartesian product

$$[x_0^1 - \delta, x_0^1 + \delta] \times [x_0^2 - \delta, x_0^2 + \delta] \times \cdots \times [x_0^m - \delta, x_0^m + \delta] \,,$$

where $x_0 = (x_0^1, x_0^2, \dots, x_0^m)$. Now define $h(x)$ to be the product of the m functions

$$h(x^1, x^2, \dots, x^m) = \varphi_{x_0^1 - \delta, x_0^1 + \delta}(x^1) \varphi_{x_0^2 - \delta, x_0^2 + \delta}(x^2) \cdots \varphi_{x_0^m - \delta, x_0^m + \delta}(x^m) \,,$$

where $\varphi_{a,b}(x)$ are defined by equation (5). We leave it as an exercise to the reader to convince herself that this function does the trick. □

Exercise 11.1. *Prove the multidimensional version of the Fundamental Lemma for $n > 1$.*

11.3 Two examples

11.3.1 The Dirichlet problem for harmonic functions

Let's look at a classic problem from partial differential equations: the **Dirichlet problem** for Laplace's equation in the plane. The problem is to solve Laplace's equation $\triangle\phi = 0$ for some function $\phi : D \to \mathbb{R}$ defined on the unit disk $D \subset \mathbb{R}^2$ subject to $\phi|_{\partial D} = f$, for some function $f : \partial D \to \mathbb{R}$ defined on the unit circle ∂D. Explicitly, the equation is

$$\phi_{xx} + \phi_{yy} = 0 \,.$$

This is the Euler–Lagrange equation associated to the functional

$$E[\phi] = \int_D \tfrac{1}{2}\left(\phi_x^2 + \phi_y^2\right) dx\, dy \,,$$

where $\phi_x = \partial\phi/\partial x$ and $\phi_y = \partial\phi/\partial y$. The functional E is called the **energy** in the partial differential equation literature, and what this shows is that the solutions to the Dirichlet problem for Laplace's equation on the unit disk extremise the energy. In fact, one can show that they actually minimise it.

The Dirichlet boundary condition is not the only one which admits a variational treatment. Other boundary conditions (e.g., Neumann, Robin,...) can be treated variationally by adding suitable boundary terms to the action functional. This is explored in the problem set.

11.3.2 Minimal surfaces

Let us now do an example from Geometry. Let $f : D \subset \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto f(x,y)$, be a twice differentiable function. The graph of $z = f(x,y)$ defines a surface $\Sigma \subset \mathbb{R}^3$. The area of this surface is the functional of f given by

$$S[f] = \int_D \sqrt{1 + f_x^2 + f_y^2}\, dxdy \,,$$

where $f_x = \partial f/\partial x$ and $f_y = \partial f/\partial y$. If f is an extremal of this functional, we say that Σ is a **minimal surface**. To find the Euler–Lagrange equation, we first differentiate the lagrangian $L(f_x, f_y) = \sqrt{1 + f_x^2 + f_y^2}$. We notice that it does not depend explicitly on f, whence the Euler–Lagrange equation is simply

$$\frac{\partial}{\partial x}\frac{\partial L}{\partial f_x} + \frac{\partial}{\partial y}\frac{\partial L}{\partial f_y} = 0 \,.$$

Now,

$$\frac{\partial L}{\partial f_x} = \frac{f_x}{\sqrt{1 + f_x^2 + f_y^2}} \qquad \text{and} \qquad \frac{\partial L}{\partial f_y} = \frac{f_y}{\sqrt{1 + f_x^2 + f_y^2}} \,.$$

Into the Euler–Lagrange equation,

$$\begin{aligned}
0 &= \frac{\partial}{\partial x}\frac{\partial \mathrm{L}}{\partial f_x} + \frac{\partial}{\partial y}\frac{\partial \mathrm{L}}{\partial f_y} \\
&= \frac{\partial}{\partial x}\left(\frac{f_x}{\sqrt{1+f_x^2+f_y^2}}\right) + \frac{\partial}{\partial y}\left(\frac{f_y}{\sqrt{1+f_x^2+f_y^2}}\right) \\
&= \frac{f_{xx}}{\sqrt{1+f_x^2+f_y^2}} - \frac{f_x(f_x f_{xx}+f_y f_{xy})}{(1+f_x^2+f_y^2)^{3/2}} + \frac{f_{yy}}{\sqrt{1+f_x^2+f_y^2}} - \frac{f_y(f_x f_{xy}+f_y f_{yy})}{(1+f_x^2+f_y^2)^{3/2}} \\
&= \frac{f_{xx}(1+f_x^2+f_y^2) - f_x^2 f_{xx} - f_x f_y f_{xy} + f_{yy}(1+f_x^2+f_y^2) - f_x f_y f_{xy} - f_y^2 f_{yy})}{(1+f_x^2+f_y^2)^{3/2}} \\
&= \frac{(1+f_y^2)f_{xx} + (1+f_x^2)f_{yy} - 2f_x f_y f_{xy}}{(1+f_x^2+f_y^2)^{3/2}} \,.
\end{aligned}$$

Multiplying through by the denominator, which cannot be zero, we see that the Euler–Lagrange equation becomes the partial differential equation

$$(1+f_y^2)f_{xx} + (1+f_x^2)f_{yy} - 2f_x f_y f_{xy} = 0 \,. \tag{47}$$

The mathematical problem of finding a minimal surface with prescribed boundaries is called the **Plateau problem**, in honour of Joseph Plateau, a Belgian physicist, who studied the formation of soap bubbles. It turns out that what keeps soap bubbles in shape is the surface tension, which they try to minimise. The surface tension basically its area, so they tend to form minimal surfaces.

For example, suppose that we have two concentric rings of radius r a distance 2ℓ apart and we ask for the minimal surface with boundary these two rings. For example, a possible surface between these two rings could be the cylinder of radius r and length 2ℓ, which has an area $4\pi\ell r$. The axial symmetry of problem suggests that we try a surface of revolution: thus reducing the expected partial differential equation to an ordinary differential equation. This is typical and highlights the uses of symmetry in the solution of partial differential equations.

Let $y : [-\ell, \ell] \to \mathbb{R}$ be the profile of the surface of revolution, as shown in Figure 14. The boundary conditions are $y(\pm\ell) = r$.

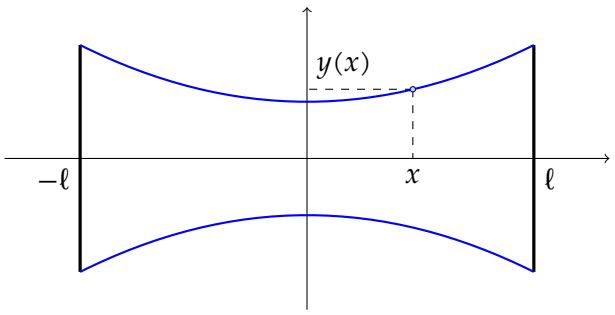


Figure 14: Minimal surface between two concentric rings

The area of the resulting surface of revolution is given by the functional

$$J[y] = \int_{-\ell}^{\ell} 2\pi y(x)\sqrt{1+y'(x)^2}dx\,,$$

which we have already met when we discussed the catenary in Section 9.4, except that there we supplemented it with an isoperimetric constraint, which is absent here. Still, we can reuse those calculations (with $\lambda = 0$) and arrive at the equation

$$y'^2 = (y/c)^2 - 1\,,$$

which has as general solution

$$y(x) = c\cosh\left(\frac{x-x_0}{c}\right)\,.$$

The resulting surface of revolution is called a **catenoid**. The boundary conditions say that $x_0 = 0$ and that

$$\cosh(\ell/c) = r/c\,.$$

The first thing one should notice about this equation is that it does *not* have solutions for some values of r/ℓ. Indeed, let $\zeta = \ell/c$. Then the above equation becomes

$$\cosh\zeta = \frac{r}{\ell}\zeta\,,$$

which has a solution if and only if the line through the origin of slope $\frac{r}{\ell}$ intersects the graph of $\cosh\zeta$. Since $\cosh\zeta \geq 1$, this will only happen if $\frac{r}{\ell} \geq \rho_c$, for some critical slope $\rho_c \simeq 1.51$, as shown in Figure 15.

