Preprint typeset in JHEP style - PAPER VERSION

Easter 2018

Variational Principles
Part 1A Mathematics Tripos

Prof. P.K. Townsend

DAMTP, CMS
Wilberforce Road, Cambridge CB3 0WA
E-mail: P.K.Townsend@damtp.cam.ac.uk

Abstract:

See schedules. There are two example sheets.

RECOMMENDED BOOKS:

(1) D.S. Lemons, *Perfect Form*, Princeton University Press.

(2) I.M. Gelfand and S.V. Fomin, *Calculus of Variations*, Dover

Contents

1. Variational problems and variational principles

We often want to know how to maximize or minimize some quantity by varying other quantities on which it depends. Problems of this kind arise in many ways. Given a finite number of things to vary, it may be possible to formulate the problem mathematically so that the solution involves finding the maximum or minimum of a function of many variables; this is the first topic that we shall study in this course. However, the main focus will be on solving such "variational problems" when the quantity to be maximized or minimised depends on a continuous infinity of variables.

An example from antiquity is Dido's *isoperimetric* problem: how do you maximise an area given a constraint on its perimeter (Q.I.12). To solve it we need to consider how the area behaves under arbitrary variations of the function used to describe the perimeter curve. Another variational problem, posed and solved by Newton in his *Principia* of 1687 (although his method of solution was given only in an appendix to a later edition) is to find the shape of a ship's hull that minimises the drag as it moves through water. At least, that was Newton's stated motivation for the problem he actually solved, which was the first non-trivial variational problem for which a correct solution was found.

A simpler but more famous variational problem from around the same time is the *Brachistochrone* problem, initially posed (but not correctly solved) by Galileo. The brachistochrone is the curve assumed by a frictionless wire that minimises the time for a bead on it to fall from rest to some horizontally displaced point (see Q.I.9 for a closely related problem). Apparently unaware of Galileo's efforts, this problem was posed in 1696 by Johann Bernoulli as a challenge to the other mathematicians of Europe, especially his brother Jacob. His brother solved it, as did Leibnitz and Newton. Newton published his (geometric) solution anonymously, but "the lion is known by his claw" said Johann Bernoulli.

Work on generalizations of this problem and other variational problems, such as one posed as a revenge challenge by Jacob Bernouilli, eventually led, in 1744, to a treatise by Euler that systematised the methods of solution, which coupled calculus with geometrical reasoning. In 1745, the 19 year old Lagrange wrote to Euler to describe a general method that did not rely on geometrical insight. Euler's response was to abandon his methods in favour of those of Lagrange, which he called "the calculus of variations". Much of this course will be about the calculus of variations, essentially as presented by Lagrange in his *Mechánique Analytique* of 1788, which recast mechanics in terms of differential equations. Lagrange was proud of the fact that this work contains no diagrams; this is in stark contrast to Newton's *Principia*, which contains no equations[1].

[1]Newton's second law, for example, is expressed in words that are equivalent to $\mathbf{F} = \dot{\mathbf{p}}$ in modern notation; the dot notation for time derivative was introduced by Newton in earlier works.

In the physical sciences, many variational problems arise from the application of a variational principle. The first variational principle was formulated about 2000 years ago, by Hero of Alexandria. If an object is viewed in a plane mirror then we can trace a ray from the object to the eye, bouncing off the mirror. Hero stated, as a principle, that the ray's path is the shortest one, and he deduced from this principle that the angles of incidence and reflection (the angles that the incoming and outgoing rays make with the normal to the mirror) are equal. It sounds like a reasonable principle. After all, light travels in a straight line in the absence of mirrors, and a straight line is the shortest path between two points. Hero was just generalizing this idea to include mirrors. His principle is indeed valid for plane mirrors but light rays bouncing off curved mirrors don't always take the shortest path! A counter-example can be found on p.8 of the book by Lemons. However, even for curved mirrors it is still true that the path length is unchanged to first order by a small change in the path, so we could reformulate the principle: light rays travel on a path with a length that is stationary with respect to small changes of the path.

If light travels with finite speed then the shortest path is also the one for which the travel time is shortest. In the mid 17th century Fermat proposed this as the fundamental principle governing light rays, and applied it to refraction as well as reflection. In particular, he used it in 1662 to show that when a light ray crosses a boundary from one transparent medium to another, the angles of incidence and refraction are such that

$$\frac{\sin\theta_1}{v_1} = \frac{\sin\theta_2}{v_2}$$

where v_i is the velocity of light in the ith medium. This is usually called Snell's law of refraction. It can also be written as

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

where $n_i \propto 1/v_i$ are the refractive indices[2].

Fermat's principle of least time can also be applied to a medium with a varying index of refraction $n(\mathbf{x})$. In this case the principle is equivalent to the statement that the path taken is one that minimises the "optical path length"

$$P = \int_p n(\mathbf{x}) d\ell \,,$$

where the integral is over a specified path p with length element $d\ell$. The value of the integral can vary continuously as we vary the path, so applications of Fermat's principle to media with known (or proposed) variable refractive indices give rise to variational problems of the type that can be solved using the calculus of variations.

[2]Fermat was assuming that light slows down as it enters a denser medium; his principle is compatible with a wave theory of light if his light velocity is taken to be the phase velocity rather than the group velocity.

Fermat's work is what led Johann Bernoulli to his solution of the brachistochrone problem, and it is also what led Euler, Maupertuis[3] and D'Alembert to the *principle of least action*, which aimed to do for mechanics what Fermat had done for geometric optics. In their formulation of the principle, which was that used by Lagrange in his *Mechánique Analytique*, energy was assumed to be conserved and the paths considered were those of fixed energy. A more powerful version of the principle of least action, based on a different meaning of "action", was found by Hamilton in the 1830s, and this is the version that we use today. It was dubbed "Hamilton's principle" by Jacobi, who significantly extended Hamilton's ideas. Nobody else took much notice in the 19th century, because variational principles had been tainted in the 18th century by association with dubious theological ideas (such as Leibnitz's suggestion from 1710, parodied by Voltaire in *Candide*, that we live in "the best of all possible worlds"). Hamilton's principle was viewed as just a clever way to arrive at some equations that could, and should, be considered as the better starting point.

That verdict was overturned in the 20th century. This was partly because of Noether's theorem, published in 1918, relating continuous symmetries of the action to conservation laws, and the increasing relevance of continuous symmetries in particle physics theories from the 1960s onward, and partly because Hamilton's principle arises naturally in Feynman's 1948 formulation of QM: all paths are "tried out" and their relative "weights" are determined by Hamilton's action, which is such that the path of least action dominates in the classical limit. This is also how Fermat's principle arises in the geometric optics approximation to Maxwell's wave equations for light propagation.

1.1 Calculus for functions of many variables

Consider a function $f : \mathbb{R}^n \to \mathbb{R}$. In coordinates $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ for $\mathbb{R}^n$,

$$\mathbf{x} \mapsto f(\mathbf{x}) \,.$$

We shall suppose f to be sufficiently smooth; C_2 (twice differentiable) will usually suffice. Stationary points of f are those points in $\mathbb{R}^n$ for which $\boldsymbol{\nabla} f = \mathbf{0}$, i.e.

$$\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n} \right) = (0, 0, \cdots, 0) \ .$$

Expanding f in a Taylor series about a stationary point $\mathbf{x} = \mathbf{a}$, we have

$$f(\mathbf{x}) = f(\mathbf{a}) + \frac{1}{2} \sum_{i,j} (x_i - a_i)(x_j - a_j) H_{ij}(\mathbf{a}) + \mathcal{O}\left(|\mathbf{x} - \mathbf{a}|^3 \right) \ ,$$

where

$$H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j} \,.$$

[3]The man who confirmed that the Earth is squashed, as Newton predicted, and not squeezed, as Descartes predicted, thus settling remaining doubts about Newton's theory of gravity.

Notice that there is no $(\mathbf{x}-\mathbf{a})\cdot\boldsymbol{\nabla} f$ term because $\boldsymbol{\nabla} f$ is zero at a stationary point, so the first non-constant term in the expansion is the one with coefficients H_{ij}. These are the entries of a matrix H called the Hessian matrix[4]. Notice that H is a symmetric matrix ($H_{ij} = H_{ji}$) as a consequence of the symmetry of mixed partial derivatives.

By shifting the origin of coordinates we can arrange to put any given stationary point at the origin, so let's now suppose that $\mathbf{a} = \mathbf{0}$. Then

$$f(\mathbf{x}) - f(\mathbf{0}) = \frac{1}{2} x_i H_{ij} x_j + \mathcal{O}\left(x^3\right) .$$

Here we use the summation convention that repeated indices are summed, and $H_{ij} = H_{ij}(\mathbf{0})$. Now let $x_i = R_{ij} x'_j$ for some rotation matrix R, which we can choose such that the matrix $H' = R^T H R$ is diagonal

$$H'_{ij} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n \end{pmatrix} .$$

The diagonal entries λ_i are the n eigenvalues of the matrix H. They are all real because H is a real symmetric matrix. Neglecting the $\mathcal{O}(x^3)$ terms, we now have

$$f(\mathbf{x}) - f(\mathbf{0}) = \frac{1}{2} \sum_{i=1}^{n} \lambda_i \left(x'_i\right)^2 .$$

If all λ_i are positive, this is a positive definite quadratic form (positive for non-zero $\mathbf{x}'$ and zero only if $\mathbf{x}' = \mathbf{0}$). In this case, the stationary point of f is a local minimum. If all λ_i are negative it is a local maximum. If some are positive and the rest negative, it is a saddle point of f. If some of the λ_i are zero it is a degenerate stationary point and we need to investigate the $\mathcal{O}(x^3)$ terms in order to say more.

The case of $n = 2$ is particularly simple because then

$$\det H = \lambda_1 \lambda_2 , \qquad \mathrm{Tr}\, H = \lambda_1 + \lambda_2 .$$

This leads to the following classification for $n = 2$:

- $\det H > 0$ & $\mathrm{Tr}\, H > 0$, local minimum.
- $\det H > 0$ & $\mathrm{Tr}\, H < 0$, local maximum.
- $\det H < 0$. Saddle point.
- $\det H = 0$. Degenerate stationary point.

[4]Or simply the "Hessian", but that term is also often used for the determinant of the matrix H.

N.B. A local minimum (maximum) may also be a global minimum (maximum), but it may not be. If not then *either* there is another stationary point that is the global minimum (maximum) *or* there is no global minimum (maximum) *or* the global minimum (maximum) is at the boundary of the domain $D(f)$ of f.

Example: Find and classify the stationary points of

$$f(x,y) = x^3 + y^3 - 3xy\,.$$

$\boldsymbol{\nabla} f = (3x^2 - 3y, 3y^2 - 3x)$. This is zero if $x^2 = y$ and $y^2 = x$, which implies $y^4 = y$, so either $y = 0$ (and then $x = 0$) or $y = 1$ (and then $x = 1$) since this is the only real solution of $y^3 = 1$. So we have two stationary points: $(0,0)$ and $(1,1)$. We can determine their properties from inspection of the Hessian matrix

$$H = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix} \quad \Rightarrow \quad \det H = 9(4xy - 1)\,, \quad \mathrm{Tr}\, H = 6(x+y)\,.$$

- $(1,1)$: $\det H = 27 > 0$ & $\mathrm{Tr}\, H = 12 > 0$. Local minimum, with $f = -1$.
- $(0,0)$: $\det H = -9 < 0$. Saddle, with $f = 0$. The eigenvalues and eigenvectors of H at this stationary point are

  $$\lambda_1 = -3\,, \quad e_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}; \qquad \lambda_2 = 3\,, \quad e_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

  Near the saddle point, the function f decreases on the line $y = x$ and increases on the line $y = -x$.

In this case, there is no global minimum or maximum because f either increases or decreases without bound as $x^2 + y^2 \to \infty$.

2. Convex functions

A set $S \subset \mathbb{R}^n$ is convex if, for all $\mathbf{x}$, $\mathbf{y}$ in S,

$$(1-t)\mathbf{x} + t\mathbf{y} \in S\,, \qquad 0 < t < 1\,.$$

That is, all points on the line segment connecting any two points of the set must also be in the set.

The "graph" of a function $f : \mathbb{R}^n \to \mathbb{R}$ is the surface in $\mathbb{R}^{n+1}$ defined by

$$z = f(\mathbf{x})\,,$$

where z is the $(n+1)$th coordinate of $\mathbb{R}^{n+1}$. A "chord" of f is the line segment in $\mathbb{R}^{n+1}$ joining any two points on its graph. The function f is *convex* if:

(i) its domain $D(f) \subset \mathbb{R}^n$ is a convex set.

(ii) its graph lies below, or on, all its chords. Equivalently,

$$f\left((1-t)\mathbf{x}+t\mathbf{y}\right) \leq (1-t)f(\mathbf{x})+tf(\mathbf{y})\,, \qquad 0<t<1\,. \tag{2.1}$$

We need (i) for (ii) to make sense. If $D(f)$ were not convex then there would be points on chords of f with $\mathbb{R}^n$ coordinates for which f is not defined.

Some further definitions:

- A *strictly convex* function is a convex function for which (2.1) holds (for $\mathbf{y} \neq \mathbf{x}$) as a strict inequality:

(ii)′ $$f\left((1-t)\mathbf{x}+t\mathbf{y}\right) < (1-t)f(\mathbf{x})+tf(\mathbf{y})\,, \qquad 0<t<1\,, \quad \mathbf{y}\neq\mathbf{x}\,. \tag{2.2}$$

- A function f is (strictly) concave if $-f$ is (strictly) convex.

Let's look at some $n = 1$ examples.

1. $f(x) = x^2$ for $x \in \mathbb{R}$. This is obviously convex. In fact, it's strictly convex:

   $[(1-t)x+ty]^2-(1-t)x^2-ty^2 = -t(1-t)\,(x-y)^2 < 0$ for $0<t<1$ and $x \neq y$.

2. $f(x) = e^x$ and $f(x) = e^{-x}$ are also both strictly convex for $x \in \mathbb{R}$.

3. $f(x) = |x|$. This convex but not strictly convex (because condition (ii) holds with equality for any x, y that are both positive or both negative).

4. $f(x) = 1/x$. This is obviously convex if we restrict the domain of f to $x > 0$, but not if the domain is $\mathbb{R}^*$ (all non-zero real numbers). That's because the function $f = 1/x$ for $x < 0$ is the same as $f = -1/x$ for $x > 0$, and is therefore concave. But it's also because $\mathbb{R}^*$ is not a convex subset of $\mathbb{R}$.

2.1 First-order conditions

For a function that is once-differentiable, the convexity condition is equivalent to the following "first-order condition":

$$f(\mathbf{y}) \geq f(\mathbf{x}) + (\mathbf{y}-\mathbf{x})\cdot\boldsymbol{\nabla} f(\mathbf{x})\,. \tag{2.3}$$

This states that a convex function lies above all its tangent planes.

Corollary: If f has a stationary point then it is a global minimum. Proof: given $\boldsymbol{\nabla} f(\mathbf{x}_0) = \mathbf{0}$, we have $f(\mathbf{y}) \geq f(\mathbf{x}_0)$ for all $\mathbf{y}$.

Let's first prove, for once-differentiable functions, that convexity implies (2.3). To do this we rewrite the convexity condition (2.1) as

$$h(t) \equiv (1-t)f(\mathbf{x}) + tf(\mathbf{y}) - f\left((1-t)\mathbf{x} + t\mathbf{y}\right) \geq 0 \qquad (0 < t < 1).$$

Notice that

$$h'(0) = -f(\mathbf{x}) + f(\mathbf{y}) + (\mathbf{y} - \mathbf{x}) \cdot \boldsymbol{\nabla} f(\mathbf{x}) \,,$$

and hence that the first-order condition (2.3) is equivalent to $h'(0) \geq 0$. Now, since $h(0) = 0$ it follows from $h(t) \geq 0$ that $[h(t) - h(0)]/t \geq 0$ for $0 < t < 1$, and hence (by taking the $t \to 0$ limit) that $h'(0) \geq 0$.

