Variational Principles

Carl Turner

August 16, 2012

Abstract

These are notes for an undergraduate course on variational principles; please send corrections, suggestions and notes to courses@suchideas.com The author's homepage for all courses may be found on his website at SuchIdeas.com, which is where updated and corrected versions of these notes can also be found.

Thanks go to Dr. D. M. A. Stuart for allowing me to use his Variational Principles course (Easter 2010) as the basis for these notes.

Contents

Prerequisites

A background in some amount of analysis, and a reasonable knowledge in vector calculus.

1 Motivating Problems and Ideas

The aim of this course is to provide a set of tools to address various new types of question. To get a sense for the type of problem to be addressed, we give two fairly simple examples of questions which can be methodically answered using techniques developed in this course.

Problem 1.1. Find the curve of shortest length joining two points in a plane.

This question is generally easily answered in Euclidean space - it is a straight line. However, how to develop a simple way of answering such a question may not be immediately obvious, especially if we then introduce some new distance metric onto the space.

Problem 1.2 (Dido's problem)**.** Find the curve $y = y(x)$ of some pre-specified length $l \in (2a, \pi a]$ such that $y(\pm a) = 0$ which gives the maximum area beneath the curve.

This question has something in common with the previous problem - we have fixed end-points, and a fairly complicated (integral) property to optimize by choosing an appropriate function. However, it is a classical Greek problem, and indeed admits comparatively simple methods to solve it, giving the answer to be the unique arc of a circle passing through $(\pm a, 0)$ which has the desired length l. But again we can imagine that a simple generalization or change of this problem might make it unapproachable using more elementary methods.

We will develop a systematic way of obtaining solutions to this type of problem by showing any solution must satisfy a specific differential equation. The underlying concept is analogous to the relationship in basic calculus between stationary points $\nabla f(\mathbf{x}) = \mathbf{0}$ and minimizing or maximizing $f(\mathbf{x})$, but instead of a normal vector $\mathbf{x}$ we have a function[1] y, and we want to have some entity $I[y]$ which acts on the function to give us the quantity to optimize.

Definition 1.3. A variable $I[y]$ which assigns a scalar to a function is called a *functional*. (Therefore, it is a special type of *operator*, something which assigns another vector to a vector.)

Example 1.4. In 1.2, the area is

$$A = \int_{-a}^{a} y(x)\,\mathrm{d}x$$

and the (fixed) length is

$$L = \int_{-a}^{a} \sqrt{1 + y'^2}\mathrm{d}x$$

Both $A[y]$ and $L[y]$ are *functionals*.

Remark. Note that we are allowed to use derivatives of the function in calculating the scalar - $L = L[y]$ is an integral involving y'. In general, we can perform any operation on the argument, so long as we agree only to apply them to functions for which the functional is defined.

[1] In fact, a function is just an element of a vector space of functions, so actually in some sense the function is a normal vector. However, such a space is *much larger* in some sense than the usual Euclidean vector spaces $\mathbb{R}^n$ or $\mathbb{C}^n$ (in fact, it is infinite dimensional) so there is a conceptual difference worth noting.

The class of problems arising from functional constraints and particularly functional quantities to optimize is called the *calculus of variations*, and forms a key part of this course.

To clarify what is meant by the above statement about using relationship between stationary points and extreme points, consider the following problem.

Problem 1.5. Show there exists a real number $x \in \mathbb{R}$ such that $x + x^9 = b$ for any $b \in \mathbb{R}$.

This problem could be addressed straightforwardly using methods from analysis, applying the intermediate value theorem to the continuous function $x + x^9$. However, we are interested in a more sophisticated approach which will come in useful in problems less tractable by basic analysis.

We construct the function $f(x) = \frac{x^2}{2} + \frac{x^{10}}{10} - bx$, so that $f'(x) = x + x^9 - b$; we want to show $f'(x) = 0$ for some x. We know that if we can find an extreme point of f, we are done.

But clearly $f \to +\infty$ as $x \to \pm\infty$, and on any bounded interval f attains its minimum, so therefore taking an interval such that $f(x) \geq 1$ outside, inside f must have its minimum at a stationary point, as $f(0) = 0 < 1$.

This is referred to as the *Direct Method* for variational problems.

2 Functions on Finite-Dimensional Real Spaces

2.1 Partial Derivatives

We begin by considering functions $f : \mathbb{R}^n \to \mathbb{R}$.