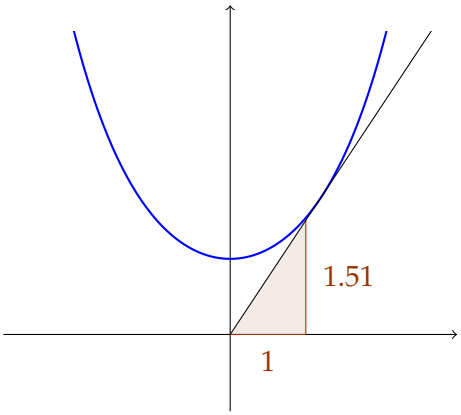


Figure 15: Critical slope

What happens for $\frac{r}{\ell} < \rho_c$*?* Alas, the calculus of variations does not tell us. The functional $J[y]$ might be minimised by functions y which are not in the class of functions we are considering. For example, it may well be the case that the two disks at $x = \pm\ell$ minimise the area, but this corresponds to a discontinuous $y(x)$. In fact, the two disks are not actually surfaces of revolution, so this analysis would never have found them. But there is a bigger difference still between these two surfaces: one is connected and the other one is not. Both extremise the functional, but one cannot go continuously from one to the other. One often says that the surfaces belong to different *topological sectors*.

(a) Boundary (b) Catenoid (c) Disks

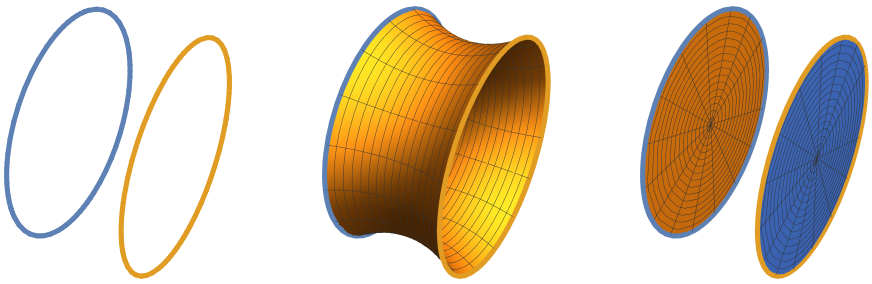


Figure 16: Two minimal surfaces with the same boundary

Exercise 11.2. *Show that half of the catenoid can be thought of as a graph of the function*

$$x(y,z) = c\,\operatorname{arccosh}\left(\frac{\sqrt{y^2+z^2}}{c}\right).$$

Verify that this function obeys the partial differential equation for a minimal surface, written as

$$(1+x_z^2)x_{yy} + (1+x_y^2)x_{zz} - 2x_y x_z x_{yz} = 0\,.$$

Exercise 11.3. *Let $u : \mathbb{R}^2 \to \mathbb{R}^2$ be a smooth vector field. Let $u_\mu(x)$, for $\mu = 1,2$ denote its component functions. Define $\phi_{\mu\nu} = \frac{\partial u_\nu}{\partial x^\mu} - \frac{\partial u_\mu}{\partial x^\nu}$, for all $\mu,\nu = 1,2$. Consider the following lagrangian*

$$\mathrm{L}(u,\nabla u) = \tfrac{1}{4}\sum_{\mu,\nu=1}^{2} (\phi_{\mu\nu})^2\,.$$

Prove that the Euler–Lagrange equation implies that $\phi_{\mu\nu}$ is constant.
Repeat this for $u : \mathbb{R}^m \to \mathbb{R}^m$ with the same definitions for $\phi_{\mu\nu}$ and $\mathrm{L}(u,\nabla u)$, except that now $\mu,\nu = 1,\dots,m$. Show that the Euler–Lagrange equation implies that $\phi_{\mu\nu}$ is ***harmonic****; that is,*

$$\triangle\phi_{\mu\nu} = 0 \qquad \text{for all } \mu,\nu,$$

where $\triangle = \sum_{\mu=1}^{m} \frac{\partial^2}{\partial x^\mu \partial x^\mu}$. (The converse is not true: the Euler–Lagrange equation imposes a weaker condition. Which one?)

12 Noether's theorem revisited

> A piece of penetrating mathematical thinking.
>
> — Albert Einstein

In this section we revisit Noether's theorem for the case of multidimensional lagrangians. The main difference is that instead of a conserved quantity, Noether's theorem results in a *Noether current*, which is a vector field whose divergence is zero when the Euler–Lagrange equations are satisfied. Throughout this section, we will let $\mathrm{D} \subset \mathbb{R}^m$ denote a bounded set with (piecewise) smooth boundary $\partial\mathrm{D}$, so that we can apply the Divergence Theorem, and we will let $u : \mathrm{D} \to \mathbb{R}^n$ denote a C^1 function which is also defined on the boundary.

12.1 First version of Noether's Theorem

Let $\mathrm{L}(u, \nabla u, x)$ be a lagrangian which is invariant under a one-parameter group of diffeomorphisms $\varphi_s : \mathbb{R}^n \to \mathbb{R}^n$, for $s \in \mathbb{R}$. Let $z : \mathrm{D} \times \mathbb{R} \to \mathbb{R}^n$ be defined by $z(x, s) = \varphi_s(u(x))$. Since the lagrangian is invariant, the same argument as in the proof of Lemma 7.1 says that if $u(x)$ solves the Euler–Lagrange equations, then so does $z(x, s)$ for all s.

Since $\mathrm{L}(z, \nabla z, x)$ is actually independent of s, taking the derivative with respect to s we get zero:

$$0 = \frac{d}{ds}\mathrm{L}(z, \nabla z, x) = \sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial z^i}\frac{\partial z^i}{\partial s} + \sum_{i=1}^{n}\sum_{\mu=1}^{m} \frac{\partial \mathrm{L}}{\partial z^i_\mu}\frac{\partial z^i_\mu}{\partial s} ,$$

where $z^i_\mu = \partial z^i / \partial x^\mu$. Using the Euler–Lagrange equations (46), we can rewrite this as

$$\begin{aligned} 0 &= \sum_{i=1}^{n}\sum_{\mu=1}^{m} \frac{\partial}{\partial x^\mu}\left(\frac{\partial \mathrm{L}}{\partial z^i_\mu}\right)\frac{\partial z^i}{\partial s} + \sum_{i=1}^{n}\sum_{\mu=1}^{m} \frac{\partial \mathrm{L}}{\partial z^i_\mu}\frac{\partial z^i_\mu}{\partial s} \\ &= \sum_{\mu=1}^{m} \frac{\partial}{\partial x^\mu}\left(\sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial z^i_\mu}\frac{\partial z^i}{\partial s}\right) . \end{aligned}$$

Finally we evaluate at $s = 0$, using that $z(x, 0) = u(x)$, to arrive at

$$\sum_{\mu=1}^{m} \frac{\partial}{\partial x^\mu}\left(\sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial u^i_\mu}\frac{\partial z^i}{\partial s}\Big|_{s=0}\right) = 0 .$$

We can state this result in the form of a theorem.