To prove the converse, i.e. that the first-order condition (2.3) implies convexity, we begin by rewriting (2.3) in terms of the vector variable pairs $(\mathbf{z}, \mathbf{x})$ and $(\mathbf{z}, \mathbf{y})$ instead of $(\mathbf{x}, \mathbf{y})$:

$$f(\mathbf{x}) \geq f(\mathbf{z}) + (\mathbf{x} - \mathbf{z}) \cdot \boldsymbol{\nabla} f(\mathbf{z}) \,,$$
$$f(\mathbf{y}) \geq f(\mathbf{z}) + (\mathbf{y} - \mathbf{z}) \cdot \boldsymbol{\nabla} f(\mathbf{z}) \,.$$

By taking a linear combination of these two equations, we can deduce that

$$(1-t)f(\mathbf{x}) + tf(\mathbf{y}) \geq f(\mathbf{z}) + [(1-t)\mathbf{x} + t\mathbf{y} - \mathbf{z}] \cdot \boldsymbol{\nabla} f(\mathbf{z}) \,.$$

The convexity property (2.1) now follows on choosing $\mathbf{z} = (1-t)\mathbf{x} + t\mathbf{y}$.

2.1.1 An alternative first-order condition

The first-order convexity condition (2.3) can be rewritten as

$$(\mathbf{x} - \mathbf{y}) \cdot \boldsymbol{\nabla} f(\mathbf{x}) \geq f(\mathbf{x}) - f(\mathbf{y}) \,,$$

and hence as

$$(\mathbf{x} - \mathbf{y}) \cdot [\boldsymbol{\nabla} f(\mathbf{x}) - \boldsymbol{\nabla} f(\mathbf{y})] \geq f(\mathbf{x}) - f(\mathbf{y}) - (\mathbf{x} - \mathbf{y}) \cdot \boldsymbol{\nabla} f(\mathbf{y}) \,.$$

The RHS of this equation is non-negative as a consequence of (2.3), so (2.3) implies the new first-order condition

$$(\mathbf{y} - \mathbf{x}) \cdot (\boldsymbol{\nabla} f(\mathbf{y}) - \boldsymbol{\nabla} f(\mathbf{x})) \geq 0 \,. \tag{2.4}$$

To get an idea of what this condition means, consider the $n = 1$ case, for which it becomes

$$(y - x)\left[f'(y) - f'(x)\right] \geq 0 \,.$$

This is equivalent to the statement that $f'(y) \geq f'(x)$ whenever $y > x$, so $f'(x)$ is a monotonically increasing (never decreasing) function.

To establish equivalence of the two first-order conditions we still need to show that (2.4) implies (2.3). To do so, it is convenient to again define

$$\mathbf{z} = (1-t)\mathbf{x} + t\mathbf{y}\,.$$

Observe that

$$f(\mathbf{y}) - f(\mathbf{x}) = [f(\mathbf{z})]_0^1 = \int_0^1 dt \frac{d}{dt} f(\mathbf{z}) = \int_0^1 dt\,(\mathbf{y}-\mathbf{x})\cdot\boldsymbol{\nabla} f(\mathbf{z})\,,$$

and hence that

$$f(\mathbf{y}) - f(\mathbf{x}) - (\mathbf{y}-\mathbf{x})\cdot\boldsymbol{\nabla} f(\mathbf{x}) = \int_0^1 dt\,(\mathbf{y}-\mathbf{x})\cdot[\boldsymbol{\nabla} f(\mathbf{z}) - \boldsymbol{\nabla} f(\mathbf{x})]\ . \tag{2.5}$$

Now, by replacing $\mathbf{y}$ by $\mathbf{z}$ in (2.4), and then dividing by t, we find that

$$(\mathbf{y}-\mathbf{x})\cdot[\boldsymbol{\nabla} f(\mathbf{z}) - \boldsymbol{\nabla} f(\mathbf{x})] \geq 0\,.$$

Using this in (2.5) we deduce (2.3).

2.2 The Hessian and a second-order condition

For a function f that is everywhere twice differentiable, the convexity condition (2.1) is equivalent to positivity of the Hessian matrix. More precisely, the function is convex iff the Hessian matrix never has a negative eigenvalue[5].

The necessity of this "second-order" condition follows directly from the first-order condition (2.4). Choose $\mathbf{y} = \mathbf{x} + \mathbf{h}$ (for constant $\mathbf{h}$) to get

$$\mathbf{h}\cdot[\boldsymbol{\nabla} f(\mathbf{x}+\mathbf{h}) - \boldsymbol{\nabla} f(\mathbf{x})] \geq 0\,.$$

For small $|\mathbf{h}|$ we have

$$\nabla_i f(\mathbf{x}+\mathbf{h}) = \nabla_i f(\mathbf{x}) + h_j H_{ij}(\mathbf{x}) + \mathcal{O}(h^2)\,,$$

where H_{ij} is the Hessian, and hence

$$h_i h_j H_{ij}(\mathbf{x}) + \mathcal{O}(h^3) \geq 0\,.$$

Since we can take $|\mathbf{h}|$ as small as we wish, this can be true for all $\mathbf{h}$ only if the Hessian matrix has no negative eigenvalues, and this condition must be satisfied for all $\mathbf{x} \in D(f)$.

Positivity of the Hessian matrix is also sufficient for convexity of a twice-differentiable function. Consider the $n = 1$ case, for which we have a function $f(x)$ such that

[5] If all eigenvalues are everywhere strictly positive then the function is strictly convex, but this is only a sufficient condition for strict convexity, not a necessary one. For example, the function $f(x) = x^4$ is strictly convex despite the fact that $f''(x)$ is zero at $x = 0$.

$f''(x) \geq 0$ for all x in some convex domain. It follows (for $x+h$ in the domain of f) that

$$0 \leq (\operatorname{sign} h) \int_x^{x+h} f''(z)dz = (\operatorname{sign} h)\left[f'(x+h) - f'(x)\right] .$$

Now integrate over h, from 0 to $y-x$ (which requires $y > x$ if $h > 0$ and $y < x$ if $h < 0$) to get

$$0 \leq \int_0^{y-x} \left[f'(x+h) - f'(x)\right] dh = f(y) - f(x) - (y-x)f'(x) .$$

Equivalently, $f(y) \geq f(x) + (y-x)f'(x)$, which is the first-order condition (2.3) for $n = 1$.

Example: $f(x,y) = (xy)^{-1}$ for $x > 0$ and $y > 0$. The graph of f is the positive quadrant of a hyperboloid. The Hessian is

$$H = \left(\frac{1}{xy}\right) \begin{pmatrix} \frac{2}{x^2} & \frac{1}{xy} \\ \frac{1}{xy} & \frac{2}{y^2} \end{pmatrix}$$

which gives

$$\det H = \frac{3}{x^4 y^4} > 0 \quad \& \quad \operatorname{Tr} H = \frac{2}{xy}\left(\frac{1}{x^2} + \frac{1}{y^2}\right) > 0 ,$$

so both eigenvalues of H are everywhere strictly positive in $D(f)$. The function is therefore strictly convex.

Notice that H is still positive if we take the domain of f to be $xy > 0$ but this domain is not a convex set, so f would *not* be a convex function in this case!

3. Legendre transform

The Legendre transform of a function $f : \mathbb{R}^n \to \mathbb{R}$ is a new function f^* defined by[6]

$$f^*(\mathbf{p}) = \sup_{\mathbf{x}} \left[\mathbf{p}\cdot\mathbf{x} - f(\mathbf{x})\right] .$$

The domain of f^* is the subset of $\mathbb{R}^n$ for which the right hand side is finite; in other words, for which a supremum exists. One immediate consequence of this definition is that f^* is a convex function:

$$\begin{aligned} f^*\left((1-t)\mathbf{p} + t\mathbf{q}\right) &= \sup_{\mathbf{x}} \left[(1-t)\mathbf{p}\cdot\mathbf{x} + t\mathbf{q}\cdot\mathbf{x} - f(\mathbf{x})\right] \\ &= \sup_{\mathbf{x}} \left[(1-t)\left\{\mathbf{p}\cdot\mathbf{x} - f(\mathbf{x})\right\} + t\left\{\mathbf{q}\cdot\mathbf{x} - f(\mathbf{x})\right\}\right] . \end{aligned}$$

[6]The "supremum" of a set of real numbers is its lowest upper bound. It's the same as the maximum when there is a maximum, but some sets have a supremum but no maximum. For example, there is no maximum number in the open interval $(0,1)$ but the supremum is 1. For practical purposes you may substitute "maximum" for "supremum".

But the supremum of the sum of any two functions cannot exceed the sum of their individual suprema, so

$$RHS \leq (1-t)\sup_{\mathbf{x}}\left[\mathbf{p}\cdot\mathbf{x}-f(\mathbf{x})\right]+t\sup_{\mathbf{x}}\left[\mathbf{q}\cdot\mathbf{x}-f(\mathbf{x})\right]=(1-t)f^*(\mathbf{p})+tf^*(\mathbf{q})\,,$$

and hence

$$f^*\left((1-t)\mathbf{p}+t\mathbf{q}\right)\leq 1-t)f^*(\mathbf{p})+tf^*(\mathbf{q})\,.$$

This shows, firstly, that the LHS is finite if $f^*(\mathbf{p})$ and $f^*(\mathbf{q})$ are finite, so that $(1-t)\mathbf{p}+t\mathbf{q}\in D(f^*)$ if $\mathbf{p},\mathbf{q}\in D(f^*)$, which means that $D(f^*)$ is convex. Secondly, it shows that f^* satisfies the convexity condition (2.1). So f^* is convex.

N.B. If $f(\mathbf{x})$ is convex then so is $F_{\mathbf{p}}(\mathbf{x})=f(\mathbf{x})-\mathbf{p}\cdot\mathbf{x}$ for any $\mathbf{p}$. Let's check this for once-differentiable f; in this case

$$F_{\mathbf{p}}(\mathbf{y})-F_{\mathbf{p}}(\mathbf{x})-(\mathbf{y}-\mathbf{x})\cdot\boldsymbol{\nabla}F_{\mathbf{p}}(\mathbf{x})=f(\mathbf{y})-f(\mathbf{x})-(\mathbf{y}-\mathbf{x})\cdot\boldsymbol{\nabla}f(\mathbf{x})\geq 0\,.$$

Corollary: if $f(\mathbf{x})$ is convex and once-differentiable then any stationary point of $\mathbf{p}\cdot\mathbf{x}-f(\mathbf{x})$ is a global maximum, which occurs at a solution $\mathbf{x}(\mathbf{p})$ of

$$\boldsymbol{\nabla}f(\mathbf{x})=\mathbf{p}\,. \tag{3.1}$$

The Legendre transform of f is then

$$f^*(\mathbf{p})=\mathbf{p}\cdot\mathbf{x}(\mathbf{p})-f\left(\mathbf{x}(\mathbf{p})\right)\,.$$

- If f is strictly convex function then the solution $\mathbf{x}(\mathbf{p})$ of (3.1) is unique.

  Consider the $n=1$ case, for which $f^*(p)=px(p)-f\left(x(p)\right)$, where $x(p)$ is the solution of $f'(x)=p$. Since f' is a monotonically increasing function of x, there can be only one value of x for a given value of p, and the solution exists for $p\in D(f)$ by definition of this domain. If f is convex but not strictly convex there will be a solution but it will not be unique for all $p\in D(f^*)$.

Let's now look at some $n=1$ examples:

1. $f=\frac{1}{2}ax^2$ with $a>0$. In this case $p=ax$, and hence $x=p/a$ at the maximum of $px-f(x)$, which is then $f^*(p)$. So

   $$f^*(p)=p(p/a)-\frac{1}{2}a(p/a)^2=\frac{1}{2a}p^2\qquad p\in\mathbb{R}$$

   The curve $z=f(x)$ describes a parabola in $\mathbb{R}^2$, so we have just shown that the Legendre transform of a parabola is another parabola.

2. $f = -\sqrt{1-v^2}$ for $|v| < 1$. The curve $z = f(v)$ is now a unit semi-circle. In this case, $p = \frac{v}{\sqrt{1-v^2}}$ and hence $v = \frac{p}{\sqrt{1+p^2}}$ for $p \in \mathbb{R}$, so

$$f^*(p) = \frac{p^2}{\sqrt{1+p^2}} + \frac{1}{\sqrt{1+p^2}} = \sqrt{1+p^2}\,.$$

The curve $z = f^*(p)$ is the upper branch of a hyperbola, so the Legendre transform of a semi-circle is one branch of a hyperbola. This example is relevant to the unit-mass relativistic point particle of velocity v.

3. $f = cx$. This is convex (although not strictly convex) for $c > 0$; the graph of f is a line in $\mathbb{R}^2$. In this case $px - f(x) = (p-c)x$, which has no maximum with respect to variations of x unless $p = c$, in which case x is undetermined. The domain of f^* is now the one point $p = c$, and $f^*(c) = 0$: The Legendre transform of a line is a point.

Theorem: If f is a convex differentiable function with Legendre transform f^* (which is usually called the "conjugate function") then

$$f^{**} = f \tag{3.2}$$

Proof: Given $f^*(\mathbf{p}) = \mathbf{p}\cdot\mathbf{x} - f(\mathbf{x})$ where $\mathbf{x}$ is the solution of $\boldsymbol{\nabla} f(\mathbf{x}) = \mathbf{p}$, it follows from the chain rule that

$$\begin{aligned}\frac{\partial f^*(\mathbf{p})}{\partial p_i} &= x_i + p_j \frac{\partial x_j(\mathbf{p})}{\partial p_i} - \frac{\partial f(\mathbf{x})}{\partial x_j}\Big|_{\mathbf{x}=\mathbf{x}(\mathbf{p})} \frac{\partial x_j(\mathbf{p})}{\partial p_i} \\ &= x_i + \left[\mathbf{p} - \boldsymbol{\nabla} f(\mathbf{x})\Big|_{\mathbf{x}=\mathbf{x}(\mathbf{p})}\right] \cdot \frac{\partial \mathbf{x}(\mathbf{p})}{\partial p_i} = x_i\,,\end{aligned} \tag{3.3}$$

where the last equality uses the fact that $\mathbf{x}(\mathbf{p})$ solves $\boldsymbol{\nabla} f = \mathbf{p}$. We have now shown that the function $\mathbf{p}(\mathbf{x})$ inverse to $\mathbf{x}(\mathbf{p})$, satisfies

$$\boldsymbol{\nabla} f^*(\mathbf{p}) = \mathbf{x}\,,$$

where, here, $\boldsymbol{\nabla}$ is a derivative with respect to $\mathbf{p}$.

Now we take the Legendre transform of f^*:

$$\begin{aligned} f^{**}(\mathbf{x}) &= \mathbf{x}\cdot\mathbf{p}(\mathbf{x}) - f^*(\mathbf{p}(\mathbf{x})) \\ &= \mathbf{x}\cdot\mathbf{p}(\mathbf{x}) - [\mathbf{p}(\mathbf{x})\cdot\mathbf{x} - f(\mathbf{x}(\mathbf{p}(\mathbf{x})))] \\ &= f(\mathbf{x}(\mathbf{p}(\mathbf{x}))) = f(\mathbf{x})\,,\end{aligned}$$

since $\mathbf{x}(\mathbf{p}(\mathbf{x})) = \mathbf{x}$.

N.B. We know that the Legendre transform of any function is a convex function, so if we know that f is the Legendre transform of f^* then we know that it is convex.

This shows that the convexity condition on f is necessary for the validity of the theorem. However, strict convexity is not needed. For example, we saw that the function $f(x) = cx$ has Legendre transform $f^*(p) = 0$ with a single point $p = c$ as its domain. Take another Legendre transform:

$$f^{**}(x) = [xp - f^*(p)]_{p=c} = cx = f(x)\,.$$

So the Legendre transform exchanges a point with a line.