We write elements in the domain as $\mathbf{x} = \sum x_j \mathbf{e}_j = (x_1, \cdots, x_n)$, so that $\mathbf{e}_j = (0, \cdots, 1, \cdots, 0)$. We shall denote the norm of a vector by $\|\mathbf{x}\| = \left(\sum x_j^2\right)^{1/2}$.

> **Definition 2.1.** $L : \mathbb{R}^n \to \mathbb{R}$ is *linear* if $L(\alpha \mathbf{V} + \beta \mathbf{W}) = \alpha L(\mathbf{V}) + \beta L(\mathbf{W})$ for all $\alpha, \beta \in \mathbb{R}$ and all $\mathbf{V}, \mathbf{W} \in \mathbb{R}^n$.

It follows from this definition that

$$L(\mathbf{x}) = \sum x_i L(\mathbf{e}_j) = \sum L_j x_j = \mathbf{L} \cdot \mathbf{x}$$

where we define $\mathbf{L} = (L_1, \cdots, L_n) = (L(\mathbf{e}_1), \cdots, L(\mathbf{e}_n))$.

A function $f : \mathbb{R}^n \to \mathbb{R}$ is *differentiable* at $\mathbf{x}$ if it can be well approximated by a linear function L near $\mathbf{x}$ in the sense that

$$f(\mathbf{x} + \mathbf{v}) - f(\mathbf{x}) - L\mathbf{v} = o(\|\mathbf{v}\|)$$

or equivalently, $\forall \epsilon > 0\ \exists \delta > 0$ such that

$$0 < \|\mathbf{v}\| < \delta \implies |f(\mathbf{x} + \mathbf{v}) - f(\mathbf{x}) - L\mathbf{v}| < \epsilon \|\mathbf{v}\|$$

It is reasonably clear that in the case $n = 1$ this is identical to the ϵ -δ formulation of differentiability familiar from one-dimensional analysis.

Now we investigate partial derivatives - in the above definition, put $\mathbf{v} = t\mathbf{e}_j$. Then we obtain the one-dimensional case, and so if f is differentiable at $\mathbf{x}$ then

$$\lim_{t \to 0} \frac{f(\mathbf{x} + t\mathbf{e}_j) - f(\mathbf{x})}{t}$$

exists and is finite, and is equal to $L\mathbf{e}_j = L_j$ by the linearity of L.

Proposition 2.2.

(i) If f is differentiable at $\mathbf{x}$, then the partial derivatives $\frac{\partial f}{\partial x_j}$ exist, and the linear map L approximating f is

$$L = \left(\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n}\right) = \boldsymbol{\nabla} f(\mathbf{x})$$

(ii) If all partial derivatives exist, and are continuous on $\mathbb{R}^n$, then f is differentiable at each $\mathbf{x} \in \mathbb{R}^n$, and

$$L = \boldsymbol{\nabla} f(\mathbf{x})$$

We have established the first rule - the second rule is established in a multi-dimensional analysis course (Analysis II).

Remark. Continuity throughout $\mathbb{R}^n$ (or at least some suitable domain within it) is required for f to be totally differentiable. It is worth noting that for $f : \mathbb{R}^2 \to \mathbb{R}$ that, even if all directional derivatives exist, f is not necessarily differentiable, or even continuous: consider

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & x^2 + y^2 \neq 0 \\ 0 & x = y = 0 \end{cases}$$

for an example.

We will define the following useful notation:

> **Definition 2.3.** $C^1(\mathbb{R}^n; \mathbb{R})$ is the set of real valued, continuous functions on $\mathbb{R}^n$ all of whose partial derivatives are continuous on $\mathbb{R}^n$.
>
> Similarly, $C^r(\mathbb{R}^n; \mathbb{R})$ has continuous partial derivatives up to order r.

It is a well-known result which in some ways underlies the entirety of this course that *extrema* are always *critical points*, as expressed by the following lemma:

> **Lemma 2.4.** *If* $f(\mathbf{x}) \geq f(\mathbf{y})$ *for all* $\mathbf{y} \in \mathbb{R}^n$ then
>
> $$\boldsymbol{\nabla} f(\mathbf{x}) = 0$$
>
> whenever $f \in C^1$.

This is the first-order necessary condition for a global maximum. Minima have the same condition, whilst for a local extremum the qualification becomes 'for all $\mathbf{y}$ in some ball[2] about $\mathbf{x}$'.