Theorem 12.1 (Multidimensional Noether's Theorem)**.** *Let*

$$\mathrm{S}[u] = \int_{\mathrm{D}} \mathrm{L}(u, \nabla u, x) d^m x$$

be an action functional for C^1 *maps* $u : D \subset \mathbb{R}^m \to \mathbb{R}^n$, *and let* L *be invariant under a one-parameter group of diffeomorphisms* $\{\varphi_s\}$ *of* $\mathbb{R}^n$. *Then the* ***Noether current*** $J : D \to \mathbb{R}^m$, *with components* J^μ *given by*

$$\boxed{J^\mu = \sum_{i=1}^{n} \frac{\partial L}{\partial u^i_\mu} \frac{\partial z^i}{\partial s}\bigg|_{s=0}}$$

is conserved; that is, its divergence vanishes:

$$\sum_{\mu=1}^{m} \frac{\partial J^\mu}{\partial x^\mu} = 0 \,.$$

You may recognise this conservation law as a **continuity equation**. Let us try to understand what this says. Let us assume that $m = 2$ so that the Noether current J is a vector field defined in a region $D \subset \mathbb{R}^2$ in the plane. Let C be any smooth simple closed curve in D. Let N denote the outward normal to C. Then using the Divergence Theorem and the fact that J is conserved, we deduce that

$$\int_C \langle N, J \rangle \, ds = 0 \,,$$

where ds is the infinitesimal arclength element in C. We can interpret $\langle N, J \rangle$ as the outward flux per unit length of the current J. The conservation law of the current simply says that the net flux is zero: as much flux comes into the region enclosed by C as comes out.

The following exercise is a continuation of Exercise 11.3.

Exercise 12.1. *In the notation of Exercise 11.3, let* $u : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^2$, $(x, s) \mapsto u_\mu(x, s)$, *be a one-parameter family of vector fields in* $\mathbb{R}^2$ *defined as follows:*

$$u_\mu(x, s) = u_\mu(x) + s \frac{\partial \Lambda}{\partial x^\mu} \,,$$

where $\Lambda : \mathbb{R}^2 \to \mathbb{R}$ *is a smooth function. Prove that the lagrangian*

$$L(u, \nabla u) = \tfrac{1}{4} \sum_{\mu,\nu=1}^{2} (\phi_{\mu\nu}(x, s))^2$$

is independent of s *and derive the expression for the corresponding Noether current. Prove directly (without appealing to Noether's theorem) that it is conserved when the Euler–Lagrange equations are satisfied.*
(Hint: The answer is

$$J^\mu = \sum_\nu \phi_{\mu\nu} \frac{\partial \Lambda}{\partial x^\nu} \,,$$

and its conservation follows trivially using the equations of motion.)

12.2 Noether's theorem for invariance under reparametrisations

Functionals may also be invariant under diffeomorphisms of the parameter space $D \subset \mathbb{R}^m$. For example, suppose that $\varphi_s : D \to D$ is a one-parameter family of

diffeomorphisms of D sending $x \in \mathrm{D}$ to $\bar{x}(x,s) = \varphi_s(x)$ such that the action functional

$$\mathrm{S}[u] = \int_{\mathrm{D}} \mathrm{L}(u, \nabla u, x) d^m x$$

is invariant. This means that

$$\mathrm{L}(u(\bar{x}), \bar{\nabla} u, \bar{x}) \left(\det \frac{\partial \bar{x}}{\partial x} \right) = \mathrm{L}(u, \nabla u, x) \ , \tag{48}$$

where $\det \frac{\partial \bar{x}}{\partial x}$ is the jacobian determinant of φ_s, which is positive. (Why?) Let s be small and let us expand $\bar{x}^\mu(x,s)$ in a Taylor series about $s = 0$:

$$\bar{x}^\mu(x,s) = x^\mu + \chi^\mu(x) s + \mathrm{O}(s^2) \ ,$$

so that $\chi^\mu = \left. \frac{\partial \bar{x}^\mu(x,s)}{\partial s} \right|_{s=0}$.

As we will show below in a more general situation still, differentiating equation (48) with respect to s at $s = 0$ one obtains

$$\sum_{\mu=1}^{m} \chi^\mu \frac{\partial \mathrm{L}}{\partial x^\mu} + \sum_{\mu=1}^{m} \frac{\partial \chi^\mu}{\partial x^\mu} \mathrm{L} - \sum_{i=1}^{n} \sum_{\mu,\nu=1}^{m} \frac{\partial \mathrm{L}}{\partial u^i_\mu} u^i_\nu \frac{\partial \chi^\nu}{\partial x^\mu} = 0 \ .$$

Exercise 12.2. *Show that the resulting Noether current is given by*

$$\mathrm{J}^\mu = \mathrm{L}\chi^\mu - \sum_{i=1}^{n} \frac{\partial \mathrm{L}}{\partial u^i_\mu} \sum_{\nu=1}^{m} u^i_\nu \chi^\nu$$

and show that it is divergenceless if u obeys the Euler–Lagrange equation.

12.3 The general Noether's Theorem

There is no reason why we should consider continuous symmetries which either act only on the dependent variables or only on the independent variables. In fact, we can consider one-parameter groups of symmetries which transform both sets of variables; that is,

$$u \mapsto \bar{u}(u,x,s) \qquad \text{and} \qquad x \mapsto \bar{x}(u,x,s) \ ,$$

for $s \in \mathbb{R}$. Suppose again that the action is invariant, so that the lagrangian obeys

$$\mathrm{L}(\bar{u}, \bar{\nabla} \bar{u}, \bar{x}) \left(\det \frac{d\bar{x}}{dx} \right) = \mathrm{L}(u, \nabla u, x) \ . \tag{49}$$

Notice that the Jacobian matrix is now written $\frac{d\bar{x}}{dx}$ and not $\frac{\partial \bar{x}}{\partial x}$, since the dependence of $\bar{x}(u(x),x,s)$ on x is not just explicit, but also through the dependence of u on x.

As before we expand around $s = 0$,

$$\begin{aligned} \bar{x}^\mu(u,x,s) &= x^\mu + \chi^\mu(u,x) s + \mathrm{O}(s^2) \\ \bar{u}^i(u,x,s) &= u^i + \zeta^i(u,x) s + \mathrm{O}(s^2) \ . \end{aligned}$$

The following discussion generalises that in Section 7.3 to higher dimensions; that is, when there are more than one independent variable.

We first calculate some further derivatives at $s = 0$:

$$\left.\frac{\partial \bar{u}^i}{\partial u^j}\right|_{s=0} = \delta^i_j \qquad \left.\frac{\partial \bar{u}^i}{\partial x^\mu}\right|_{s=0} = 0 \qquad \left.\frac{\partial \bar{x}^\mu}{\partial x^\nu}\right|_{s=0} = \delta^\mu_\nu \qquad \left.\frac{\partial \bar{x}^\mu}{\partial u^i}\right|_{s=0} = 0\,,$$

and

$$\left.\frac{\partial^2 \bar{u}^i}{\partial s \partial u^j}\right|_{s=0} = \frac{\partial \zeta^i}{\partial u^j} \qquad \left.\frac{\partial^2 \bar{x}^\mu}{\partial s \partial x^\nu}\right|_{s=0} = \frac{\partial \chi^\mu}{\partial x^\nu}$$
$$\left.\frac{\partial^2 \bar{u}^i}{\partial s \partial x^\mu}\right|_{s=0} = \frac{\partial \zeta^i}{\partial x^\mu} \qquad \left.\frac{\partial^2 \bar{x}^\mu}{\partial s \partial u^i}\right|_{s=0} = \frac{\partial \chi^\mu}{\partial u^i}\,,$$

where δ^i_j and δ^μ_ν are Kronecker deltas as in Section 7.3. We will use the shorthands u^i_μ for the derivative $\frac{\partial u^i}{\partial x^\mu}$ and similarly $\bar{u}^i_\mu$ for $\frac{\partial \bar{u}^i}{\partial \bar{x}^\mu}$. In addition we will let $\mathrm{D} := \det \frac{d\bar{x}}{dx}$.