3.1 Application to Thermodynamics

The first law of thermodynamics is

$$dE = TdS - PdV\,, \tag{3.4}$$

which states that a small change in the energy of a system in thermal equilibrium at temperature[7] T and pressure P is the sum of a heat energy term (TdS), due to a change in the entropy S, and a mechanical work energy term ($-PdV$), due to a change in the volume V. The formula also shows that the total energy E is a function of the two "extensive" variables (S, V), so called because these variables scale with the size of the system. This is in contrast to the "intensive" variables (T, P), which can be defined as[8]

$$T = \left(\frac{\partial E}{\partial S}\right)_V\,, \qquad P = -\left(\frac{\partial E}{\partial V}\right)_S \tag{3.5}$$

The pairs (T, S) and $(-P, V)$ are said to be conjugate pairs of thermodynamic variables. In more general forms of the first law there can be more pairs of conjugate variables, e.g. chemical potential μ and particle number N.

For a process occurring at fixed entropy the first law becomes $dE + PdV = 0$, which tells us that work done by the system will lead to a corresponding reduction of its energy E. However, many processes of interest occur at fixed temperature, not at fixed entropy, and in such cases it is more useful to consider (T, V) as the independent variables. We can arrange for this by taking the Legendre transform of $E(S, V)$ with respect to S (the volume variable V just goes along for the ride here). We will call this new function $-F(T, V)$, so

$$-F(T, V) = \sup_S \left[TS - E(S, V)\right]\,. \tag{3.6}$$

Strictly speaking, we do not yet know that the new independent variable T is the temperature appearing in the first law. However, the maximum of the RHS w.r.t. variations of S occurs when $T = (\partial E/\partial S)_V$, and this is indeed the temperature as

[7] This is absolute temperature, i.e. zero at absolute zero.

[8] The variable held fixed is indicated explicitly because otherwise it can get confusing.

defined in (3.5). Solving $T = (\partial E/\partial S)_V$ for $S = S(T, V)$ then gives F as a function of T and V. It then follows that

$$dF = \left(\frac{\partial F}{\partial T}\right)_V dT + \left(\frac{\partial F}{\partial V}\right)_T dV \,.$$

On the other hand, it follows from (3.6) that

$$dF = -SdT - TdS + dE = -SdT - PdV \,,$$

where the second equality follows from the first law. This confirms that $F = F(T, V)$ and it tells us that

$$S = -\left(\frac{\partial F}{\partial T}\right)_V \,, \qquad P = -\left(\frac{\partial F}{\partial V}\right)_T \,.$$

We now have an alternative version of the first law:

$$dF = -SdT - PdV \,.$$

For a process at fixed T, this reduces to $dF + PdV = 0$, which tells us that work done by the system at fixed T implies a corresponding reduction in F, which is therefore the energy that is *available* to do work at fixed temperature; this is *less than the total energy* E because $F = E - TS$ and both T and S are positive. This "available energy", as it is sometimes called, is more usually called the *Helmholtz free energy*, or just "free energy".

It is also possible to take the Legendre transform of $E(S, V)$ with respect to the volume V. This gives a new function

$$-H(S, P) \equiv E^*(S, -P) = \sup_V \left[(-P)V - E\right] \,.$$

Varying with respect to V yields $-P = (\partial E/\partial V)_S$, in agreement with the definition of P in (3.5), and this explains why we take the new variable for the Legendre transform to be $-P$. Solving this equation for V in terms of P (now S goes along for the ride) we indeed get a function of S and P, which implies that

$$dH = \left(\frac{\partial H}{\partial S}\right)_P dS + \left(\frac{\partial H}{\partial P}\right)_S dP \,.$$

On the other hand, we have

$$dH = VdP + PdV + dE = TdS + VdP \,,$$

where the second equality follows from the first law. This tells us that

$$T = \left(\frac{\partial H}{\partial S}\right)_P \,, \qquad V = \left(\frac{\partial H}{\partial P}\right)_S \,.$$

The function $H(S,P)$ is called the "enthalpy". It is useful for chemistry because chemical reactions often take place at fixed (e.g. atmospheric) pressure P, in which case $dH = TdS$, which tells that a transfer of heat to a substance raises its enthalpy by a corresponding amount.

The functions E, F, H are called "thermodynamic potentials". There is one more of them: the Gibbs free energy $G(T,P)$. It is found either by taking the Legendre transform of $F(T,V)$ with respect to V or, equivalently, by taking the Legendre transform of $H(S,P)$ with respect to S. In the latter case

$$G(T,P) = H(S,P) - TS\,,$$

where S solves $(\partial H/\partial S)_P = T$. Now we find that

$$dG = [VdP + TdS] - TdS - SdT = VdP - SdT\,,$$

and hence

$$S = -\left(\frac{\partial G}{\partial P}\right)_T\,, \qquad V = \left(\frac{\partial G}{\partial T}\right)_P\,.$$

Maxwell Relations: e.g. from $F(T,V)$. We have

$$\frac{\partial^2 F}{\partial T \partial V} = \frac{\partial}{\partial T}\Big|_V \left(\frac{\partial F}{\partial V}\right)_T = -\left(\frac{\partial P}{\partial T}\right)_V\,,$$

but we also have

$$\frac{\partial^2 F}{\partial V \partial T} = \frac{\partial}{\partial V}\Big|_T \left(\frac{\partial F}{\partial T}\right)_V = -\left(\frac{\partial S}{\partial V}\right)_T\,.$$

Using the symmetry of mixed partial derivatives, we deduce that

$$\left(\frac{\partial P}{\partial T}\right)_V = \left(\frac{\partial S}{\partial V}\right)_T\,.$$

Starting from the other three thermodynamic potentials we can derive three more such "Maxwell relations".

4. Constrained variation and Lagrange multipliers

Let $f(x,y)$ be height above the ground (the x-y plane). At the highest (hilltop) point a small change in position does not change the altitude so

$$0 = df = \boldsymbol{\nabla} f \cdot \mathbf{dl}\,.$$

In other words, at the hilltop $\boldsymbol{\nabla} f$ is orthogonal to all possible displacement vectors $\mathbf{dl}$, and hence zero there, which means that we have to solve the two equations $\boldsymbol{\nabla} f = \mathbf{0}$ for (x,y) to find the position of the hilltop.

Now suppose that we have a path threading through this hilly landscape, specified by the constraint $p(x, y) = 0$, and suppose that we want to find the highest point on this path. We still need $\boldsymbol{\nabla} f \cdot \mathbf{dl} = 0$ but $\mathbf{dl}$ is no longer arbitrary; it is restricted by the fact that to remain on the path we must have

$$0 = dp = \boldsymbol{\nabla} p \cdot \mathbf{dl} \,.$$

In other words, we need

$$\boldsymbol{\nabla} f \cdot \mathbf{dl}_\perp = 0 \,, \tag{4.1}$$

where $\mathbf{dl}_\perp$ is orthogonal to $\boldsymbol{\nabla} p$. This implies not that $\boldsymbol{\nabla} f = 0$ but instead that $(\boldsymbol{\nabla} f)_\perp = 0$; we learn nothing about the projection of $\boldsymbol{\nabla} f$ on to the direction defined by $\boldsymbol{\nabla} p$. At the hilltop we therefore have

$$\boldsymbol{\nabla} f - \lambda \boldsymbol{\nabla} p = 0 \,, \qquad p = 0 \,, \tag{4.2}$$

where λ is some unknown constant, corresponding to the unknown magnitude of the component of $\boldsymbol{\nabla} f$ parallel to ∇p. We now have an additional equation to solve ($p = 0$) but also an additional variable (λ). These equations are those that give the stationary points, *without constraint*, of the function of three variables

$$\phi(x, y, \lambda) = f(x, y) - \lambda p(x, y) \,.$$

Variation with respect to λ, which is called a *Lagrange multiplier* in this context, gives us the constraint $p = 0$ of the original variational problem. Variation with respect to (x, y) gives us the other equations of (4.2).

By means of a Lagrange multiplier we have therefore turned a constrained variational problem into an unconstrained variational problem. To do so we have had to weaken the problem from finding a maximum to finding a stationary point, so we have to determine by other means which stationary point, if any, is the one we need. However, this is usually easy to sort out, and a bonus is that the value of the Lagrange multiplier λ often has some significance that aids understanding of the problem.

Example 1: An open shoe box has sides of length (x, y) and height z. Given that its volume is $L^3/2$ for fixed L, find the dimensions (x, y, z) of the box that minimise its surface area.

The surface area A and volume V are

$$A = 2z(x + y) + xy \,, \qquad V = xyz \,.$$

We therefore have to minimise $A(x, y, z)$ subject to the constraint that $xyz = L^3/2$. This problem can be solved directly, by solving the constraint, or indirectly, using the Lagrange multiplier method. Let's do it both ways:

- **Direct method: solve the constraint**. We can do this by expressing one of the three variables (x, y, z) in terms of the other two. The height dimension z is clearly special because the box has a bottom but not a top, which suggests that we 'solve' the constraint for z by rewriting it as $z = L^3/(2xy)$. Substituting into $A(x, y, z)$ we then get the area as a function of (x, y) alone:

$$A(x, y) = A\left(x, y, z\left(x, y\right)\right) = \frac{L^3}{x} + \frac{L^3}{y} + xy\,.$$

We now have an unconstrained variational problem. The function $A(x, y)$ is stationary when

$$0 = \frac{\partial A}{\partial x} = -\frac{L^3}{x^2} + y\,, \qquad 0 = \frac{\partial A}{\partial y} = -\frac{L^3}{y^2} + x\,,$$

or $x^2y = xy^2 = L^3$. These equations tell us that $x = y = L$, which gives us $z = L/2$, so there is one stationary point at

$$(x, y, z) = (L, L, L/2)\,.$$

The stationary point is obviously a minimum, and this can be checked by computing the Hessian matrix of A at the stationary point: Its eigenvalues are 1 and 3, both positive, as required for a minimum.

At this minimum, both A and V are particular functions of the length L:

$$A(L) = 3L^2\,, \quad V(L) = L^3/2\,, \qquad \Rightarrow \quad dA = 6L dL\,, \quad dV = \frac{3}{2}L^2 dL\,,$$

so the minimum area changes when we change the fixed volume according to the the formula

$$dA_{min}/dV_{fixed} = 4/L\,.$$

- **Lagrange multiplier method**. In this case we need to find the stationary points, without constraint, of the function

$$\phi(x, y, z, \lambda) = A(x, y, z) - \lambda\left(xyz - L^3/2\right) = 2z(x + y) + xy - \lambda xyz + \lambda L^3/2\,.$$

From the variation respect to z we get

$$0 = 2(x + y) - \lambda xy \quad \Rightarrow\ \lambda = 2\frac{(x + y)}{xy}\,.$$

From the variation with respect to x and y (and then using the above result for λ) we get

$$0 = 2z + y - \lambda yz = \frac{y}{x}\left(x - 2z\right) \quad \Rightarrow\ x = 2z$$
$$0 = 2z + x - \lambda xz = \frac{x}{y}\left(y - 2z\right) \quad \Rightarrow\ y = 2z$$

We thus learn that $x = y = 2z$. Finally, variation with respect to λ yields the constraint; in terms of z this is now $z = L/2$, so the stationary point is at

$$(x, y, z) = (L, L, L/2)\,, \qquad \lambda = 4/L \equiv dA_{min}/dV_{fixed}\,.$$

We get the same stationary point as before, and the value of the Lagrange multiplier at this point tells us something else about the nature of the solution to the problem.

A disadvantage of the Lagrange multiplier method is that the the stationary point is always a saddle point of ϕ. This may be verified for the above example by computing the determinant of the 4×4 Hessian matrix of ϕ at the stationary point; it is negative, which implies that the Hessian matrix has a odd number of negative eigenvalues. The main advantage of the method is that it can be used when the direct method cannot be used because the constraint is too complicated to allow an *explicit* solution.

Even when the constraint *can* be solved explicitly, it might not be convenient to do so, as the next two examples illustrate.

Example 2: *For* $\mathbf{x} \in \mathbb{R}^n$*, find the minumum of the quadratic form* $f(\mathbf{x}) = x_i A_{ij} x_j$ *on the surface* $|\mathbf{x}|^2 = 1$.

We *could* solve the constraint; e.g. $x_n = \sqrt{1 - x_1^2 - \ldots - x_{n-1}^2}$, but this solution arbitrarily picks out x_n as special, and it also introduces non-linearities that are not intrinsic to the problem. It is simpler to use the Lagrange multiplier method, according to which we have to find the stationary values, without constraint, of the function

$$\phi\,(\mathbf{x}, \lambda) = x_i A_{ij} x_j - \lambda\left(|\mathbf{x}|^2 - 1\right)\,.$$

The stationary points of this function are found by solving

$$A_{ij} x_j = \lambda x_i\,, \qquad |\mathbf{x}|^2 = 1\,,$$

which tells us that the stationary points are normalized eigenvectors of the symmetric matrix A with entries A_{ij}. The eigenvalues are the possible solutions for the Lagrange multiplier λ. Furthermore, at a stationary point we have

$$f(\mathbf{x}) \equiv x_i A_{ij} x_j = \lambda x_i x_i = \lambda\,,$$

so the eigenvalues of A are the values of f at its stationary points. Assuming that all eigenvalues are positive, so that f has a minimum, its absolute minimum will be the value of the lowest eigenvalue, i.e. the least possible value of the Lagrange multiplier.

- **Alternative solution**. Another way to solve this problem is based on the observation that the constraint fixes only the length of the vector $\mathbf{x} \in \mathbb{R}^n$, not its direction. If we had been asked to minimise the ratio

$$\Lambda(\mathbf{x}) = f(\mathbf{x})/g(\mathbf{x})\,, \qquad \left(g \equiv |\mathbf{x}|^2\right)\,,$$

then the constraint would have been irrelevant because $\Lambda(\mathbf{x})$ does not depend on the length of $\mathbf{x}$; it is sensitive only to changes in the *direction* of $\mathbf{x}$. On the other hand, $\Lambda(\mathbf{x}) = f(\mathbf{x})$ when the constraint is satisfied. It follows that the original problem is equivalent to the problem of finding the minimum of Λ with respect to *unconstrained* variations of $\mathbf{x}$.

Let's verify this claim; we will again that assume that A is a positive matrix (no negative eigenvalues). The stationary points of Λ are given by

$$0 = \frac{\partial \Lambda}{\partial x_i} = \frac{1}{g}\left[\nabla_i f - (f/g)\nabla_i g\right] = \frac{2}{g}\left[A_{ij}x_j - \Lambda x_i\right]\,.$$

So we are back to the eigenvalue problem. The stationary points of Λ are eigenvectors of A, but their length is now undetermined. The values of the function Λ at these stationary points are the eigenvalues of A, so the absolute minimum of Λ equals the lowest eigenvalue. Finally, because this result is independent of the length of $\mathbf{x}$, we are free to choose $|\mathbf{x}| = 1$, in which case $\Lambda = f$ and we have the solution to the original problem.

Example 3: *What probability distribution* $\{p_1, \ldots, p_n\}$*, satisfying* $\sum_{i=1}^n p_i = 1$*, maximises the information entropy* $S = -\sum_{i=1}^n p_i \log_2 p_i$*?*

We can solve this problem by finding the stationary points without constraint of the function

$$\phi(p_1, \ldots, p_n; \lambda) = S - \lambda\left(\sum_{i=1}^n p_i - 1\right) = \sum_i \left[-p_i \log_2 p_i - \lambda p_i\right] + \lambda\,.$$

The stationary points are solutions of (use $\log_2 p = \ln p / \ln 2$)

$$\log_2 p_i + \left(\frac{1}{\ln 2} + \lambda\right) = 0 \qquad i = 1, \ldots, n\,, \qquad \sum_{i=1}^n p_i = 1\,.$$

The first equation tells us that all p_i are equal (to some function of λ) and to satisfy the constraint we require

$$p_i = \frac{1}{n} \qquad i = 1, \ldots, n \quad \Rightarrow \quad S_{\max} = \log_2 n\,.$$

Multiple constraints: The method of Lagrange multipliers is easily extended to find the stationary points of $f : \mathbb{R}^n \to \mathbb{R}$ subject to $m < n$ constraints $p_k(\mathbf{x}) = 0$ $(k = 1, \ldots, m)$. In this case we need m Lagrange multipliers, one for each constraint, and we have to extremise the function

$$\phi(\mathbf{x}; \lambda_1, \ldots, \lambda_m) = f(\mathbf{x}) - \sum_{k=1}^{m} \lambda_k \, p_k(\mathbf{x})$$

with respect to the $n + m$ variables on which it depends.