For more advanced conditions, we turn to higher-order derivatives.

2.2 Second Order Conditions for Extrema

In one dimension, we are familiar with the idea that if the second derivative f'' is strictly positive at a stationary point, then that means that the slope is increasing in either direction, so the point is a minimum of f, whilst if $f'' < 0$, this is a maximum.

For the multidimensional case, it seems clear that if the slope is increasing is *all possible directions*, then the point is a minimum (similarly for maxima). But for a function $f : \mathbb{R}^n \to \mathbb{R}$, this is essentially saying that the matrix of second-order partial derivatives

$$\frac{\partial^2 f}{\partial x_i \partial x_j}$$

never reverses a vector which it acts on - i.e. when one moves away a small amount, f is increasing in the direction you went. This leads to the following definition:

[2] A ball about $\mathbf{y}$ of radius r is $B_{\mathbf{y}}(r) = \{\mathbf{a} \in \mathbb{R}^n : |\mathbf{y} - \mathbf{a}| < r\}$.

Definition 2.5. A real, $m \times m$ symmetric matrix A_{ij} is *positive definite*, $A > 0$, if

$$\mathbf{v}^T A \mathbf{v} = \sum_{i,j} A_{ij} v^i v^j > 0$$

for all vectors $\mathbf{v} \neq \mathbf{0}$ in $\mathbb{R}^m$; it is *positive semi-definite*, $A \geq 0$, if the inequality is not strict.

Negative (semi-)definite matrices are defined in much the same way.

Remark. The notation $\sum_{i,j} A_{ij} v^i v^j$ is essentially equivalent to writing $\sum_{i,j} A_{ij} v_i v_j$ - the significance of the superscripts is due to tensor properties called valence which need not concern us here.

It is worth noting that the generalization of this notion to complex spaces involves requiring $A = A^\dagger$ to be Hermitian, and taking $\mathbf{v}^\dagger A \mathbf{v} > 0$ etc. (These matrices in fact correspond to positive-definite symmetric bilinear or sesquilinear forms for the real and complex cases respectively.)

One very useful way of thinking about positive (semi-)definite matrices is in terms of their eigenvalues. It is left as an exercise to show the following:

Exercise 2.6. Show a symmetric matrix is positive semi-definite $\iff$ all its eigenvalues are greater than or equal to 0. Similarly, show A is positive definite $\iff$ all its eigenvalues are strictly positive.

Theorem 2.7. *If $f \in C^2(\mathbb{R}^n)$ and $\nabla f(\mathbf{x}) = \mathbf{0}$, then*

(i) *if $\mathbf{x}$ is a local minimum or maximum, then the matrix $A_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$ is positive semi-definite or negative semi-definite respectively.*

(ii) *if $A_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$ is positive or negative definite, then $\mathbf{x}$ is a strict local minimum or maximum respectively.*

Remark. A strict local minimum $\mathbf{x}_0$ is a point such that, in some sufficiently small open sphere (of strictly positive radius) around $\mathbf{x}_0$, there is no point such that f takes on even the same value.

Note that if $f \in C^2(\mathbb{R})$ is a function defined on the real line, and $f'(x_0) = 0$ and $f''(x_0) > 0$, then f has a *strict* local minimum at x_0. It follows, in fact, that if x_0 is the only stationary point, then it is the global minimum, by Rolle's Theorem[3].

In $\mathbb{R}^n$ for $n \geq 2$, however, there are in fact $C^2(\mathbb{R}^n)$ functions with only one stationary point which is a strict local minimum but *not* a global minimum.

2.3 Convexity

However, one class of functions does in fact have very nice properties in terms of determining global minima.

[3] Rolle's Theorem states that if a differentiable function $\mathbb{R} \to \mathbb{R}$ takes equal values at two points, its derivative is 0 at some intermediate point (Analysis I). So if $f(x_0) = f(y)$, then there is a stationary point in (x_0, y) - hence by continuity of f the function is either strictly larger or strictly smaller than $f(x_0)$ at all other points. The sign of $f''(x_0)$ then indicates whether x_0 is a global minimizer or maximizer (via an application of Taylor's Theorem).