We now differentiate equation (49) with respect to s. The right-hand side is independent of s, so we arrive at

$$\left(\sum_i \frac{\partial \mathrm{L}}{\partial \bar{u}^i}\frac{\partial \bar{u}^i}{\partial s} + \sum_i \sum_\mu \frac{\partial \mathrm{L}}{\partial \bar{u}^i_\mu}\frac{\partial \bar{u}^i_\mu}{\partial s} + \sum_\mu \frac{\partial \mathrm{L}}{\partial \bar{x}^\mu}\frac{\partial \bar{x}^\mu}{\partial s}\right)\mathrm{D} + \mathrm{L}\frac{\partial \mathrm{D}}{\partial s} = 0\,.$$

We now evaluate at $s = 0$, to obtain

$$\left(\sum_i \frac{\partial \mathrm{L}}{\partial u^i}\zeta^i + \sum_i \sum_\mu \frac{\partial \mathrm{L}}{\partial u^i_\mu}\left.\frac{\partial \bar{u}^i_\mu}{\partial s}\right|_{s=0} + \sum_\mu \frac{\partial \mathrm{L}}{\partial x^\mu}\chi^\mu\right)\mathrm{D}|_{s=0} + \mathrm{L}\left.\frac{\partial \mathrm{D}}{\partial s}\right|_{s=0} = 0\,. \quad (50)$$

There are three terms which we still to evaluate at $s = 0$. To this end let us calculate

$$\frac{d\bar{u}^i}{dx^\mu} = \frac{\partial \bar{u}^i}{\partial x^\mu} + \sum_j \frac{\partial \bar{u}^i}{\partial u^j}u^j_\mu \implies \left.\frac{d\bar{u}^i}{dx^\mu}\right|_{s=0} = u^i_\mu\,, \quad (51)$$

and

$$\frac{d\bar{x}^\mu}{dx^\nu} = \frac{\partial \bar{x}^\mu}{\partial x^\nu} + \sum_i \frac{\partial \bar{x}^\mu}{\partial u^i}u^i_\nu \implies \left.\frac{d\bar{x}^\mu}{dx^\nu}\right|_{s=0} = \delta^\mu_\nu\,, \quad (52)$$

whence we learn that $\mathrm{D}|_{s=0} = 1$, being the determinant of the identity matrix with entries δ^μ_ν. To calculate $\left.\frac{\partial \mathrm{D}}{\partial s}\right|_{s=0}$, we first differentiate the jacobian matrix itself with respect to s at $s = 0$:

$$\begin{aligned}\left.\frac{\partial}{\partial s}\frac{d\bar{x}^\mu}{dx^\nu}\right|_{s=0} &= \left.\frac{\partial^2 \bar{x}^\mu}{\partial s \partial x^\nu}\right|_{s=0} + \sum_i \left.\frac{\partial^2 \bar{x}^\mu}{\partial s \partial u^i}\right|_{s=0} u^i_\nu \\ &= \frac{\partial \chi^\mu}{\partial x^\nu} + \sum_i \frac{\partial \chi^\mu}{\partial u^i}u^i_\nu = \frac{d\chi^\mu}{dx^\nu}\,.\end{aligned}$$

Next we use that if a matrix $\mathrm{M}(s) = 1 + s\mathrm{A} + \mathrm{O}(s^2)$, then its determinant obeys $\det \mathrm{M}(s) = 1 + s\,\mathrm{tr}\,\mathrm{A} + \mathrm{O}(s^2)$, whence

$$\left.\frac{\partial \mathrm{D}}{\partial s}\right|_{s=0} = \sum_\mu \frac{d\chi^\mu}{dx^\mu}\,.$$

To compute the last remaining term in equation (50), we consider the following consequence of the chain rule

$$\frac{d\bar{u}^i}{dx^\mu} = \sum_\nu \bar{u}^i_\nu \frac{d\bar{x}^\nu}{dx^\mu} ,$$

which expands, using equations (51) and (52), to

$$\frac{\partial \bar{u}^i}{\partial x^\mu} + \sum_j \frac{\partial \bar{u}^i}{\partial u^j} u^j_\mu = \sum_\nu \bar{u}^i_\nu \left(\frac{\partial \bar{x}^\nu}{\partial x^\mu} + \sum_j \frac{\partial \bar{x}^\nu}{\partial u^j} u^j_\mu \right) .$$

Evaluating at $s = 0$ we find the identity

$$\bar{u}^i_\mu \big|_{s=0} = u^i_\mu ,$$

whereas differentiating with respect to s and then setting $s = 0$ we find

$$\frac{\partial \zeta^i}{\partial x^\mu} + \sum_j \frac{\partial \zeta^i}{\partial u^j} u^j_\mu = \left. \frac{\partial^2 \bar{u}^i}{\partial s \partial \bar{x}^\mu} \right|_{s=0} + \sum_\nu u^i_\nu \left(\frac{\partial \chi^\nu}{\partial x^\mu} + \sum_j \frac{\partial \chi^\nu}{\partial u^j} u^j_\mu \right) ,$$

from where we can solve for

$$\left. \frac{\partial^2 \bar{u}^i}{\partial s \partial \bar{x}^\mu} \right|_{s=0} = \frac{d\zeta^i}{dx^\mu} - \sum_\nu u^i_\nu \frac{d\chi^\nu}{dx^\mu} .$$

Finally, we are able to rewrite equation (50) as

$$\boxed{\sum_i \frac{\partial \mathrm{L}}{\partial u^i} \zeta^i + \sum_\mu \left(\chi^\mu \frac{\partial \mathrm{L}}{\partial x^\mu} + \frac{d\chi^\mu}{dx^\mu} \mathrm{L} + \sum_i \frac{\partial \mathrm{L}}{\partial u^i_\mu} \left(\frac{d\zeta^i}{dx^\mu} - \sum_{\nu=1}^m u^i_\nu \frac{d\chi^\nu}{dx^\mu} \right) \right) = 0} \quad . \tag{53}$$

☞ As mentioned already in Section 7.3, it follows from the theory of Lie groups that equation (53) is equivalent to equation (49) and usually much easier to verify.

We now relate partial and total derivatives by

$$\frac{\partial \mathrm{L}}{\partial x^\mu} = \frac{d\mathrm{L}}{dx^\mu} - \sum_i \frac{\partial \mathrm{L}}{\partial u^i} u^i_\mu - \sum_{i,\nu} \frac{\partial \mathrm{L}}{\partial u^i_\nu} u^i_{\mu\nu} ,$$

where we have introduced the shorthand $u^i_{\mu\nu}$ for the second partial derivative $\frac{\partial^2 u^i}{\partial x^\mu \partial x^\nu}$, to arrive at

$$\sum_\mu \frac{d}{dx^\mu} (\mathrm{L}\chi^\mu) + \sum_i \frac{\partial \mathrm{L}}{\partial u^i} \left(\zeta^i - \sum_\mu u^i_\mu \chi^\mu \right) + \sum_{i,\mu} \frac{\partial \mathrm{L}}{\partial u^i_\mu} \frac{d}{dx^\mu} \left(\zeta^i - \sum_\nu u^i_\nu \chi^\nu \right) = 0 .$$

We use the product rule in reverse in the last term to arrive at

$$\sum_\mu \frac{d}{dx^\mu} \left(\mathrm{L}\chi^\mu + \sum_i \frac{\partial \mathrm{L}}{\partial u^i_\mu} \left(\zeta^i - \sum_\nu u^i_\nu \chi^\nu \right) \right) + \sum_i \left(\frac{\partial \mathrm{L}}{\partial u^i} - \sum_\mu \frac{\partial}{\partial x^\mu} \frac{\partial \mathrm{L}}{\partial u^i_\mu} \right) \left(\zeta^i - \sum_\nu u^i_\nu \chi^\nu \right) = 0 .$$

Recognising the Euler–Lagrange equation in the second term, we see that if u obeys the Euler–Lagrange equation, the **Noether current**

$$\boxed{J^\mu = L\chi^\mu + \sum_i \frac{\partial L}{\partial u^i_\mu}\left(\zeta^i - \sum_\nu u^i_\nu \chi^\nu\right)}\,.$$

is divergenceless.