5. Functionals and the Euler-Lagrange equation

The concept of a function $f : \mathbb{R}^n \to \mathbb{R}$ still makes sense in the $n \to \infty$ limit. In that case we have a function of an infinite number of variables:

$$f(\mathbf{x}) \in \mathbb{R}\,, \qquad \mathbf{x} = \{x_i,\ i \in \mathbb{N}^+\}\,.$$

Here i is a label for a discrete infinity of variables. We can also have a continuous infinity of them, e.g. $\{x(s); s \in \mathbb{R}\}$. In this case we use the notation $F[x]$ and call it a "functional" of the function $x(s)$:

$$F[x] \in \mathbb{R}\,, \quad x = \{x(s),\ s \in \mathbb{R}\}\,.$$

The functional F depends on the function $x(s)$ but it doesn't depend on the independent variable s; that's just a label, analogous to the discrete label i for the variables $\mathbf{x}$ of the function $f(\mathbf{x})$ (we don't get a new function for each i).

Just as we can have functions of many variables, so we can have functionals of many functions:

$$F[\mathbf{x}] \in \mathbb{R}\,, \qquad \mathbf{x} = \{\mathbf{x}(s) \in \mathbb{R}^n,\ s \in \mathbb{R}\}\,.$$

We can also have functionals of a function of many variables:

$$F[x] \in \mathbb{R}\,, \qquad x = \{x(\mathbf{s})\,,\ \mathbf{s} \in \mathbb{R}^n\}\,.$$

And, you guessed it, we can have functionals of many functions of many variables.

Let's start with a functional $F[y]$ of a single function $y(x)$ defined for $\alpha \le x \le \beta$. We will assume that all functions are infinitely differentiable; i.e. "smooth". For many important cases

$$F[y] = \int_\alpha^\beta f(y, y', x)\, dx \qquad (y' = dy/dx).$$

This is a definite integral over x in the interval from $x = \alpha$ to $x = \beta > \alpha$, so it doesn't depend on x, which is just an integration variable, but the value we get

for the integral (a real number) does depend on the function y. In this example, the integrand is a function of y and y'. We could consider integrands that are also functions of y'' or of yet higher derivatives of y, but the most important case is the one we are considering, for which the integrand is restricted to be a function of y and y' only. Notice that the integrand may also have an explicit dependence on the integration variable x, in addition to its implicit x-dependence through y and y'.

Now we investigate how $F[y]$ changes when we change $y(x)$ to a "nearby" function $y(x) + \delta y(x)$. The change in $F[y]$ will be

$$\begin{aligned} F\left[y+\delta y\right] - F[y] &= \int_\alpha^\beta f\left(y+\delta y, y'+(\delta y)'; x\right) dx - \int_\alpha^\beta f(y, y'; x)\, dx \\ &= \int_\alpha^\beta \left\{ \delta y \frac{\partial f}{\partial y} + (\delta y)' \frac{\partial f}{\partial y'} \right\} dx \ + \ldots \end{aligned} \tag{5.1}$$

where the omitted terms are second-order small. Calling the first-order variation $\delta F[y]$, and integrating by parts, we have

$$\delta F[y] = \int_\alpha^\beta \left\{ \delta y \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'} \right) \right] \right\} dx + \left[\delta y \frac{\partial f}{\partial y'} \right]_\alpha^\beta . \tag{5.2}$$

We would like the boundary term to vanish, so we impose boundary conditions on the function y such that this is the case. There are three possibilities:

- Fixed end boundary conditions. We specify the values of $y(\alpha)$ and $y(\beta)$. Then $\delta y(\alpha) = \delta y(\beta) = 0$.
- Free end (or "natural") boundary conditions. These are such that $\partial f/\partial y'$ is zero at the integration endpoints. Usually, this will be the case if we set $y'(\alpha) = y'(\beta) = 0$.
- Mixed boundary conditions. Fixed at one end and free at the other.

However we choose the boundary conditions, if they are such that the boundary term is zero then we can write δF in the form[9]

$$\delta F = \int_\alpha^\beta \left\{ \delta y(x) \frac{\delta F[y]}{\delta y(x)} \right\} dx$$

where $\delta F[y]/\delta y(x)$ is called the "functional derivative" of $F[y]$ with respect to $y(x)$; in this case,

$$\frac{\delta F[y]}{\delta y(x)} = \frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'} \right) .$$

[9]Compare this with the variation of function $f(\mathbf{x})$ of many variables: $\delta f(\mathbf{x}) = \sum_i \delta x_i \partial f/\partial x_i$. The sum gets replaced by an integral for the variation of a functional.

Notice that this is a function of x. The functional F is stationary when its functional derivative is zero (assuming that the b.c.s are such that this derivative is defined) and the condition for this to be true for functionals of the form assumed here is the **Euler-Lagrange equation**

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0\,, \qquad \alpha \le x \le \beta\,.$$

This has an immediate generalisation to functionals of many functions $\mathbf{y}(x) \in \mathbb{R}^n$ for each x in the interval $[\alpha, \beta]$. The starting point is now the functional

$$F[\mathbf{y}] = \int_\alpha^\beta f(\mathbf{y}, \mathbf{y}'; x)\, dx\,.$$

Its variation is

$$\delta F[\mathbf{y}] = \int_\alpha^\beta \left\{ \sum_{i=1}^n \delta y_i \left[\frac{\partial f}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_i'}\right)\right]\right\} dx + \left[\delta y_i \frac{\partial f}{\partial y_i'}\right]_\alpha^\beta\,.$$

For boundary conditions that remove the boundary term[10], the functional is stationary for solutions of the multiple Euler-Lagrange equations

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f}{\partial y_i'}\right) = 0 \qquad i = 1, \ldots, n \quad \alpha \le x \le \beta\,.$$

5.1 Geodesics of the Euclidean plane

What is the curve of least length between two points on the Euclidean plane? No marks for guessing the right answer! Let's pretend we don't know it. The distance between points A and B on a curve $\mathcal{C}$ joining them is

$$L = \int_{\mathcal{C}} dl\,, \qquad dl = \sqrt{dx^2 + dy^2}\,.$$

To express L as a functional we need to decide how to parametrize the path. There are two standard options:

1. Use the x-coordinate (or the y-coordinate) as a parameter on the curve $\mathcal{C}$. Given that $x = \alpha$ at point A and $x = \beta$ at point B, the length of the curve is

   $$L[y] = \int_\alpha^\beta \sqrt{1 + (y')^2}\; dx\,.$$

   This is now a functional of the function $y(x)$ that determines the curve $\mathcal{C}$. We cannot consider all possible curves this way because x will not be monotonically increasing on a curve that "doubles back" on itself, and neither will it uniquely

[10] There are now more possibilities.

specify a point on such a curve, but we can still seek the minimal length curve within the allowed class for which x is a good parameter. In this case $f = \sqrt{1 + (y')^2}$, so $\partial f/\partial y = 0$ and the EL equation can be immediately once-integrated to give the "first integral"

$$\frac{y'}{\sqrt{1 + (y')^2}} = \text{constant}\,.$$

This implies that y' is a constant, and hence that

$$y = mx + b\,,$$

for constants (m, x). This is a straight line. The constants (a, b) are fixed by the boundary conditions.

2. We can use an *arbitrary* monotonically increasing parameter t such that $t = 0$ at point A and $t = 1$ at point B. The path is then specified by giving the two functions $(x(t), y(t))$, which we assume to be twice differentiable (this is now the *only* restriction). We can now write the length as

$$L[\mathbf{x}] = \int_0^1 \sqrt{|\dot{\mathbf{x}}|^2}\, dt\,, \qquad \dot{\mathbf{x}}(t) = \left(\frac{dx}{dt}, \frac{dy}{dt}\right)\,.$$

The boundary conditions fix $\mathbf{x}(t)$ at the endpoints, so the functional L is stationary for solutions of the Euler-Lagrange equation. As the integrand of L depends on $x(t)$ and $y(t)$ only through their first derivatives, the Euler Lagrange equations can be once integrated immediately to give the equations

$$\frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}} = c\,, \qquad \frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}} = s\,, \tag{5.3}$$

for constants (c, s). Squaring these equations, and then adding them, we find that $c^2 + s^2 = 1$, so we may write

$$c = \cos\theta\,, \qquad s = \sin\theta\,.$$

The equations (5.3) also imply that $\dot{y}/\dot{x} = s/c$, and hence that $dy/dx = \tan\theta$, which has the solution

$$y - y_0 = (\tan\theta)x$$

for constant y_0. The path is a straight line, with slope $\tan\theta$. The constants (y_0, θ) are fixed by requiring that this line pass through the points A and B.

6. First integrals and Fermat's principle

As the above example illustrates, the Euler-Lagrange equation for a functional

$$F[y] = \int_\alpha^\beta f(y, y', x)\, dx$$

can be trivially once-integrated to give a "first integral" whenever the integrand of $F[y]$ depends on y only through its derivatives; in other words, when $\partial f/\partial y = 0$. It is also possible to find a first integral when the integrand is special in other ways.

In general, the function f depends both *implicitly* on x through its dependence on y and y' (both functions of x) and *explicitly* on x. By the chain rule, the total derivative of f with respect to x is therefore

$$\frac{df}{dx} = \frac{\partial f}{\partial x} + y'\frac{\partial f}{\partial y} + y''\frac{\partial f}{\partial y'}\,.$$

We can rewrite this

$$\frac{df}{dx} = \frac{\partial f}{\partial x} + y'\left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right] + \frac{d}{dx}\left(y'\frac{\partial f}{\partial y'}\right)\,,$$

which is equivalent to

$$\frac{d}{dx}\left(f - y'\frac{\partial f}{\partial y'}\right) = \frac{\partial f}{\partial x} + y'\left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right]\,.$$

If we now use the Euler-Lagrange equations (on the assumption of appropriate boundary conditions) then

$$\frac{d}{dx}\left(f - y'\frac{\partial f}{\partial y'}\right) = \frac{\partial f}{\partial x}$$

We deduce from this that when f has no explicit dependence on x, i.e. $\partial f/\partial x = 0$, then the EL equations imply that

$$f - y'\frac{\partial f}{\partial y'} = \text{constant}\,. \tag{6.1}$$

In other words, this is a first-integral of the EL equations. This enables us to solve easily a number of important variational problems.

Example 1: *Use Fermat's principle to find the path of a light ray in the vertical $x-z$ plane inside a medium with a refractive index $n(z) = \sqrt{a - bz}$, where (a, b) are positive constants and z is height above the x axis.*

Recall that Fermat's principle states that light takes the path of least time, and this is on the assumption that the speed of light in a medium of refractive index n is c/n. The time T taken to go from point A to point B on a given path is therefore

$c^{-1}\int_A^B n(z)dl$, where the integral is along the path. We have to minimise this time. Equivalently, we have to minimise the "optical path length"

$$\mathcal{P} = cT = \int_A^B n(z)dl\,.$$

Supposing that $x = \alpha$ at A and $x = \beta$ at B, and that x is a good parameter for the ray, the optical path length is

$$\mathcal{P}[z] = \int_\alpha^\beta n(z)\sqrt{1+(z')^2}\,dx\,.$$

Notice that

$$f = n(z)\sqrt{1+(z')^2} \quad \Rightarrow \quad \frac{\partial f}{\partial x} = 0\,,$$

so we have the first integral

$$k = f - z'\frac{\partial f}{\partial z'} = \frac{n(z)}{\sqrt{1+(z')^2}} = \sqrt{\frac{a-bz}{1+(z')^2}}\,,$$

for some constant k. Squaring, we deduce that

$$(z')^2 = \left(b/k^2\right)(z_0 - z)\;, \qquad z_0 = \frac{a-k^2}{b}\,.$$

Taking the square root, we deduce that

$$\frac{d}{dx}\left[\sqrt{z_0 - z} \pm \frac{\sqrt{b}}{2k}\,x\right] \quad \Rightarrow \quad z = z_0 - \frac{b}{4k^2}\,(x-x_0)^2\;,$$

where x_0 is another integration constant. This is a parabola. At $x = x_0$ the ray reaches a maximum height $z = z_0$.

Does this result remind you of something? The motion of a projectile subject to the downward acceleration g due to gravity near the Earth's surface? Inspired by Fermat's work in optics, Maupertuis suggested that mechanics could be similarly based on a "principle of least action", where "action" should be the product of mass, velocity and distance (which means that it has dimensions of angular momentum). He was vague about the details, but Euler had already discovered that the motion of a body of constant total energy

$$E = \frac{1}{2}mv^2 + U(\mathbf{x}) \qquad (v = |\dot{\mathbf{x}}|)$$

would minimise the integral $A = m\int v\,dl$. Solving the above equation for v, this means that we should minimise

$$A = \int_A^B \sqrt{2m(E - U(\mathbf{x})}\,dl\,.$$

For the motion of a projectile near the surface of the Earth, we should take $U = mgz$, and $dl = \sqrt{dx^2 + dz^2}$, so we have to minimise

$$A = \int_A^B \sqrt{a - bz}\, dl\,, \qquad a = 2mE\,, \qquad b = 2m^2 g\,.$$

This is the same problem as the geometric optics problem just posed, and solved using Fermat's principle!

Example 2: The brachistochrone. *A bead slides on a frictionless wire in a vertical plane. What shape of the wire minimises the time for the bead to fall from rest at point A to a lower, and horizontally displaced, point B?*

Choose A to be the origin of coordinates in the vertical plane with x being horizontal distance from the origin and y being the distance below the origin. The bead starts with zero velocity so conservation of energy implies that its speed v at any later time is given by

$$\frac{1}{2}mv^2 = mgy \quad \Rightarrow \quad v = \sqrt{2gy}\,.$$

In other words, we have to find the path that minimises the travel time when the speed depends on position, *exactly like the optics problems to which Fermat's principle applies*. Specifically, we have to minimise

$$T = \int_A^B \frac{dl}{v} = \frac{1}{\sqrt{2g}} \int_A^B \frac{\sqrt{dx^2 + dy^2}}{\sqrt{y}}\,.$$

For simplicity, assume that x is a good coordinate on the curve, so that

$$T[y] \propto \int_0^{x_B} \sqrt{\frac{1 + (y')^2}{y}}\, dx\,, \quad \Rightarrow \quad f = \sqrt{\frac{1 + (y')^2}{y}}\,.$$

As f has no explicit x-dependence, we have the first integral

$$\text{constant} = f - y' \frac{\partial f}{\partial y'} = \frac{1}{\sqrt{y[1 + (y')^2}} \quad \Rightarrow \quad y\left[1 + (y')^2\right] = 2c\,,$$

for positive constant c. The solution of this first-order ODE with $y(0) = 0$ is given parametrically by

$$x = c\,(\theta - \sin\theta)\,, \qquad y = c\,(1 - \cos\theta)\,,$$

which is an inverted cycloid. The origin (point A) corresponds to $\theta = 0$. Requiring that the curve pass through (x_B, y_B) fixes both c and the value of θ at point B.

A cycloid is "the curve traced by a point on the rim of a circular wheel as the wheel rolls along a straight line without slippage" (Wikipedia). The cycloid was studied and named by Galileo, but Johann Bernouilli is credited with the discovery, published in 1697, that it is a Brachistochrone. Huygens had earlier shown, in 1673, that it is a Tautochrone (the curve such that the time taken for the bead to fall from rest to B is independent of the choice of A).