13 Classical field theory

> You may be a little disturbed at this stage by the importance that the time variable plays in the formalism.
>
> — Paul A.M. Dirac

The study of classical (as opposed to quantum) field theory is an important special case in the study of variational problems defined by multidimensional integrals.

In this course we discussed newtonian mechanics and its reformulation as a variational theory via Hamilton's principle of least action. Newtonian mechanics represents the birth of modern science and the newtonian paradigm held sway for more than two centuries in which the mechanistic view of the universe (the so-called *clockwork universe*) reigned supreme. Even Maxwell, who is directly responsible for the next rung in the ladder of the "ascent of man" [5], wrote his theory of electromagnetism in mechanical terms. His electric and magnetic fields were disturbances in an underlying medium: the *luminiferous aether*. This concept was abandoned for lack of experimental evidence and thanks to the conceptual breakthrough of Einstein's theory of relativity, after which the concept of the *(classical) field* as an independent entity took hold. In this section we will study field theories with a variational interpretation.

Classical fields — from now on, simply, **fields** — are functions $u : \mathbb{R}^{m+1} \to \mathbb{R}^n$, written $u(t, x^1, \dots, x^m)$. Typically $m = 3$, but the formalism does not require it. It is convenient to rename $t =: x^0$. As usual we will let ∇u denote the Jacobian matrix whose entries are the partial derivatives of the components of u relative to the x^μ, $\mu = 0, 1, \dots, m$.

We are concerned with action functionals of the form

$$S[u] = \int_\Omega \mathcal{L}(u, \nabla u, x) d^{m+1}x \,,$$

where $\mathcal{L}$ is called the **lagrangian density** and $\Omega = [0,1] \times \mathbb{R}^m$; that is, $\Omega = \{(x^0, x^1, \dots, x^m) \in \mathbb{R}^{m+1} | x^0 \in [0,1]\}$. Since the region Ω is not bounded, the integral $S[u]$ is improper and to guarantee its convergence it will be necessary to impose suitable decay conditions on u and ∇u. We will not explicitly state what these conditions are: they depend on value of m and the particular form of $\mathcal{L}$. We will assume that they are such that we may integrate by parts without picking up any "boundary" contributions. To make this more precise, notice that if the integral converges, then it agrees with the **Cauchy principal value**

$$S[u] = \lim_{R \to \infty} \int_{\Omega_R} \mathcal{L}(u, \nabla u, x) d^{m+1}x \,,$$

where $\Omega_R = [0,1] \times B_R$, where $B_R = \left\{(x^1, \dots, x^m) \middle| \sum_{i=1}^m (x^i)^2 \leq R^2\right\}$ is the ball of radius R centred at 0 in $\mathbb{R}^m$. Integrating by parts, we pick boundary contributions from the sphere $S_R = \left\{(x^1, \dots, x^m) \middle| \sum_{i=1}^m (x^i)^2 = R^2\right\}$ of radius R, and the idea is that such contributions should vanish in the limit $R \to \infty$.

As a motivating example, consider the **Klein–Gordon equation** in $\mathbb{R}^4$ (so $m = 3$),

$$\frac{\partial^2 u}{(\partial x^0)^2} - \sum_{i=1}^{3} \frac{\partial^2 u}{(\partial x^i)^2} + M^2 u = 0 \,,$$

where $M > 0$ is called the **mass**. If $M = 0$, then this is the wave equation, whence the Klein–Gordon equation is a sort of massive wave equation. It is convenient to rewrite this in a more succinct way by introducing x^μ, $\mu = 0, 1, 2, 3$, and the symmetric matrix

$$\eta^{\mu\nu} = \begin{pmatrix} -1 & 0 \\ 0 & \mathbb{1}_3 \end{pmatrix} \,,$$

and also letting $\partial_\mu = \frac{\partial}{\partial x^\mu}$ be the partial derivatives, so that the Klein–Gordon equation can be rewritten as

$$\sum_{\mu,\nu=0}^{3} \eta^{\mu\nu} \partial_\mu \partial_\nu u - M^2 u^2 = 0 \,.$$

We will show that this equation is the Euler–Lagrange equation of an action functional for the classical field u.

13.1 Calculus of variations with improper integrals

As before, let $\Omega = [0, 1] \times \mathbb{R}^m$ and consider the action

$$S[u] = \int_\Omega \mathcal{L}(u, \nabla u, x) d^{m+1} x \,,$$

where $u : \Omega \to \mathbb{R}^n$ is C^1 decaying sufficiently fast so that the integral exists. We will consider variations $\varepsilon : \Omega \to \mathbb{R}^n$ which obey

$$\varepsilon(t, x^1, \dots, x^m) = 0 \qquad \text{for } (x^1)^2 + \cdots + (x^m)^2 > R^2$$

for some R and also

$$\varepsilon(0, x^1, \dots, x^m) = \varepsilon(1, x^1, \dots, x^m) = 0 \qquad \text{for all } x^1, \dots, x^m.$$

Let us vary the action (which we write as the Cauchy principal value)

$$\begin{aligned} \frac{d}{ds} S[u + s\varepsilon]\Big|_{s=0} &= \lim_{R\to\infty} \int_{\Omega_R} \frac{d}{ds} \mathcal{L}(u + s\varepsilon, \nabla u + s\nabla\varepsilon, x)\Big|_{s=0} d^{m+1} x \\ &= \lim_{R\to\infty} \int_{\Omega_R} \sum_{i=1}^{n} \left(\frac{\partial \mathcal{L}}{\partial u^i} \varepsilon^i + \sum_{\mu=0}^{m} \frac{\partial \mathcal{L}}{\partial u^i_\mu} \partial_\mu \varepsilon^i \right) d^{m+1} x \\ &= \lim_{R\to\infty} \int_{\Omega_R} \sum_{i=1}^{n} \left(\frac{\partial \mathcal{L}}{\partial u^i} - \sum_{\mu=0}^{m} \partial_\mu \frac{\partial \mathcal{L}}{\partial u^i_\mu} \right) \varepsilon^i d^{m+1} x \,, \end{aligned}$$

where we have omitted a boundary term

$$\lim_{R\to\infty} \int_{\Omega_R} \sum_{\mu=0}^{m} \partial_\mu \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial u^i_\mu} \varepsilon^i d^{m+1} x \,,$$

which is seen to vanish after applying the Divergence Theorem and using the boundary conditions on the variation ε. Finally we use the Fundamental Lemma, which was proved for such ε, to deduce the Euler–Lagrange equation

$$\boxed{\frac{\partial \mathscr{L}}{\partial u^i} = \sum_{\mu=0}^{m} \partial_\mu \frac{\partial \mathscr{L}}{\partial u^i_\mu}} \,. \tag{54}$$

For example, consider the action functional for $u : \Omega \to \mathbb{R}$ given by the lagrangian density

$$\mathscr{L} = -\tfrac{1}{2} \sum_{\mu,\nu=0}^{3} \eta^{\mu\nu} \partial_\mu u \partial_\nu u - \tfrac{1}{2} M^2 u^2 \,. \tag{55}$$

The resulting Euler–Lagrange equation is obtained from

$$\frac{\partial \mathscr{L}}{\partial u} = -M^2 u \qquad \text{and} \qquad \frac{\partial \mathscr{L}}{\partial u_\mu} = -\sum_{\nu=0}^{3} \eta^{\mu\nu} \partial_\nu u \,,$$

and is seen to agree with the Klein–Gordon equation.