7. Constrained variation of functionals

The method of Lagrange multipliers can be used to solve variational problems with constraints when we are faced with finding the stationary values of some functional subject to some other functional constraint. For example, if we want to find the stationary points of $F[y]$ subject to the constraint $P[y] = c$, for some constant c, we may extremize, *without constraint*,

$$\Phi_\lambda[y] = F[y] - \lambda\,(P[y] - c)$$

with respect to both the function y and the variable λ. Assuming that the boundary term in the variation is zero, this yields the equations

$$\frac{\delta F}{\delta y(x)} - \lambda \frac{\delta P}{\delta y(x)} = 0\,, \qquad P[y] = c\,.$$

A well-known example is the problem of the curve assumed by a chain of fixed length hanging under its own weight; the curve of minimal energy is a *catenary* (see Q.I.13). Here we'll consider a problem related to Q.I.12.

Isoperimetric problem. *What simple closed plane curve of fixed length L maximizes the enclosed area A?*

The adjective "simple" means that the curve cannot cross itself (excludes a figure of eight) and that the region it encloses is simply connected (excludes a curve that bounds several disjoint regions). As the problem is posed, the inside region need not be convex but it is obvious that it must be to maximise the area, so we'll assume that the curve bounds a convex region in the plane.

As we move around such a curve, the x coordinate will increase monotonically from a minimum value $x = \alpha$ to a maximum value $x = \beta > \alpha$ and then decrease back to its minimum value. If we go around the curve in a clockwise sense, the semi-curve of increasing x is its upper part and the semi-curve of decreasing x is its lower part. So each value of the x coordinate in the interval (α, β) corresponds to two values of y; call them y_1 and $y_2 > y_1$. We can now write the area of the enclosed region as an integral over x of area elements of vertical strips of width dx and height $y_2(x) - y_1(x)$:

$$dA = [y(x)]_{x1}^{x_2}\,dx\,.$$

The total area is therefore

$$A[y] = \int_\alpha^\beta [y_2(x) - y_1(x)]\,dx = \oint_C y(x)dx\,.$$

We must maximize A subject to the condition that $P[y] = L$, where

$$P[y] = \oint_C dl = \oint_C \sqrt{dx^2 + dy^2} = \oint_C \sqrt{1 + (y')^2}\,dx\,,$$

Using a Lagrange multiplier to impose the constraint, we have

$$\Phi_\lambda[y] = \oint_C f_\lambda(y, y')\, dx - \lambda L\,, \qquad f(y, y') = y - \lambda\sqrt{1 + (y')^2}\,.$$

We have to find the stationary values of this functional with respect to variations of the function y and of the real variable λ.

We do not have to worry about boundary terms in the variation of Φ_λ because there is no boundary, so the Euler-Lagrange equations apply. Furthermore, $f(y, y')$ has no explicit x-dependence, so the EL equations imply that

$$\text{constant} = f - y'\frac{\partial f_\lambda}{\partial y'} = y - \frac{\lambda}{\sqrt{1 + (y')^2}}\,.$$

This is equivalent to

$$(y')^2 = \frac{\lambda^2}{(y - y_0)^2} - 1$$

for some constant y_0. This ODE has the solution $y = y_0 \pm \sqrt{\lambda^2 - (x - x_0)^2}$ for some constant x_0, so

$$(x - x_0)^2 + (y - y_0)^2 = \lambda^2\,.$$

This is a circle of radius λ, which is fixed by the equation obtained by varying λ; this gives the original constraint that the circumference is L, so $2\pi\lambda = L$.

7.1 Sturm-Liouville problem

Another important constrained variational problem is a functional version of the problem of minimising a quadratic form subject to a normalization condition. Let $\rho(x)$, $\sigma(x)$ and $w(x)$ be real functions of x, defined for $\alpha \le x \le \beta$, such that both ρ and w are positive for $\alpha < x < \beta$, and consider the following real functionals of the real function $y(x)$:

$$F[y] = \int_\alpha^\beta \left\{\rho(x)\,(y')^2 + \sigma(x)\,y^2\right\} dx\,, \qquad G[y] = \int_\alpha^\beta w(x)\,y^2\,dx\,. \tag{7.1}$$

The problem is to find the function y that minimises $F[y]$ subject to the condition that $G[y] = 1$, given that $y(x)$ is fixed at $x = \alpha$ and $x = \beta$. The first task is to find the stationary values for this problem, and this can be done by finding the stationary values of

$$\Phi_\lambda[y] = F[y] - \lambda(G[y] - 1)$$

with respect to variations of $y(x)$, and λ. The EL equation for this functional is

$$\frac{\delta F[y]}{\delta y(x)} - \lambda\frac{\delta G[y]}{\delta y(x)} = 0 \qquad (\alpha < x < \beta). \tag{7.2}$$

Let's consider separately the variations of F and G with respect to a variation of $y(x)$:

$$\delta F = 2\int_\alpha^\beta \delta y\,\{-(\rho y')' + \sigma y\}\,dx - 2\,[\delta y\,\rho y']_\alpha^\beta\,,$$
$$\delta G = 2\int_\alpha^\beta \delta y\,wy\,dx$$

The boundary term in δF is zero because of the fixed-end boundary conditions, so

$$\frac{\delta F}{\delta y} = 2\mathcal{L}y\,, \qquad \frac{\delta G}{\delta y} = 2wy\,,$$

where $\mathcal{L}$ is the differential operator

$$\mathcal{L} = -\frac{d}{dx}\left(\rho(x)\frac{d}{dx}\right) + \sigma(x)\,.$$

In other words, $\mathcal{L}y = -(\rho y')' + \sigma y$ for any (twice-differentiable) function y. The EL equation (7.2) is therefore

$$\mathcal{L}y = \lambda w y\,. \tag{7.3}$$

This is an eigenvalue problem, with eigenvalue λ. The function $w(x)$ is called a "weight function". Many important ODEs are of Sturm-Liouville form, and one can find tables of the equations and their weight functions in texts on mathematical methods.

If the function $\sigma(x)$ is positive then $F \geq 0$, so its minimum is positive. Its minimum value is the lowest eigenvalue of the associated Sturm-Liouville eigenvalue problem. This can be seen as follows. Multiply both sides of (7.3) by y and integrate to get

$$\lambda G = \int_\alpha^\beta y\mathcal{L}y dx = F - [\rho y y']_\alpha^\beta\,,$$

where the second equality comes from an integration by parts. The boundary term is zero, so $\lambda = F/G \geq 0$. The original problem is equivalent to the problem of minimising F/G because the scale fixed by the normalization constraint drops out of this ratio, so the lowest eigenvalue will be the minimum of F/G.

Notice that F/G is not a functional of the type considered so far because it is a ratio of definite integrals. Nevertheless, it is still a functional. We can solve the problem directly by minimising $\Lambda = F/G$ without constraint. The functional $\Lambda[y]$ is stationary when

$$0 = \frac{\delta\Lambda}{\delta y} = \frac{1}{G}\left[\frac{\delta F[y]}{\delta y(x)} - \frac{F}{G}\frac{\delta G[y]}{\delta y(x)}\right] = \frac{2}{G}\left[\mathcal{L}y - \Lambda w y\right]\,,$$

so the values of Λ at its stationary points are the Sturm-Liouville eigenvalues, and the minimum value of Λ is the lowest SL eigenvalue, in agreement with the conclusion above deduced using the Lagrange multiplier method.

7.2 Function constraints; geodesics on surfaces

It can happen that we want to minimise a functional $F[\mathbf{x}]$ subject to a condition that restricts the functions $\mathbf{x}(\mathbf{t})$ *for all* t. In this case we need a Lagrange multiplier *function* $\lambda(t)$.

Suppose that we want to find the geodesics on a surface in Euclidean 3-space defined by the relation $g(\mathbf{x}) = 0$. We could solve this problem by first looking for the stationary points of the functional

$$\Phi[\mathbf{x};\lambda] = \int_0^1 \left\{ \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} - \lambda g(x,y,z) \right\} dt\,.$$

Here we are parametrising curves in the Euclidean 3-space between two points by an arbitrary parameter t, and using a Lagrange multiplier *function* $\lambda(t)$ to impose the constraint that the entire curve lie in the surface $g = 0$.

Alternatively, we could first try to solve the constraint $g = 0$. For example, if $g = x^2 + y^2 + z^2 - 1$ then the surface $g = 0$ is a unit sphere and we can solve the constraint by setting

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

The problem then reduces to minimising the distance functional

$$F[\theta,\phi] = \int_0^1 \sqrt{\dot{\theta}^2 + \sin^2\theta\,\dot{\phi}^2}\; dt$$

with respect to the functions $\theta(t)$ and $\phi(t)$. Equivalently, if θ is a good parameter for the curve, we can minimise the functional

$$F[\phi] = \int_{\theta_0}^{\theta_1} \sqrt{1 + \sin^2\theta(\phi')^2}\, d\theta\,,$$

where the curve is now specified by the function $\phi(\theta)$ (see Q.I.7).

8. Hamilton's principle

The time evolution of any mechanical system can be viewed as a trajectory in some multi-dimensional *configuration space*. For example, the configuration space of N point particles in a box is a space of dimension $3N$ because it takes 3 coordinates to specify the position of each of the N particles, and each of these can be changed independently of any change in the others. We may choose any coordinates we wish to indicate position in this configuration space; call them $\mathbf{q}$. The time-evolution of the system is then specified by functions $\mathbf{q}(t)$. Lagrange, who introduced this idea in his *Mechánique Analytique*, showed how to reduce problems in mechanics to a set of ODEs once both the kinetic energy T and the potential energy V are known in

terms of configuration space position $\mathbf{q}(t)$ and configuration space velocity $\dot{\mathbf{q}}(t)$. He made use of the principle of least action as formulated by Maupertuis, Euler and D'Alembert.

A limitation of the 18th century least action principle was that it assumed conservation of energy and only allowed variations of a given fixed energy. This restriction means that the principle determines only trajectories in configuration space; it does not provide information about position on this trajectory at a given time. About 50 years after Lagrange's work, Hamilton found an improved version of the least action principle that lifts these restrictions. This is often called the "least action principle" because it is the version of this principle in use today but in Hamilton's time it was called "Hamilton's principle" in order to distinguish it from the 18th century version.

Hamilton's first step was to define what he called the "Lagrangian", in honour of his intellectual hero. This is

$$L = T - V\,,$$

i.e. the *difference* between kinetic energy and potential energy V. The "action" for a path in confguration space between point A at time t_A and point B at time t_B is then defined to be

$$I[q] = \int_{t_A}^{t_B} L(t)\,dt$$

Hamilton's principle is the statement that the actual path taken is the one for which this functional is stationary.

For example, the configuration space of a single point particle is space itself, and we may choose cartesian coordinates $\mathbf{x}$ as coordinates on this 3-dimensional configuration space. In this case, for a particle of mass m we have

$$T = \frac{1}{2}m|\dot{\mathbf{x}}|^2\,, \qquad V = V(\mathbf{x},t)\,,$$

for a potential function of position that may also depend on time, so the Lagrangian is

$$L(\mathbf{x},\dot{\mathbf{x}};t) = \frac{1}{2}m|\dot{\mathbf{x}}|^2 - V(\mathbf{x},t)\,. \tag{8.1}$$

For a trajectory that starts at point A at time t_A and ends at point B at time $t_B > t_A$, the particle's action is[11]

$$I[\mathbf{x}] = \int_{t_A}^{t_B} L(\mathbf{x},\dot{\mathbf{x}};t)\,dt\,.$$

The Euler-Lagrange equations for this action are

$$0 = \frac{d}{dt}\frac{\partial L}{\partial \dot{x}_i} - \frac{\partial L}{\partial x_i} = \frac{d}{dt}\left(m\dot{x}_i\right) + \nabla_i V = m\ddot{x}_i - F_i\,.$$

[11]Notice that this has the same dimensions (ML^2/T) as the 18th century "action"; these are also the dimensions of Planck's constant.

In other words, for boundary conditions that make the boundary term in δI zero, Hamilton's principle implies Newton's second law:

$$\mathbf{F} = m\ddot{\mathbf{x}} \,, \qquad \mathbf{F} = -\boldsymbol{\nabla} V(\mathbf{x}, t) \,.$$

Notice that we allow the potential energy to be time-dependent. However, if it happens to be time-independent then the Lagrangian has no explicit dependence on t, which implies that there is a first-integral of the EL equations. The argument is just a repeat of one given earlier: the chain rule gives

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + \sum_{i=1}^{3} \left\{ \dot{x}_i \frac{\partial L}{\partial x_i} + \ddot{x}_i \frac{\partial L}{\partial \dot{x}_i} \right\} \,.$$

Using the EL eqs to rewrite the first term in the sum, we deduce that

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + \frac{d}{dt} \sum_{i=1}^{3} \dot{x}_i \frac{\partial L}{\partial \dot{x}_i} \,,$$

and hence that

$$\frac{d}{dt} \left[L - \sum_{i=1}^{3} \dot{x}_i \frac{\partial L}{\partial \dot{x}_i} \right] = \frac{\partial L}{\partial t} \,.$$

Given that $\partial L/\partial t = 0$ we deduce that

$$\text{constant} = \sum_{i=1}^{3} \dot{x}_i \frac{\partial L}{\partial \dot{X}_i} - L = m|\dot{\mathbf{x}}|^2 - L = \frac{1}{2} m|\dot{\mathbf{x}}|^2 + V(\mathbf{x}) \,,$$

so the constant of motion is the total energy $E = T + V$.

8.1 Central force fields

It was not necessary to use cartesian coordinates. In the important special case of a (time-independent) central force field, the potential V depends only on the distance r from the centre, and we will want to work with spherical polar coordinates (r, θ, φ), related to Cartesian coordinates by

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

We could apply Newton's second law in spherical polar coordinates but it is easier to first find the Lagrangian in these coordinates, using

$$\dot{x}^2 + \dot{y}^2 + \dot{z}^2 = \dot{r}^2 + r^2 \left(\dot{\theta}^2 + \sin^2\theta \dot{\phi}^2 \right) \,,$$

and then apply Hamilton's principle.

We will simplify the task by using the fact that the motion is planar (the plane being that normal to the constant angular momentum vector). Because of the spherical symmetry of the problem, we may choose the plane $\theta = \pi/2$ without loss of generality. In this case the Lagrangian simplifies to

$$L(r, \dot{r}, \phi, \dot{\varphi}; t) = \frac{1}{2} m \dot{r}^2 + \frac{1}{2} m r^2 \dot{\varphi}^2 - V(r) \,.$$

Notice that $\partial L / \partial \phi = 0$, so we have the first integral

$$\text{const.} = \frac{\partial L}{\partial \dot{\phi}} = m r^2 \dot{\phi} \quad \Rightarrow \quad \dot{\phi} = \frac{h}{r^2}$$

for a constant h that can be interpreted as angular momentum divided by the mass.

Notice too that $\partial L / \partial t = 0$, so we have another first integral :

$$\text{const.} = L - \dot{\phi} \frac{\partial L}{\partial \dot{\phi}} - \dot{r} \frac{\partial L}{\partial r} = -(T + V)$$

from which we deduce that

$$\frac{1}{2} m \dot{r}^2 + \frac{1}{2} m r^2 \left(\frac{h}{r^2} \right)^2 + V(r) = E$$

for constant E (total energy). We can rewrite this as

$$m \dot{r} = \sqrt{2m \left[E - V_{\text{eff}}(r) \right]} \,, \qquad V_{\text{eff}}(r) = V(r) + \frac{m h^2}{2 r^2} \,.$$

We now have a simple first-order ODE for r. Given a solution we can then solve the other first-order ODE $\dot{\phi} = h / r^2$ to find ϕ.

An important example of a central potential is

$$V(r) = -\frac{GMm}{r} \,,$$

where G is Newton's gravitational constant and M the mass of the sun. In this case

$$V_{\text{eff}}(r) = m \left(-\frac{GM}{r} + \frac{h^2}{2 r^2} \right) \,.$$

The structure of this "effective potential" leads to the following conclusions:

- Because $V_{\text{eff}} \propto m$, the motion of the "particle" of mass m (e.g. planet) will be independent of its mass m.
- The term proportional to h^2 is known as the "centrifugal barrier". It prevents any particle with non-zero angular momentum from reaching $r = 0$.