13.2 Noether's theorem revisited

The multidimensional Noether Theorem can again be understood as giving rise to conserved quantities. Recall that if the action functional is invariant under a continuous one-parameter symmetry, then there is a Noether current J^μ which has zero divergence when the Euler–Lagrange equation is satisfied. Let us consider an action functional for a classical field $u : \Omega \to \mathbb{R}^n$ which is invariant under a continuous one-parameter symmetry with Noether current J^μ. Let us integrate the (zero) divergence of the current on a cylindrical region $C_R = [a, b] \times B_R \subset \Omega$ and apply the divergence theorem:

$$0 = \int_{C_R} \sum_{\mu=0}^{m} \partial_\mu J^\mu d^{m+1}x = \int_{\partial C_R} \sum_{\mu=0}^{m} N_\mu J^\mu d^m x \,,$$

where N is the outward normal to the boundary ∂C_R. What is the boundary of C_R? It consists of the "sides" $[a, b] \times S_R$ of the cylinder, where S_R is the sphere $(x^1)^2 + \cdots + (x^m)^2 = R^2$, and the top and bottom caps of the cylinder $\{b\} \times B_R$ and $\{a\} \times B_R$, respectively. We can rewrite the above equation as

$$0 = \int_{B_R} J^0(b, x^1, \dots, x^m) d^m x - \int_{B_R} J^0(a, x^1, \dots, x^m) d^m x + \int_a^b \int_{S_R} \sum_{\mu=1}^{m} N^\mu J^\mu(x^0, x^1, \dots, x^m) d^{m+1}x \,,$$

where we have used that the outward normal on the bottom cap of the cylinder is pointing along the negative x^0-axis, which accounts for the sign of the second

term. We now take the limit $R \to \infty$. Using the boundary conditions on the field, which mean that $J^\mu \to 0$ on $[a,b] \times S_R$ as $R \to \infty$, we see that the last term vanishes and hence that

$$\lim_{R\to\infty} \int_{B_R} J^0(b, x^1, \dots, x^m) d^m x = \lim_{R\to\infty} \int_{B_R} J^0(a, x^1, \dots, x^m) d^m x \, .$$

Since a, b were arbitrary, that says that the quantity

$$\boxed{Q = \lim_{R\to\infty} \int_{B_R} J^0(t, x^1, \dots, x^m) d^m x}$$

is constant in t, so it is conserved.

13.3 Maxwell equations

Maxwell equations are familiar to every student in this course: you step on them every time you make your way to the JCMB Lecture Theatres. We are going to see that they too have a variational origin. Let us write Maxwell equations in Heaviside form:

$$\begin{aligned} \boldsymbol{\nabla} \cdot \mathbf{B} &= 0 & \boldsymbol{\nabla} \cdot \mathbf{E} &= \rho \\ \boldsymbol{\nabla} \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} & \boldsymbol{\nabla} \times \mathbf{B} &= \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J} \, , \end{aligned}$$

where $\mathbf{B}$ is the **magnetic field**, $\mathbf{E}$ is the **electric field**, ρ is the **electric charge density** and $\mathbf{J}$ is the **electric current density**. Let us concentrate on the two equations which do not involve ρ or $\mathbf{J}$. The equation $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$ can be solved by writing $\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}$, where $\mathbf{A}$ is called the **magnetic potential**. This does not determine $\mathbf{A}$ uniquely, since we can add to $\mathbf{A}$ the gradient of any function without altering $\mathbf{B}$. The ambiguity

$$\mathbf{A} \mapsto \mathbf{A} + \boldsymbol{\nabla}\theta \, ,$$

is called a **gauge transformation**, a concept due to Weyl in a different context. Inserting $\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}$ into the other Maxwell equation independent from the sources $\rho, \mathbf{J}$, we find

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial}{\partial t} \boldsymbol{\nabla} \times \mathbf{A} = -\boldsymbol{\nabla} \times \frac{\partial \mathbf{A}}{\partial t} \implies \boldsymbol{\nabla} \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right) = 0 \, .$$

This means that there exists a function ϕ, called the **electric potential**, such that

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\boldsymbol{\nabla}\phi \, .$$

If we perform a gauge transformation on $\mathbf{A}$, $\mathbf{E}$ will change unless we also transform ϕ as

$$\phi \mapsto \phi - \frac{\partial \theta}{\partial t} \, .$$

In summary, two of Maxwell equations can be solved by

$$\mathbf{B} = \nabla \times \mathbf{A} \qquad \text{and} \qquad \mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \nabla \phi \,,$$

where $\mathbf{A}$ and ϕ are defined up to gauge transformations

$$\mathbf{A} \mapsto \mathbf{A} + \nabla \theta \qquad \text{and} \qquad \phi \mapsto \phi - \frac{\partial \theta}{\partial t} \,,$$

for some function θ.

This ambiguity in the definition of $\mathbf{A}$ and ϕ can be exploited to impose the **Lorenz gauge** condition

$$\nabla \cdot \mathbf{A} + \frac{\partial \phi}{\partial t} = 0 \,.$$

Exercise 13.1. *Show that in the Lorenz gauge, the remaining two Maxwell equations become wave equations with sources:*

$$\frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi = \rho \qquad \textit{and} \qquad \frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla^2 \mathbf{A} = \mathbf{J} \,.$$

We will now show that Maxwell equations are variational. Let $\rho = 0$ and $\mathbf{J} = 0$ at first. We consider the lagrangian density

$$\mathcal{L} = \tfrac{1}{2} \left(\|\mathbf{E}\|^2 - \|\mathbf{B}\|^2 \right) \,,$$

but thought of as a lagrangian density for $\mathbf{A}$ and ϕ. In other words, using the expressions for $\mathbf{E}$ and $\mathbf{B}$ in terms of $\mathbf{A}$ and ϕ,

$$\mathcal{L} = \tfrac{1}{2} \sum_{i=1}^{3} \left((\partial_i \phi)^2 + (\partial_t A_i)^2 + 2 \partial_i \phi \partial_t A_i \right) - \tfrac{1}{2} \sum_{i,j=1}^{3} \left((\partial_i A_j)^2 - \partial_i A_j \partial_j A_i \right) \,,$$

where we have written $\mathbf{A} = (A_1, A_2, A_3)$ and used the shorthands $\partial_t = \frac{\partial}{\partial t}$ and $\partial_i = \frac{\partial}{\partial x^i}$ for $i = 1, 2, 3$. We find

$$\frac{\partial \mathcal{L}}{\partial (\partial_i \phi)} = \partial_i \phi + \partial_t A_i = -E_i \qquad \frac{\partial \mathcal{L}}{\partial (\partial_t A_i)} = \partial_i \phi + \partial_t A_i = -E_i$$

and

$$\frac{\partial \mathcal{L}}{\partial (\partial_i A_j)} = -\partial_i A_j + \partial_j A_i = -\sum_{k=1}^{3} \epsilon_{ijk} B_k \,,$$

where $\mathbf{E} = (E_1, E_2, E_3)$ and $\mathbf{B} = (B_1, B_2, B_3)$ and ϵ_{ijk} is the Levi-Civita symbol. Since the lagrangian density does not depend explicitly on $\mathbf{A}$ or ϕ, but only on their derivatives, the Euler–Lagrange equations are

$$\sum_{i=1}^{3} \partial_i \frac{\partial \mathcal{L}}{\partial (\partial_i \phi)} = 0 \implies \sum_{i=1}^{3} \partial_i E_i = 0 \,,$$

and

$$\sum_{i=1}^{3} \partial_j \frac{\partial \mathscr{L}}{\partial(\partial_j A_i)} + \partial_t \frac{\partial \mathscr{L}}{\partial(\partial_t A_i)} = 0 \implies (\boldsymbol{\nabla} \times \mathbf{B})_i - \partial_t E_i = 0 \,,$$

which are precisely the two remaining Maxwell equations when $\rho = 0$ and $\mathbf{J} = 0$.