- The effective potential has one stationary point (for $h \neq 0$): a global minimum, at
$$r = \frac{h^2}{GM}$$
This implies a stable circular orbit at this radius.

- Non-circular but stable orbits exist for $E < 0$. They are ellipses but to prove this requires more effort.

8.2 The Hamiltonian and Hamilton's equations

The Lagrangian $L = T - V$ is a function of position and velocities. Usually, we assume that L is a convex function of the velocities, as it is for our point particle example. So let's take its Legendre transform with respect to the velocity $\mathbf{v} = \dot{\mathbf{x}}$. This gives the *Hamiltonian*

$$H(\mathbf{x}, \mathbf{p}; t) = [\mathbf{p} \cdot \mathbf{v} - L(\mathbf{x}, \mathbf{v})]_{\mathbf{v}=\mathbf{v}(\mathbf{p})} \tag{8.2}$$

where $\mathbf{v}(\mathbf{p})$ is the solution to $\partial L/\partial \mathbf{v} = \mathbf{p}$. For our point particle example, this equation is $m\mathbf{v} = \mathbf{p}$, so $\mathbf{p}$ is the particle's momentum, and

$$\mathbf{v}(\mathbf{p}) = \mathbf{p}/m\,.$$

The Hamiltonian for this case is

$$H(\mathbf{x}, \mathbf{p}) = \mathbf{p} \cdot \mathbf{v}(\mathbf{p}) - \frac{1}{2} m |\mathbf{v}(\mathbf{p})|^2 + V(\mathbf{x}, t) = \frac{|\mathbf{p}|^2}{2m} + V(\mathbf{x}, t)\,.$$

This is the total energy $T + V$ but expressed in terms of position and momenta rather than position and velocities. More generally, every position variable q has its "conjugate momentum" variable $p = \partial L/\partial \dot{q}$, and the Hamiltonian will be a function of these "conjugate pairs", which are "phase-space" coordinates.

Let's take the partial derivatives of the Hamiltonian, as given in (8.2), with respect to $\mathbf{x}$ and $\mathbf{p}$. We have

$$\frac{\partial H}{\partial p_i} = v_i + \left(\mathbf{p} - \frac{\partial L}{\partial \mathbf{v}}\right) \cdot \frac{\partial \mathbf{v}(\mathbf{p})}{\partial p_i} = v_i\,,$$

where the last equality follows from the fact that $\mathbf{p} = \partial L/\partial \mathbf{v}$ when $\mathbf{v} = \mathbf{v}(\mathbf{p})$. We also have

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

Using the EL equation we can rewrite this as

$$\frac{\partial H}{\partial \mathbf{x}} = -\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{x}}}\right) = -\dot{\mathbf{p}}\,.$$

To summarize: the EL equations imply **Hamilton's equations**

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

Notice that these equations are the EL equations for the "phase-space action" functional

$$I[\mathbf{x},\mathbf{p}] = \int \{\dot{\mathbf{x}}\cdot\mathbf{p} - H(\mathbf{x},\mathbf{p};t)\}\,dt\,.$$

Variation of $\mathbf{p}$ yields the first of Hamilton's equations, which we can use to solve for $\mathbf{p}$ in terms of $\dot{\mathbf{x}}$. Substitution then yields the integral of the Lagrangian (i.e. the action $I[\mathbf{x}]$) by the $f^{**} = f$ theorem that we proved for the Legendre transform. This shows that Hamilton's principle can be applied either to the action defined as the integral of the Lagrangian or to the above "phase-space action" .

9. Symmetries and Noether's theorem

We consider a system with Lagrangian of the form $L(\mathbf{q},\dot{\mathbf{q}};t)$. Let $\mathbf{Q}(t)$ denote a function of $\mathbf{q}(t)$ and its derivatives and, possibly, other given functions of t. If

$$L(\mathbf{Q},\dot{\mathbf{Q}};t) = L(\mathbf{q},\dot{\mathbf{q}};t) + \frac{dK}{dt} \tag{9.1}$$

for any function $K(t)$ (where the t-dependence may be explicit, through given functions of t, or implicit, through $\mathbf{q}(t)$ and its derivatives) then the equations of motion for $\mathbf{Q}$ will be identical to those for $\mathbf{q}$ (because dK/dt contributes only to the endpoints of the action integral). In this case we say that the "transformation" $\mathbf{q}(t) \to \mathbf{Q}(t)$ is a "symmetry" of the system.

Here we are interested in *continuous* symmetries, with continuous families of transformations that include the identity transformation $\mathbf{q}(t) \to \mathbf{q}(t)$. Let $\mathbf{q}(t) \to \mathbf{Q}_s(t)$ be a one-parameter family of transformations with $s = 0$ being the identity transformation. Then, for small s,

$$\mathbf{q}(t) \to \mathbf{Q}_s(t) = \mathbf{q}(t) + \delta_s\mathbf{q}(t) + \mathcal{O}(s^2)\,,$$

where $\delta_s\mathbf{q}(t)$ is the change in $\mathbf{q}(t)$ to first-order in the parameter s:

$$\delta_s\mathbf{q}(t) = s\boldsymbol{\xi}(t)\,, \qquad \boldsymbol{\xi}(t) = \left[d\mathbf{Q}_s(t)/ds\right]_{s=0}\,.$$

The change $\delta_s\mathbf{q}$ in $\mathbf{q}$ induces a corresponding change $\delta_s L$ in $L(\mathbf{q},\dot{\mathbf{q}};t)$:

$$L\left(\mathbf{Q}_s(t),\dot{\mathbf{Q}}_s(t);t\right) - L(\mathbf{q},\dot{\mathbf{q}};t) = \delta_s L + \mathcal{O}(s^2)\,, \tag{9.2}$$

where, by the chain rule,

$$\delta_s L = s\left[\boldsymbol{\xi}\cdot\frac{\partial L}{\partial \mathbf{q}} + \dot{\boldsymbol{\xi}}\cdot\frac{\partial L}{\partial \dot{\mathbf{q}}}\right]\,. \tag{9.3}$$

So far we used only the fact that $\mathbf{q}(t) \to \mathbf{Q}_s(t)$ is a one-parameter family of transformations, but if this family of transformations is also a family of symmetries then (9.1) holds in the form

$$L(\mathbf{Q}_s, \dot{\mathbf{Q}}_s; t) = L(\mathbf{q}, \dot{\mathbf{q}}; t) + \frac{dK_s}{dt}\,, \tag{9.4}$$

where $K_s(t)$ is a function of the parameter s (in addition to being a function of t) such that $K_0(t) \equiv 0$. For small s we therefore have

$$K_s(t) = sk(t) + \mathcal{O}(s^2)\,, \tag{9.5}$$

and hence

$$\delta_s L = s\frac{dk}{dt}\,. \tag{9.6}$$

By comparing this with (9.3) we deduce that for a continuous symmetry we must have

$$\boldsymbol{\xi} \cdot \frac{\partial L}{\partial \mathbf{q}} + \dot{\boldsymbol{\xi}} \cdot \frac{\partial L}{\partial \dot{\mathbf{q}}} = \frac{dk}{dt}\,, \tag{9.7}$$

for some function $k(t)$ (where again, the t-dependence may be explicit and/or implicit). We may rewrite this equation as

$$\boldsymbol{\xi} \cdot \left[\frac{\partial L}{\partial \mathbf{q}} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right)\right] + \frac{d}{dt}\left[\boldsymbol{\xi} \cdot \frac{\partial L}{\partial \mathbf{q}} - k\right] = 0\,.$$

For solutions of the Euler-Lagrange equations, this reduces to

$$\frac{d}{dt}\left[\boldsymbol{\xi} \cdot \frac{\partial L}{\partial \mathbf{q}} - k\right] = 0\,,$$

from which we deduce:

- **Noether's Theorem** (for Lagrangian mechanics). If $\mathbf{q} \to \mathbf{Q}_s$ is a one-parameter family of symmetries for a dynamical system with Lagrangiam L, as explained above, then

  $$\boldsymbol{\xi}(t) \cdot \frac{\partial L}{\partial \dot{\mathbf{q}}(t)} - k(t) \tag{9.8}$$

  is a constant of the motion.

Let us look some examples for Lagrangian $L(q, \dot{q}; t)$

1. Translation in configuration space. This is the transformation

   $$q(t) \to Q_s(t) = q(t) + s\,.$$

   In this case, $\delta_s q = s$ and $\delta_s L = s(\partial L/\partial q)$, so we have a symmetry (with $k = 0$) when $\partial L/\partial q = 0$, and the corresponding constant of the motion is

   $$\frac{\partial L}{\partial \dot{q}}\,,$$

   which is the momentum p conjugate to q.

2. Time translation. This is the transformation

$$q(t) \to Q_s(t) = q(t+s) = q(t) + s\dot{q}(t) + \mathcal{O}(s^2)\,.$$

In this case, $\delta_s q = s\dot{q}$ and

$$\delta_s L = s\left[\dot{q}\frac{\partial L}{\partial q} + \ddot{q}\frac{\partial L}{\partial \dot{q}}\right] = s\left[\frac{dL}{dt} - \frac{\partial L}{\partial t}\right]\,.$$

We see that

$$\frac{\partial L}{\partial t} = 0 \quad \Rightarrow \quad \delta_s L = s\frac{dL}{dt}$$

and hence we have a symmetry (with $k = L$) when L has no explicit time dependence. The corresponding constant of motion is

$$\dot{q}\frac{\partial L}{\partial \dot{q}} - L\,,$$

which is the energy.

9.0.1 A shortcut

Let us 'promote' the parameter s to a function $s(t)$. Then eq. (9.3) becomes

$$\begin{aligned}\delta_s L &= s\boldsymbol{\xi}\cdot\frac{\partial L}{\partial \mathbf{q}} + \frac{d}{dt}\left(s\boldsymbol{\xi}\right)\cdot\frac{\partial L}{\partial \dot{\mathbf{q}}} \\ &= s\left[\boldsymbol{\xi}\cdot\frac{\partial L}{\partial \mathbf{q}} + \dot{\boldsymbol{\xi}}\cdot\frac{\partial L}{\partial \dot{\mathbf{q}}}\right] + \dot{s}\left(\boldsymbol{\xi}\cdot\frac{\partial L}{\partial \dot{\mathbf{q}}}\right)\,. \end{aligned} \tag{9.9}$$

If the transformation $\mathbf{q} \to \mathbf{Q}_s$ is a symmetry for $\dot{s} = 0$ then we know that

$$\boldsymbol{\xi}\cdot\frac{\partial L}{\partial \mathbf{q}} + \dot{\boldsymbol{\xi}}\cdot\frac{\partial L}{\partial \dot{\mathbf{q}}} = \dot{k}\,,$$

for some function $k(t)$. Using this in (9.9) we have

$$\delta_s L = \dot{s}\left(\boldsymbol{\xi}\cdot\frac{\partial}{\partial \dot{\mathbf{q}}} - k\right) + \frac{d(sk)}{dt}\,.$$

This reduces to (9.6) when $\dot{s} = 0$, as expected, and the coefficient of $\dot{s}$ is the constant of motion corresponding to the symmetry transformation for constant parameter s. To summarize:

- By allowing $s \to s(t)$ we can both check that $\delta_s L = s\dot{k}$ when $\dot{s} = 0$, *and* read off the corresponding constant of motion from the coefficient of $\dot{s}$ in $\delta_s L$.

In what follows, we shall use this as a shortcut.

N.B. What we are calling "Noether's theorem" is sometimes called "Noether's first theorem". Noether's "second theorem" (not part of this course) applies in the special case that the constant of motion is zero, in which case $\delta_s L = d(sk)/dt$ for arbitrary parameter *function* $s(t)$. This used to be called a "symmetry of the second kind" but is now called a "gauge invariance".

9.1 Application to Hamiltonian mechanics

Consider a particle in a force field $\mathbf{F} = -\boldsymbol{\nabla} V$. The phase-space action is

$$I[\mathbf{x}, \mathbf{p}] = \int \{\mathbf{p} \cdot \dot{\mathbf{x}} - H(\mathbf{x}, \mathbf{p})\}\, dt\,, \qquad H(\mathbf{x}, \mathbf{p}) = \frac{1}{2m}|\mathbf{p}|^2 + V(\mathbf{x}, t)\,.$$

The symmetries of this action depend on properties of V. We shall consider three cases.

1. **Space translation invariance**. For the transformation $\mathbf{x} \to \mathbf{x} + \mathbf{a}$, we find that

   $$\delta_{\mathbf{a}} I = \int \{-\mathbf{a} \cdot \boldsymbol{\nabla} V + \dot{\mathbf{a}} \cdot \mathbf{p}\}\, dt\,.$$

   So we have a symmetry for constant $\mathbf{a}$ if V is position independent, and then $\mathbf{p}$ is a constant of the motion. *Translation invariance implies conservation of momentum.*

2. **Rotation invariance**. If the potential V depends on position only through distance $|\mathbf{x}|$ from the origin, the action I is unchanged, to first-order in $\boldsymbol{\omega}$, by the transformation

   $$\mathbf{x} \to \mathbf{x} + \boldsymbol{\omega} \times \mathbf{x}\,, \qquad \mathbf{p} \to \mathbf{p} + \boldsymbol{\omega} \times \mathbf{p}\,,$$

   for constant vector $\boldsymbol{\omega}$. This is a rotation. Allowing $\boldsymbol{\omega}$ to be time-dependent, one then finds, to first order in $\boldsymbol{\omega}$, that

   $$\delta_{\boldsymbol{\omega}} I = \int \dot{\boldsymbol{\omega}} \cdot \mathbf{L}\, dt\,, \qquad \mathbf{L} = \mathbf{x} \times \mathbf{p}\,.$$

   It follows from Noether's theorem that the vector $\mathbf{L}$, which is the particle's angular momentum, is constant as a consequence of Hamilton's equations (and this is easily verified). *Rotation invariance implies conservation of angular momentum.*

3. **Time translation invariance**.

   Time translation is equivalent to the transformation

   $$\begin{aligned}\mathbf{x}(t) &\to \mathbf{x}(t+s) = \mathbf{x}(t) + s\dot{\mathbf{x}}(t) + (s^2)\\ \mathbf{p}(t) &\to \mathbf{p}(t+s) = \mathbf{p}(t) + s\dot{\mathbf{p}}(t) + (s^2)\,,\end{aligned}$$

   To leading order in s, this induces the following change in the action:

   $$\begin{aligned}\delta_s I &= \int \left\{-s\dot{\mathbf{x}} \cdot \frac{\partial H}{\partial \mathbf{x}} + s\dot{\mathbf{p}} \cdot \left(\dot{\mathbf{x}} - \frac{\partial H}{\partial \mathbf{p}}\right) + \frac{d}{dt}\,(s\dot{\mathbf{x}}) \cdot \mathbf{p}\right\} dt\\ &= \int \left\{-s\left(\dot{\mathbf{x}} \cdot \frac{\partial H}{\partial \mathbf{x}} + \mathbf{p} \cdot \frac{\partial H}{\partial \mathbf{p}}\right) + \frac{d}{dt}\,(s\dot{\mathbf{x}} \cdot \mathbf{p})\right\} dt\,.\end{aligned}$$

Using the identity

$$\dot{\mathbf{x}} \cdot \frac{\partial H}{\partial \mathbf{x}} + \mathbf{p} \cdot \frac{\partial H}{\partial \mathbf{p}} \equiv \frac{dH}{dt} - \frac{\partial H}{\partial t}$$

we have

$$\delta_s I = \int \left\{ s \frac{\partial H}{\partial t} + \frac{d}{dt} \left[s(\dot{\mathbf{x}} \cdot \mathbf{p} - H) \right] + \dot{s} H \right\} dt \, .$$

If we now suppose that the potential V has no explicit time dependence then neither does the Hamiltonian, and hence

$$\delta_s I = \int \{ \dot{s} H \} \, dt + \left[s(\dot{\mathbf{x}} \cdot \mathbf{p} - H) \right]_{t_A}^{t_B} \, .$$

For $\dot{s} = 0$ the action changes by a boundary term, which does not affect the equations of motion, so we have a symmetry, and from the coefficient of $\dot{s}$ we see that the corresponding constant of motion is the Hamiltonian. This is the total energy, as we saw previously.

Time translation invariance implies conservation of energy.

10. PDEs from variational principles

Now we consider functionals for functions of more than one independent variable. The general case that we consider is functionals of functions $\mathbf{y} : \mathbb{R}^m \to \mathbb{R}^n$, for $m > 1$, expressed as integrals of the form

$$F[\mathbf{y}] = \int dx_1 \cdots \int dx_m \, f(\mathbf{y}, \nabla \mathbf{y}; x_1, \cdots, x_m) \, ,$$

where

$$\nabla \mathbf{y} = \left(\frac{\partial \mathbf{y}}{\partial x_1}, \ldots, \frac{\partial \mathbf{y}}{\partial x_m} \right) \, .$$

Stationary points of such functions are solutions of PDEs in m variables for the n functions $\mathbf{y}$. In principle, it is possible to derive a generalisation of the EL equation for such functionals, but it is as easy to consider the variation of F on a case by case basis. In what follows we consider a few important examples. Little attention will be paid to boundary terms or boundary conditions.

10.1 Minimal surfaces

A minimal surface is a higher-dimensional analog of a geodesic. Instead of asking for a curve of minimal length we ask for a surface of minimal area. Consider a surface S in $\mathbb{E}^3$ specified by a constraint $g(\mathbf{x}) = 0$ on the cartesian coordinates $\mathbf{x} = (x, y, z)$. Suppose now that this constraint can be solved in the form

$$z = h(x, y) \, ,$$

where h is a height function. This assumes that (x, y) are "good" coordinates for the surface S, which may not be true everywhere on the surface, so let's restrict to a region D of the x-y plane for which it is true. The area of that part of the surface S above this region is given by

$$A[h] = \iint_D dxdy \sqrt{1 + h_x^2 + h_y^2}\,, \qquad h_x = \frac{\partial h}{\partial x}\,, \quad h_y = \frac{\partial h}{\partial y}\,. \tag{10.1}$$

Here we have a functional of a function $h(x, y)$ of two variables.

Aside. Here is how the above formula is arrived at. The squared length element for a curve in S is

$$d\ell^2 = dx^2 + dy^2 + (h_x dx + h_y dy)^2 = d\mathbf{x}^T g\, d\mathbf{x}\,, \qquad g = \begin{pmatrix} 1 + h_x^2 & h_x h_y \\ h_x h_y & 1 + h_y^2 \end{pmatrix}\,.$$

The area element on S is $dA = \sqrt{\det g}\, dxdy = \sqrt{1 + h_x^2 + h_y^2}\, dxdy$. Integrating over D we arrive at (10.1).

Suppose that we wish to find the surfaces for which A is a minimum for specified boundary conditions; these are called "minimal surfaces". Then we must first find the functions $h(x, y)$ that make stationary the functional $A[h]$. Consider a variation $h(x, y) \to h(x, y) + \delta h(x.y)$. This gives

$$h_\alpha(x, y) \to h_\alpha(x, y) + \nabla_\alpha \delta h(x, y)\,, \qquad \alpha = 1, 2\,,$$

and hence

$$A[h] \to A[h] + \iint_D dxdy \left\{ \frac{h_x \nabla_x \delta h + h_y \nabla_y \delta h}{\sqrt{1 + h_x^2 + h_y^2}} \right\} + \mathcal{O}(\delta h^2)\,.$$

Call the integral expression $\delta A[y]$. It is the first-order change in $A[h]$. Integrating by parts in this integral, we have

$$\delta A[h] = -\iint_D dxdy \left\{ \delta h \left[\nabla_x \left(\frac{h_x}{\sqrt{1 + h_x^2 + h_y^2}} \right) + \nabla_y \left(\frac{h_y}{\sqrt{1 + h_x^2 + h_y^2}} \right) \right] \right\} + \text{b.t.}\,,$$

where "b.t." is a boundary term. To deal properly with the boundary term we should consider whether we actually want to impose boundary conditions at the boundary of the region D or whether the surface should be considered to extend beyond the region D, where we will need a different parametrisation of it. In any case, it is now clear that any minimal surface will satisfy the non-linear PDE

$$\nabla_x \left(\frac{h_x}{\sqrt{1 + h_x^2 + h_y^2}} \right) + \nabla_y \left(\frac{h_y}{\sqrt{1 + h_x^2 + h_y^2}} \right) = 0\,.$$

This is equivalent to the *minimal surface equation*

$$\left(1+h_y^2\right)h_{xx} + \left(1+h_x^2\right)h_{yy} - 2h_x h_y h_{xy} = 0\,.$$

If we can ignore the non-linearities on the grounds that $|\boldsymbol{\nabla} h| \ll 1$, then the minimal surface equation becomes $h_{xx} + h_{yy} = 0$, which is the Laplace equation

$$\nabla^2 h = 0\,.$$

One obvious solution of the non-linear PDE is

$$h(x,y) = Ax + By + C$$

for constants (A, B, C). This is the equation of a plane in $\mathbb{E}^3$. Less obvious solutions are hard to find.

Solutions with circular symmetry (a surface of revolution) can be found by supposing that

$$h(x,y) = z(r)\,, \qquad r = \sqrt{x^2+y^2}$$

The minimal surface equation then reduces to an ODE for $z(r)$:

$$rz'' + z' + (z')^3 = 0\,. \tag{10.2}$$

This looks a bit difficult to solve, so let's consider an alternative procedure (see Q.I.8). We first substitute $h(x,y) = z(r)$ into the function of (10.1) to get the simpler functional

$$A[z] = 2\pi \int \left\{ r\sqrt{1+(z')^2} \right\} dr\,.$$

The integrand $f = r\sqrt{1+(z')^2}$ depends on z only through its derivative, so we have the first-integral

$$\frac{d}{dr}\left[\frac{rz'}{\sqrt{1+(z')^2}}\right] = 0 \quad \Rightarrow \quad \frac{rz'}{\sqrt{1+(z')^2}} = r_0\,,$$

for constant r_0. Any solution of this first-order ODE will solve (10.2), as you may verify by taking the derivative of both sides. The first order ODE is easily solved, and the solution is $z = z_0 + r_0 \cosh^{-1}(r/r_0)$ for integration constant z_0; equivalently,

$$r = r_0 \cosh\left(\frac{z-z_0}{r_0}\right)\,.$$

This is a "catenoid", the minimal surface of revolution found by Euler in his treatise of 1744.

10.2 Small amplitude oscillations of a uniform string

For a string of uniform tension T and uniform mass density ρ, stretched on the x axis between the origin and $x = a$, small displacements in the y direction are associated with the following kinetic and potential energies:

$$K.E. = \frac{1}{2}\rho \int_0^a \dot{y}^2 dx\,, \qquad P.E. = \frac{1}{2}T \int_0^a (y')^2 dx\,.$$

Here, an overdot means partial derivative with respect to t and a prime means partial derivative with respect to x. The action for this system is therefore

$$S[y] = \frac{1}{2}\int dt \int_0^a \left\{\rho\dot{y}^2 - Ty')^2\right\} dx\,.$$

The variation of $S[y]$, given a variation δy in y, is

$$\delta S = \int dt \int_0^a \left\{\rho\dot{y}\frac{\partial \delta y}{\partial t} - Ty'\frac{\partial \delta y}{\partial x}\right\} dx\,.$$

Integrating by parts, and assuming that the boundary conditions are such that the boundary terms are zero[12], we have

$$\delta S = \int dt \int_0^a \left\{\delta y\left[-\rho\ddot{y} + Ty''\right]\right\} dx$$

The action is stationary for arbitrary $\delta y(t, x)$ (subject to boundary conditions) iff

$$\ddot{y} - v^2 y'' = 0\,, \qquad v \equiv \sqrt{T/\rho}\,.$$

Notice that this equation can be written in factorised form as

$$\left(\frac{\partial}{\partial t} - v\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial t} + v\frac{\partial}{\partial x}\right) y = 0\,,$$

which shows that *either* $\dot{y} = vy'$ *or* $\dot{v} = -vy'$. The general solution is therefore

$$y(x, t) = f_+(x + vt) + f_-(x - vt)$$

for functions $f_\pm$ of a single variable. This is a superposition of two wave profiles, one moving to the left and the other to the right, both with speed v.

10.3 Maxwell's equations from Hamilton's principle

Consider the action

$$S[\mathbf{A}, \varphi] = \int dt \left\{T[\mathbf{A}, \phi] - V[\mathbf{A}, \phi]\right\}\,,$$

[12]For example, fixed y at $x = 0, a$ and at initial and final times.

where the kinetic and potential energies are functionals of a vector field $\mathbf{A}(\mathbf{x},t)$ and a scalar field $\phi(\mathbf{x},t)$. Specifically,

$$T = \frac{1}{2}\int d^3x\ \left\{|\mathbf{E}|^2 + \mathbf{A}\cdot\mathbf{j}\right\}\ , \qquad V = \frac{1}{2}\int d^3x\ \left\{|\mathbf{B}|^2 + \phi\rho\right\}\ ,$$

where $\mathbf{j}(\mathbf{x},t)$ is a given vector field, $\rho(\mathbf{x},t)$ is a given scalar field, and

$$\mathbf{E} = -\boldsymbol{\nabla}\phi - \partial_t\mathbf{A}\,, \qquad \mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A} \qquad \left(\partial_t = \frac{\partial}{\partial t}\right).$$

We will interpret $\mathbf{E}$ as the electric field and $\mathbf{B}$ as the magnetic field; in this case ρ is the electric charge density and $\mathbf{j}$ the electric current density. These definitions of electric and magnetic fields in terms of the "vector potential" $\mathbf{A}$ and "scalar potential" φ imply the two equations

$$\boldsymbol{\nabla}\cdot\mathbf{B} = 0\,, \qquad \boldsymbol{\nabla}\times\mathbf{E} = -\partial_t\mathbf{B}\,. \tag{10.3}$$

The first of these equations says that there are no magnetic monopoles. The second equation is Faraday's law of induction.

Let us now apply Hamilton's principle to this action. A variation of $\mathbf{A}$ and φ induces the following variation of S:

$$\begin{aligned}\delta S &= \int dt\int d^3x\,\{-\mathbf{E}\cdot[\partial_t(\delta\mathbf{A}) + \boldsymbol{\nabla}(\delta\phi)] - \mathbf{B}\cdot\boldsymbol{\nabla}\times(\delta\mathbf{A}) + \delta\mathbf{A}\cdot J - \delta\phi\,\rho\}\\ &= \int dt\int d^3x\left\{\delta\mathbf{A}\cdot\left[\dot{\mathbf{E}} - \boldsymbol{\nabla}\times\mathbf{B} + \mathbf{j}\right] + \delta\phi\,(\boldsymbol{\nabla}\cdot\mathbf{E} - \rho)\right\} + b.t.\end{aligned}$$

Assuming that the boundary conditions are such that the boundary term is zero, we see that the action is stationary for solutions of the equations

$$\boldsymbol{\nabla}\cdot\mathbf{E} = \rho\,, \qquad \boldsymbol{\nabla}\times\mathbf{B} = \mathbf{j} + \partial_t\mathbf{E}\,. \tag{10.4}$$

The first of these equations is the Gauss law of electrostatics. Without the $\partial_t\mathbf{E}$ term, the second is Ampère's law; Maxwell discovered that the combined laws of electricity and magnetism were not mutually consistent unless Ampère's law was modified to include the extra term, which he called the "displacement current". For some (non-standard) choice of units, the combined equations of (10.3) and (10.4) are what are now known as Maxwell's equations.

11. The second variation

We are now going to consider the expansion of a functional $F[y+\delta y]$ to second order in δy. Let's first return to the case of a function $f(\mathbf{x})$ of many variables $\mathbf{x}$ and write $\delta\mathbf{x} = \epsilon\boldsymbol{\xi}$ for "small" ϵ. We then have

$$f(\mathbf{x}+\epsilon\boldsymbol{\xi}) - f(\mathbf{x}) = \epsilon\,\boldsymbol{\xi}\cdot\nabla f(\mathbf{x}) + \frac{1}{2}\epsilon^2\xi_i H_{ij}(\mathbf{x})\xi_j + \mathcal{O}(\epsilon^3)$$

If the Hessian matrix $H(\mathbf{x})$ is positive (no zero eigenvalues) then the quadratic form $\xi_i H_{ij}(\mathbf{x})\xi_j$ will be non-negative.

A similar expansion for functionals $F[y] = \int_\alpha^\beta f(y, y'; x)dx$ will give us information about the nature of stationary points of $F[y]$, so we write

$$\delta y(x) = \epsilon\xi(x)\,,$$

and expand the integrand of F to second-order in ϵ:

$$\begin{aligned} f(y+\epsilon\xi, y'+\epsilon\xi'; x) = f(y, y'; x) &+ \epsilon\left\{\xi\left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right] + \frac{d}{dx}\left[\xi\,\frac{\partial f}{\partial y'}\right]\right\} \\ &+ \frac{1}{2}\epsilon^2\left\{\xi^2\frac{\partial^2 f}{\partial y^2} + 2\xi\xi'\frac{\partial^2 f}{\partial y\partial y'} + \xi'^2\frac{\partial^2 f}{\partial y'^2}\right\} + \mathcal{O}(\epsilon^3) \end{aligned}$$

We restrict to fixed-end boundary conditions at $x = \alpha, \beta$, which means that

$$\xi(\alpha) = \xi(\beta) = 0\,. \tag{11.1}$$

In this case, we have

$$F[y+\epsilon\xi] - F[y] = \epsilon\int_\alpha^\beta \xi(x)\frac{\delta F[y]}{\delta y(x)}\,dx \;+ \epsilon^2\,\delta^2 F[y,\xi] + \mathcal{O}(\epsilon^3)$$

where

$$\delta^2 F[y,\xi] = \frac{1}{2}\int_\alpha^\beta \left\{\xi^2\frac{\partial^2 f}{\partial y^2} + 2\xi\xi'\frac{\partial^2 f}{\partial y\partial y'} + \xi'^2\frac{\partial^2 f}{\partial y'^2}\right\}dx\,.$$

Using the fact that $\xi\xi' = (\xi^2)'/2$, and integrating by parts, we can rewrite this "second variation" as

$$\delta^2 F[y,\xi] = \frac{1}{2}\int_\alpha^\beta \left\{\xi^2\left[\frac{\partial^2 f}{\partial y^2} - \frac{d}{dx}\left(\frac{\partial^2 f}{\partial y\partial y'}\right)\right] + \xi'^2\frac{\partial^2 f}{\partial y'^2}\right\}dx\,. \tag{11.2}$$

This is a functional of both $y(x)$ and $\xi(x)$.

Recall that if $H(\mathbf{x})$ is positive for all $\mathbf{x}$ then $f(\mathbf{x})$ is convex and hence any stationary point of f will be an absolute minimum. An analogous result holds for functionals $F[y]$:

- If $\delta^2 F[y,\xi] \geq 0$ for all allowed functions ξ and all functions y satisfying appropriate boundary (and differentiability) conditions then $F[y]$ has an absolute minimum for the solution $y_0(x)$ of the EL equation that satisfies the chosen boundary conditions.

  For $\xi(x)$ to be "allowed" the function $y(x)+\epsilon\xi(x)$ must satisfy the same boundary (and differentiability) conditions as $y(x)$. For fixed-end boundary conditions at $x = \alpha, \beta$ this means that $\xi(x)$ must satisfy (11.1). It then follows that any allowed function $\xi(x)$ that is not identically zero in the interval $[\alpha, \beta]$ must be such that $\xi'(x) \neq 0$ in some sub-interval $I \subset [\alpha, \beta]$.