Exercise 13.2. *Show that Maxwell equations with* ρ *and* $\mathbf{J}$ *nonzero can be be obtained from the modified lagrangian density*

$$\mathscr{L} = \tfrac{1}{2}\left(\|\mathbf{E}\|^2 - \|\mathbf{B}\|^2\right) + \mathbf{A} \cdot \mathbf{J} - \phi\rho \,.$$

We can rewrite the Maxwell lagrangian density in a more succinct way by introducing the **electromagnetic 4-potential**

$$A_\mu = (-\phi, \mathbf{A}) \,,$$

where $\mu = 0, 1, 2, 3$, so that $A_0 = -\phi$ and the **4-current**

$$J_\mu = (-\rho, \mathbf{J}) \,,$$

so that $J_0 = -\rho$. We start by defining the **fieldstrength**

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \,,$$

which obeys $F_{\mu\nu} = -F_{\nu\mu}$. This allows us to think of $F_{\mu\nu}$ as the entries of a 4×4 antisymmetric matrix:

$$[F_{\mu\nu}] = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix} ,$$

where we have used that

$$F_{0i} = \partial_0 A_i + \partial_i \phi = -E_i \qquad \text{and} \qquad F_{ij} = \partial_i A_j - \partial_j A_i = \sum_{k=1}^{3} \epsilon_{ijk} B_k \,.$$

In terms of the fieldstrength, the lagrangian for Maxwell's equations can be written as

$$\boxed{\mathscr{L} = -\tfrac{1}{4} \sum_{\mu,\nu=0}^{3} F_{\mu\nu} F^{\mu\nu} + \sum_{\mu=0}^{3} \eta^{\mu\nu} A_\mu J_\nu} \,, \tag{56}$$

where we have "raised the indices" of $F_{\mu\nu}$ with $\eta^{\mu\nu}$ as follows:

$$F^{\mu\nu} = \sum_{\sigma,\tau=0}^{3} \eta^{\mu\sigma} \eta^{\nu\tau} F_{\sigma\tau} \,.$$

Exercise 13.3. *Show that the Euler–Lagrange equations of the lagrangian density* (56) *are given by*

$$-\sum_{\mu,\nu=0}^{3}\eta^{\mu\nu}\partial_\mu F_{\nu\rho}=J_\rho\,.$$

Notice that in this formulation, gauge transformations are

$$A_\mu\mapsto A_\mu+\partial_\mu\theta\,,$$

under which $F_{\mu\nu}$ remains invariant. In the absence of sources, so when $J_\mu=0$, the lagrangian density is **gauge invariant**.

Exercise 13.4. *Consider a fixed function* θ *and a one-parameter family of gauge transformations* $A_\mu\mapsto A_\mu+s\partial_\mu\theta$. *This is a continuous symmetry of the lagrangian density with* $J_\mu=0$. *What is the corresponding Noether current?*

13.4 The abelian Higgs model

We have seen two examples of action functionals for field theories: the Klein–Gordon lagrangian density in equation (55) and the Maxwell lagrangian density in equation (56). We will now "couple" them to construct the so-called *abelian Higgs model*, which earned Peter Higgs (jointly with the Belgian physicist François Englert) the Nobel Prize in Physics in 2013, half a century after the papers [6, 7] were written.

We will consider the Maxwell lagrangian density without sources, i.e., with $J_\mu=0$. As we will see after we couple it to the Klein–Gordon fields, these will act as sources for the Maxwell field. The attentive reader will have noticed that I wrote "Klein–Gordon fields", and that's because we will consider not one but two Klein–Gordon fields u,v, which we will think as the real and imaginary parts of a complex Klein–Gordon field $\varphi=u+iv$. This is purely for notational convenience. The Klein–Gordon lagrangian density (assuming that the mass M is the same for both u and v) is then

$$\mathcal{L}=-\tfrac{1}{2}\sum_{\mu,\nu=0}^{3}\eta^{\mu\nu}\overline{\partial_\mu\varphi}\partial_\nu\varphi-\tfrac{1}{2}M^2|\varphi|^2\,,$$

where $|\cdot|$ denotes the modulus of a complex number and the bar denotes complex conjugation. Notice that this lagrangian density is invariant under the following continuous symmetry

$$\varphi\mapsto e^{i\theta}\varphi\,, \tag{57}$$

where $\theta\in\mathbb{R}$.

Exercise 13.5. *Show that resulting Noether current is given by*

$$J_\mu=(u\partial_\mu v-v\partial_\mu u)\,. \tag{58}$$

From now on, we will use the shorthand notation

$$|\partial\varphi|^2 := \sum_{\mu,\nu=0}^{3} \eta^{\mu\nu}\overline{\partial_\mu\varphi}\partial_\nu\varphi \,,$$

although despite the suggestive notation it is not a positive quantity, due to the fact that $\eta^{\mu\nu}$ is not positive-definite. In fact, more generally, if V_μ is a complex vector,

$$|V|^2 := \sum_{\mu,\nu=0}^{3} \eta^{\mu\nu}\overline{V_\mu}V_\nu \,.$$

We will modify the above Klein–Gordon lagrangian density by replacing the mass term $\frac{1}{2}M^2|\varphi|^2$ by a *Mexican hat* potential

$$V(|\varphi|) = \lambda(|\varphi|^2 - \rho^2)^2 \,,$$

where λ, ρ are positive real numbers. The potential is still invariant under the continuous transformation (57), so we still have a continuous symmetry and hence the same Noether current.

The reason for the moniker should be clear after plotting the function. In Figure 17 we see the profile, that is, the graph of $V(|\varphi|)$ as a function of $|\varphi|$, and the potential itself, which is the surface of revolution with that profile. The potential has an unstable critical point at $\varphi = 0$ and a circle of semistable critical points at $|\varphi| = \rho$. These are the essential characteristics in the model: any other potential with a similar behaviour would work equally well.

(a) Profile (b) Potential

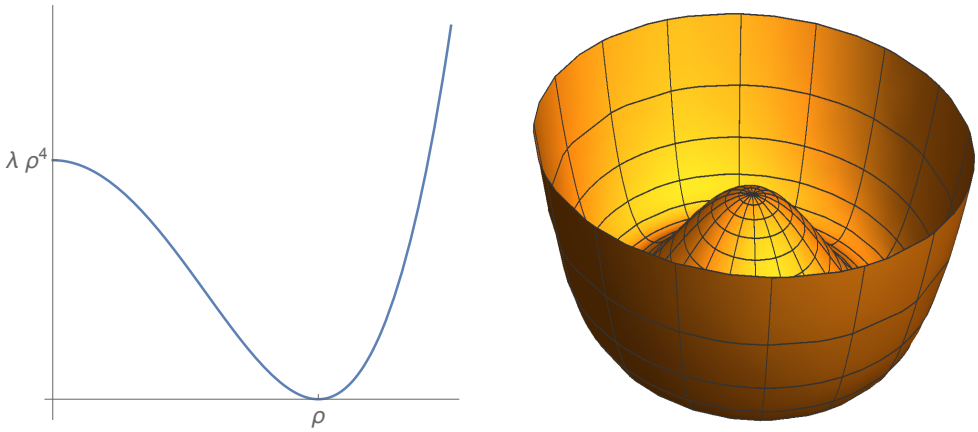


Figure 17: The Mexican hat potential

We now couple the two systems. As a first step we can simply add the lagrangian densities to arrive at

$$\mathcal{L}_0 = -\tfrac{1}{4}\sum_{\mu,\nu=0}^{3} F^{\mu\nu}F_{\mu\nu} - \tfrac{1}{2}|\partial\varphi|^2 - V(|\varphi|) \,.$$