Example: geodesics in the Euclidean plane. We know that a straight line solves the EL equation of the path-length functional $F[y] = \int_\alpha^\beta \sqrt{1+(y')^2}\,dx$, and also that this solution is unique once we specify the two endpoints, but does this straight line actually *minimise* the distance between the points? Of course it does, but let's check this by looking at the second variation. In this case

$$\frac{\partial^2 f}{\partial y^2} = 0\,, \qquad \frac{\partial^2 f}{\partial y \partial y'} = 0\,, \qquad \frac{\partial^2 f}{\partial y' \partial y'} = \left[1+(y')^2\right]^{-\frac{3}{2}}\,,$$

so that

$$\delta^2 F[y,\xi] = \frac{1}{2}\int_\alpha^\beta \left\{ \left[1+(y')^2\right]^{-\frac{3}{2}} (\xi')^2 \right\} dx\,.$$

This is positive for all ξ and y so a straight line really does minimise the distance between two points.

N.B. Q.II.9 asks you to make the analogous computation for geodesics on the sphere using the distance functional $F[\phi] = \int \sqrt{1+\sin^2\theta\,\phi'^2}\,d\theta$. We know that the EL equation is solved by a path that is an arc of a great circle, but there are two such paths, and one is longer than the other (unless the two endpoints are antipodes). However, θ is never a good parameter on the longer arc (it will not increase or decrease monotonically) so the validity of your result holds only for arcs for which $F \leq \pi$.

Returning to the general case, suppose we have found that $F[y]$ is stationary for solution $y_0(x)$ of the EL equation. There is a simple test to see whether $y_0(x)$ stands a chance of being a local minimum of $F[y]$. If it *is* a local minimum then it must satisfy the *Legendre condition*

$$\left.\frac{\partial^2 f}{\partial y' \partial y'}\right|_{y_0} \geq 0\,. \tag{11.3}$$

The idea of the proof is as follows. If the Legendre condition is not satisfied in some interval then there will be a negative contribution to $\delta^2 F[y_0,\xi]$ in this interval, weighted by ξ'^2. There may also be a positive contribution weighted by ξ^2, not to mention positive contributions from elsewhere. However, if we can choose ξ to be a smooth bounded function such that ξ varies arbitrarily rapidly within the interval in which $\partial^2 f/\partial y'^2 < 0$ then the integral of $(\xi')^2 \partial^2 f/\partial y'^2$ over this interval can be made arbitrarily negative. The necessity of the Legendre condition is then proved by exhibiting a smooth functions ξ with these properties (see e.g. the formula (15) in Chapter 5 of Gelfand and Fomin).

The Legendre condition is obviously satisfied by the distance functional in the plane and on the sphere because in these cases the LHS of (11.3) is non-negative for all y, not just for the solution y_0 of the EL equations. In other cases, it can allow rapid elimination of the possibility that a minimum exists. Here is an example

- $F[y] = \int_{-1}^{1} x\sqrt{1+y'^2}\,dx$. In this case,

$$\frac{\partial^2 f}{\partial y' \partial y'} = \left[1+y'^2\right]^{-\frac{3}{2}} x$$

  which changes sign at $x = 0$. So we can say in advance of solving the EL equation that any solution we find will *not* be a minimum of the functional, although it could be a minimum of the same integral with different integration limits (and hence different boundary conditions).

Given that we have a solution $y = y_0$ for which $F[y]$ is stationary, it is convenient to rewrite the second-variation formula (11.2) in the Sturm-Liouville form,

$$\delta^2 F[y_0, \xi] = \frac{1}{2}\int_{\alpha}^{\beta} \left\{\rho(x)\xi'^2 + \sigma(x)\xi^2\right\} dx \tag{11.4}$$

where

$$\rho(x) = \left.\frac{\partial^2 f}{\partial y'^2}\right|_{y=y_0}, \qquad \sigma(x) = \left[\frac{\partial^2 f}{\partial y^2} - \frac{d}{dx}\left(\frac{\partial^2 f}{\partial y \partial y'}\right)\right]\Bigg|_{y=y_0}.$$

As we have just seen, a *necessary* condition for y_0 to minimise $F[y]$ is that $\rho \geq 0$, but (as we shall see later) this condition is *not sufficient*.

A condition that *is* sufficient for $F[y]$ to have a local minimum at $y = y_0$ is

$$\rho(x) > 0 \quad \& \quad \sigma(x) \geq 0 \qquad \alpha < x < \beta$$

because in this case $\delta^2 F[y_0, \xi] > 0$ for all allowed functions ξ that are not identically zero. The qualification "allowed" is important here because in the case that $\sigma(x) \equiv 0$ we have $\delta^2 F[y_0, \xi] = 0$ for $\xi =$ const. but the only "allowed" constant is zero, which would make ξ identically zero. From this we see that if $\xi(x)$ is not identically zero then $\xi'(x) \neq 0$ somewhere in the interval $[\alpha, \beta]$, and for sufficiently smooth functions this implies that it will be non-zero in some sub-interval, in which case $\delta^2 F[y_0, \xi] > 0$.

Example: brachistochrone: Recall that the time for the bead to go from A to B can be expressed as functional $T[y]$ of the function $y(x)$ giving the vertical distance dropped in terms of horizontal distance travelled, and that $T[y]$ is stationary for a cycloid $y = y_0(x)$ satisying $y[1 + (y')^2] = 2c$ for positive constant c. Now we investigate the second variation of $T[y]$ at this solution.

Recall that the integrand of $T[y]$ is proportional to $f = \sqrt{\frac{1+(y')^2}{y}}$ for $y > 0$. The first derivatives of f are

$$\frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{y[1+(y')^2]}}, \qquad \frac{\partial f}{\partial y} = -\frac{f}{2y}.$$

The second derivatives are

$$\frac{\partial^2 f}{\partial y' \partial y'} = \frac{1}{\sqrt{y[1+(y')^2]^3}}, \qquad \frac{\partial^2 f}{\partial y \partial y'} = -\frac{1}{2y}\frac{\partial f}{\partial y'}, \qquad \frac{\partial^2 f}{\partial y \partial y} = \frac{3f}{4y^2}.$$

Using these results, and then the EL equation, we find that

$$\frac{\partial^2 f}{\partial y^2} - \frac{d}{dx}\left(\frac{\partial^2 f}{\partial y \partial y'}\right) = \frac{\partial^2 f}{\partial y \partial y} + \frac{d}{dx}\left(\frac{1}{2y}\frac{\partial f}{\partial y'}\right) = \frac{\partial^2 f}{\partial y \partial y} - \frac{y'}{2y^2}\frac{\partial f}{\partial y'} + \frac{1}{2y}\frac{\partial f}{\partial y}$$
$$= \frac{f}{2y^2} - \frac{(y')^2}{2y^2\sqrt{y[1+(y')^2]}} = \frac{1}{2y^2\sqrt{y[1+(y')^2]}}\,. \qquad (11.5)$$

This gives

$$\rho = \frac{1}{\sqrt{y[1+(y')^2]^3}}\bigg|_{y=y_0} > 0\,, \qquad \sigma(x) = \frac{1}{2y^2\sqrt{y[1+(y')^2]}}\bigg|_{y=y_0} > 0\,,$$

so $\delta^2 T[y_0, \xi] > 0$ and hence that the cycloid is (at least) a local minimiser of $T[y]$.

12. The Jacobi condition

The Legendre condition $\rho \geq 0$ is necessary for $\delta^2 F[y_0, \xi]$ to be positive, and hence for $F[y]$ to have a local minimum at $y = y_0$. Legendre attempted to prove that the stronger condition (now called the "strong Legendre condition")

$$\rho(x) > 0\,, \qquad \alpha < x < \beta$$

is sufficient for $\delta^2 F[y_0, \xi]$ to be positive. He failed, because it isn't sufficient, but his idea was a good one, as we'll now see.

First we observe that (because $\xi(\alpha) = \xi(\beta) = 0$)

$$0 = \int_\alpha^\beta \left(\varphi\xi^2\right)' dx = \int_\alpha^\beta \left[2\varphi\xi\xi' + \varphi'\xi^2\right] dx$$

for any function $\varphi(x)$. This allows us to rewrite the second variation at $y = y_0$ as

$$\delta^2 F[y_0, \xi] = \frac{1}{2}\int_\alpha^\beta \left\{\rho(\xi')^2 + 2\varphi\xi\xi' + (\sigma + \varphi')\,\xi^2\right\} dx\,.$$

Given $\rho > 0$, we can complete the square in ξ' to get

$$\delta^2 F[y_0, \xi] = \frac{1}{2}\int_\alpha^\beta \left\{\rho\left(\xi' + \frac{\varphi}{\rho}\xi\right)^2 + \left(\sigma + \varphi' - \frac{\varphi^2}{\rho}\right)\xi^2\right\} dx\,.$$

This is manifestly positive if φ is chosen to satisfy the "Ricatti" equation[13]

$$\varphi^2 = \rho(\sigma + \varphi')\,. \qquad (12.1)$$

[13] A Ricatti equation for function φ equates φ' to an expression quadratic in φ with known functions as coefficients.

Moreover, $\delta^2 F[y_0, \xi]$ can be zero only if

$$\xi' + (\varphi/\rho)\xi = 0 \quad \Rightarrow \quad \xi(x) = C \exp\left\{ - \int_\alpha^x \frac{\varphi(s)}{\rho(s)} ds \right\} ,$$

for constant C, but this satisfies the boundary conditions on ξ only if $C = 0$, so $\delta^2 F[y_0, \xi] > 0$ for any allowed non-zero ξ.

It seems that we have now proved that the strong Legendre condition is sufficient for y_0 to be a minimum of $F[y]$ but the proof hinges upon the existence of a solution $\varphi(x)$ to the Ricatti equation (12.1). We can recast this first-order but non-linear ODE as a linear, but second-order, ODE for a new function $u(x)$, by setting

$$\varphi = -\rho \frac{u'}{u} .$$

Substitution into (12.1) gives

$$\rho \left(\frac{u'}{u} \right)^2 \; = \; \sigma - \left(\frac{\rho u'}{u} \right)' \; = \; \sigma - \frac{(\rho u')'}{u} + \rho \left(\frac{u'}{u} \right)^2 ,$$

and hence

$$- (\rho u')' + \sigma u = 0 . \tag{12.2}$$

This called the *Jacobi accessory equation*. If we can find a solution to it with the property that

$$u(x) \neq 0 , \qquad \alpha < x < \beta ,$$

then we have a solution to (12.1) for φ. We need to impose this condition because otherwise φ becomes infinite somewhere in the interval (α, β), but u is *not subject to any boundary conditions at* $x = \alpha, \beta$.

The conditions under which (12.2) has a nowhere-zero solution for u were studied by Jacobi in the 19th century. By "nowhere" we mean here nowhere in the interval (α, β). Such a solution always exists (given $\rho > 0$) for a sufficiently small interval but this may change when the interval becomes too large. We shall illustrate this with an example:

Example: Geodesics on the sphere. The length L of a curve C on the unit-radius sphere is

$$L = \int_C \sqrt{d\theta^2 + \sin^2\theta \, d\phi^2} .$$

Previously we have used the polar angle θ as a parameter on the curve, so that L becomes a functional of the function $\phi(\theta)$ used to specify the path. But we can also use the azimuthal angle ϕ as a parameter, in which case L becomes a functional of the function $\theta(\phi)$:

$$L[\theta] = \int_{\phi_1}^{\phi_2} \sqrt{(\theta')^2 + \sin^2\theta} \; d\phi \quad \Rightarrow \quad f(\theta, \theta'; \phi) = \sqrt{(\theta')^2 + \sin^2\theta} .$$

In this case, we have

$$\frac{\partial f}{\partial \theta} = \frac{\sin\theta\cos\theta}{\sqrt{(\theta')^2+\sin^2\theta}}\,, \qquad \frac{\partial f}{\partial \theta'} = \frac{\theta'}{\sqrt{(\theta')^2+\sin^2\theta}}\,.$$

Since $\frac{\partial f}{\partial \phi} = 0$, the EL equations have the first integral

$$\text{constant} = f - \theta'\frac{\partial f}{\partial \theta'} = \frac{\sin^2\theta}{\sqrt{(\theta')^2+\sin^2\theta}}\,,$$

which is equivalent to

$$\sqrt{(\theta')^2+\sin^2\theta} = c\sin^2\theta \qquad (c \geq 1).$$

The solution for $c = 1$ is $\theta_0(\phi) = \pi/2$, which connects any two points on the equator. However, by a rotation of the sphere we can arrange for any two points to lie on the equator, so we may set $c = 1$ without loss of generality. The $c = 1$ solution is really two solutions because any two points on the equator are connected by two equatorial arcs. Unless the two endpoints are antipodal one arc is shorter than the other. Which, if any, of these two equatorial arcs minimises $L[\theta]$?

To answer this question we look at the second derivatives. First we compute

$$\frac{\partial^2 f}{\partial\theta'\partial\theta'} = \frac{\sin^2\theta}{[\sqrt{(\theta')^2+\sin^2\theta}]^3} = \frac{1}{c^3\sin^4\theta} = 1\,,$$
$$\frac{\partial^2 f}{\partial\theta\partial\theta} = \frac{\cos 2\theta}{\sqrt{(\theta')^2+\sin^2\theta}} - \frac{\sin^2\theta\cos^2\theta}{[\sqrt{(\theta')^2+\sin^2\theta}]^3} = -\frac{1}{c\sin^2\theta} = -1\,,$$

where we use the fact that $c = 1 = \sin\theta$ and $\cos\theta = 0$ for our equatorial path. We also need

$$\frac{\partial^2 f}{\partial\theta'\partial\theta} = -\frac{\sin\theta\cos\theta\theta'}{[\sqrt{(\theta')^2+\sin^2\theta}]^3} = 0\,.$$

So, for this equatorial solution, the functions ρ and σ are extremely simple:

$$\rho(\phi) = 1\,, \qquad \sigma(\phi) = -1\,,$$

and the second variation is therefore

$$\delta^2 L[\theta_0,\xi] = \frac{1}{2}\int_{\phi_1}^{\phi_2} \left\{(\xi')^2 - \xi^2\right\} d\phi\,.$$

Is this positive? To answer this we have to look at the Jacobi accessory equation, which is

$$u'' + u = 0 \quad \Rightarrow \quad u \;\propto\; \sin\phi - \gamma\cos\phi\,,$$

for any constant γ, so $u = 0$ when

$$\tan\phi = \gamma\,.$$

Consider $\gamma = 0$; in this case $u = 0$ when $\sin\phi = 0$, i.e. when

$$\phi = 0\,,\ \pm\pi\,,\ \pm 2\pi\,,\ \cdots$$

Notice that the separation $\Delta\phi$ between zeros of u is π. By choosing another value for γ we can shift the position of any given zero of u to any value of ϕ that we wish, but the zeros will still have separation

$$\Delta\phi = \pi\,.$$

This has the following consequences:

(i) We *can* find a nowhere-zero solution to the Jacobi accessory equation when $\Delta\phi < \pi$. This implies that the second variation is positive when $\Delta\phi < \pi$, and hence that the shorter of the two solutions of the EL equations for the given b.c.s is a local minimiser of the length functional $L[\theta]$.

(ii) We *cannot* find a nowhere-zero solution to the Jacobi accessory equation when $\Delta\phi > \pi$. This does not prove that the the longer arc is not a local minimiser of the distance functional L. However, in this case it is easy to find a function ξ for which $\delta^2 L < 0$. Consider

$$\xi(\phi) = A\sin\left(\frac{(\phi-\phi_1)\pi}{\phi_2-\phi_1}\right)\,.$$

Substitution gives

$$\delta^2 L = \frac{A^2}{4(\phi_2-\phi_1)}\left[\pi^2 - (\phi_2-\phi_1)^2\right] < 0\,.$$

Corollary: Point (i) shows that the condition $\sigma > 0$ is not necessary for positivity of $\delta^2 F$. Point (ii) shows that the strong Legendre condition $\rho > 0$ is not sufficient for positivity of $\delta^2 F$.