This is not quite coupled, because there are no terms which involve both A_μ and φ. To write such terms we let gauge invariance guide us. Recall that $\varphi \mapsto e^{i\theta}\varphi$ is a symmetry of the Klein–Gordon action provided that θ is a constant, whereas $A_\mu \mapsto A_\mu + \partial_\mu\theta$ is an invariance of the Maxwell action for θ an arbitrary differentiable function. What happens if we demand that the coupled action be invariant also under $\varphi \mapsto e^{i\theta}\varphi$ when θ is a function? The potential term $V(|\varphi|)$ is still invariant, since $|\varphi|$ does not change even if θ is a function. However the term with a derivative on φ does change:

$$\partial_\mu\varphi \mapsto e^{i\theta}(\partial_\mu\varphi + i\varphi\partial_\mu\theta) \,.$$

But we notice that it changes by a term proportional to $\partial_\mu\theta$ which is precisely the way A_μ changes. This means that the combination $D_\mu\varphi := \partial_\mu\varphi - iA_\mu\varphi$ changes in a nice way under both $\varphi \mapsto e^{i\theta}\varphi$ and $A_\mu \mapsto A_\mu + \partial_\mu\theta$:

$$D_\mu\varphi \mapsto e^{i\theta}D_\mu\varphi \,,$$

and hence $|D\varphi|^2$ is invariant! In summary, the lagrangian density

$$\boxed{\mathcal{L} = -\tfrac{1}{4}\sum_{\mu,\nu=0}^{3} F^{\mu\nu}F_{\mu\nu} - \tfrac{1}{2}|D\varphi|^2 - V(|\varphi|)} \tag{59}$$

is gauge invariant: it defines the **abelian Higgs model**.

The replacement of $\partial_\mu\varphi$ by $D_\mu\varphi = \partial_\mu\varphi - iA_\mu\varphi$ in the lagrangian density is known as **minimal coupling**.

Exercise 13.6. *Show that the Euler–Lagrange equation for the electromagnetic field* A_μ *in the abelian Higgs model is given by*

$$-\sum_{\mu,\nu=0}^{3} \eta^{\mu\nu}\partial_\mu F_{\nu\rho} = J_\rho - |\varphi|^2 A_\rho \,,$$

where J_ρ *is the Noether current* (58) *associated to the continuous symmetry of the Klein–Gordon action.*

The point $\varphi = 0$ corresponds to an unstable equilibrium for the potential. Indeed, expanding the potential around $\varphi = 0$ and comparing with the Klein–Gordon action we find a *negative* mass:

$$V \approx \lambda\rho^4 - 2\lambda\rho^2|\varphi|^2 + \cdots$$

Let us instead expand the lagrangian density around one of the semistable minima. To do this let us parametrise φ as follows:

$$\varphi = (\rho + h)e^{-i\theta/\rho} \,,$$

where h and θ are real fields. Let us write the lagrangian density in terms of h and θ. We find

$$D_\mu\varphi = (\partial_\mu - iA_\mu)\left((\rho+h)e^{-i\theta/\rho}\right) = e^{-i\theta/\rho}\left(\partial_\mu h - i(A_\mu + \tfrac{1}{\rho}\partial_\mu\theta)(\rho+h)\right) .$$

Let us define $C_\mu := A_\mu + \frac{1}{\rho}\partial_\mu\theta$. This is a gauge transformation of A_μ, hence the fieldstrength $F_{\mu\nu} = \partial_\mu C_\nu - \partial_\nu C_\mu$. In terms of C_μ,

$$D_\mu\varphi = e^{i\theta/\rho}(\partial_\mu h - iC_\mu(\rho+h)) ,$$

whence

$$|D\varphi|^2 = |\partial h|^2 + (\rho+h)^2\,|C|^2 ,$$

where we have used that h, ρ and C_μ are real. Similarly, the potential becomes $V(|\varphi|) = V(\rho+h)$. In these new "coordinates", the lagrangian density of the abelian Higgs model becomes

$$\mathcal{L} = -\tfrac{1}{4}\sum_{\mu,\nu=0}^{3} F^{\mu\nu}F_{\mu\nu} - \tfrac{1}{2}\,|\partial h|^2 - \tfrac{1}{2}(\rho+h)^2|C|^2 - \lambda((\rho+h)^2-\rho^2)^2 .$$

It looks more complicated than before, but notice that it has one less field: θ no longer appears. The two fields are the **Proca field** C_μ and the **Higgs field** h. Expanding to quadratic order (this allows us to read off the masses of the fields), we find

$$\mathcal{L} = \underbrace{-\tfrac{1}{4}\sum_{\mu,\nu=0}^{3} F^{\mu\nu}F_{\mu\nu} - \tfrac{1}{2}\rho^2|C|^2}_{\text{massive vector field (Proca)}} - \underbrace{\tfrac{1}{2}\,|\partial h|^2 - 4\lambda\rho^2h^2}_{\text{massive scalar field (Higgs)}} + \cdots ,$$

whence we read off that C_μ has mass ρ and h has mass $\sqrt{8\lambda}\rho$.

In summary, the Higgs mechanism has broken the gauge symmetry of the abelian Higgs model, given mass to the "gauge field" and as a tell-tale sign leaves behind a "Higgs boson," which was found in 2012!

Figure 18: Higgs Nobel medal

References

[1] R. Montgomery, "A new solution to the three-body problem," *Notices Amer. Math. Soc.* **48** (2001), no. 5, 471–481.

[2] I. M. Gelfand and S. V. Fomin, *Calculus of variations*. Revised English edition translated and edited by Richard A. Silverman. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.

[3] W. B. Gordon, "On the completeness of Hamiltonian vector fields," *Proc. Amer. Math. Soc.* **26** (1970) 329–331.

[4] M. Puta, "The completeness of some Hamiltonian vector fields on a Poisson manifold," *An. Univ. Timişoara Ser. Mat.-Inform.* **32** (1994), no. 2, 93–98.

[5] J. Bronowski, *The ascent of man*. London : British Broadcasting Corporation, 1973.

[6] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys.Rev.Lett.* **13** (1964) 508–509.

[7] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys.Rev.Lett.* **13** (1964) 321–323.

[8] K. R. Symon, *Mechanics*. 2nd ed. Addison-Wesley Series in Physics. Addison-Wesley Publishing Co., Inc., Reading, Mass.-London, 1960.

[9] L. D. Landau and E. M. Lifshitz, *Mechanics*. Course of Theoretical Physics, Vol. 1. Translated from the Russian by J. B. Bell. Pergamon Press, Oxford-London-New York-Paris; Addison-Wesley Publishing Co., Inc., Reading, Mass., 1960.

[10] D. Lovelock and H. Rund, *Tensor, differential forms, and variational principles*. Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, 1975. Pure and Applied Mathematics.

[11] V. I. Arnol'd, *Mathematical methods of classical mechanics*, vol. 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 199? Translated from the 1974 Russian original by K. Vogtmann and A. Weinstein, Corrected reprint of the second (1989) edition.

[12] J. L. Troutman, *Variational calculus and optimal control*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, second ed., 1996. With the assistance of William Hrusa, Optimization with elementary convexity.

[13] B. van Brunt, *The calculus of variations*. Universitext. Springer-Verlag, New York, 2004.

[14] H. Goldstein, *Classical mechanics*. Addison-Wesley Publishing Co., Reading, Mass., second ed., 1980. Addison-Wesley Series in Physics.

[15] H. H. Goldstine, *A history of the calculus of variations from the 17th through the 19th century*, vol. 5 of *Studies in the History of Mathematics and Physical Sciences*. Springer-Verlag, New York-Berlin, 1980.

[16] O. de Oliveira, "The implicit and inverse function theorems: easy proofs," *Real Anal. Exchange* **39** (2013), no. 1, 207–218.

Index