Definition 2.8. A set $S \subset \mathbb{R}^n$ is *convex* if whenever $\mathbf{x}, \mathbf{y} \in S$, and $\theta \in (0,1)$,

$$\theta \mathbf{x} + (1-\theta)\mathbf{y} \in S$$

A function $f : \mathbb{R}^n \to \mathbb{R}$ is *convex* if similarly

$$f((1-\theta)\mathbf{x} + \theta \mathbf{y}) \leq (1-\theta) f(\mathbf{x}) + \theta f(\mathbf{y})$$

It is strictly *convex* if and only if this is a strict inequality.

It is important to note that $(1-\theta)\mathbf{x} + \theta\mathbf{y}$ must be in the domain of the function f for this definition to make any sense. Therefore, in the case of a function $f : D \to \mathbb{R}$ where $D \subseteq \mathbb{R}^n$, f can only be convex if the set D on which it is defined is a *convex set.*

Remark. The *epigraph* is the set of points which lie above the graph of the function, as shown in Figure 2.1; i.e.

$$E_f = \{(z, \mathbf{x}) : z \geq f(\mathbf{x})\} \subset \mathbb{R}^{1+n}$$

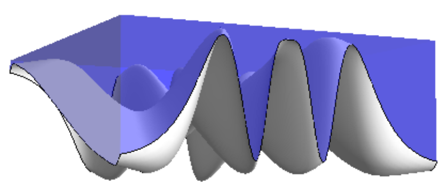

Figure 2.1: The epigraph of a (non-convex) function $f : \mathbb{R}^2 \to \mathbb{R}$

It can be shown that E_f is convex if and only if f is convex. One can also verify that if all of the line cuts (vertical cross-sections in the $\mathbb{R}^2 \to \mathbb{R}$ case shown) of the form $g(s) = f(\mathbf{x} + s\mathbf{v})$ are convex, then f is also convex.

A function f is *concave* if and only if $(-f)$ is convex.

Proposition 2.9. *If $f \in C^1(\mathbb{R}^n)$, then the following are equivalent:*

(i) *f is convex*

(ii) $f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x})$ *for all* $\mathbf{x}$ *and* $\mathbf{y}$

(iii) $[\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})] \cdot (\mathbf{x} - \mathbf{y}) \geq 0$

Proof.

$(i) \implies (ii)$: Let $H(t) = (1-t) f(\mathbf{x}) + t f(\mathbf{y}) - f((1-t)\mathbf{x} + t\mathbf{y}) \geq 0$. Note $H(0) = 0$, so $\dot{H}(0) \geq$

0. Then

$$\begin{aligned}\dot{H}(0) &= \lim_{t\to 0^+}\frac{H(t)-H(0)}{t}\\ &= -f(\mathbf{x})+f(\mathbf{y})-(\mathbf{y}-\mathbf{x})\cdot\boldsymbol{\nabla} f(\mathbf{x})\\ &\geq 0\end{aligned}$$

$(ii) \implies (i)$: We have

$$\begin{aligned}f(\mathbf{y}) &\geq f(\mathbf{z})+\boldsymbol{\nabla} f(\mathbf{z})\cdot(\mathbf{y}-\mathbf{z})\\ f(\mathbf{x}) &\geq f(\mathbf{z})+\boldsymbol{\nabla} f(\mathbf{z})\cdot(\mathbf{x}-\mathbf{z})\end{aligned}$$

and therefore

$$\begin{aligned}(1-t)f(\mathbf{y})+tf(\mathbf{x}) &\geq (1-t+t)f(\mathbf{z})+\boldsymbol{\nabla} f(\mathbf{z})\left[(1-t)(\mathbf{y}-\mathbf{z})+t(\mathbf{x}-\mathbf{z})\right]\\ &= f(\mathbf{z})\end{aligned}$$

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

$(ii) \implies (iii)$: Add

$$\begin{aligned}f(\mathbf{y}) &\geq f(\mathbf{x})+\boldsymbol{\nabla} f(\mathbf{x})\cdot[\mathbf{y}-\mathbf{x}]\\ f(\mathbf{x}) &\geq f(\mathbf{y})+\boldsymbol{\nabla} f(\mathbf{y})\cdot[\mathbf{x}-\mathbf{y}]\end{aligned}$$

$(iii) \implies (ii)$: Left as an exercise.

$\square$

It is perhaps worth developing some intuition about the latter two equivalent statements. The second states that the function always lies above all of its tangent planes; the third part is a generalization notion of the derivative being monotone non-decreasing.

Proposition 2.10. *f is convex $\iff$ f lies above all of its tangent planes $\iff$ $\boldsymbol{\nabla} f$ is 'monotone non-decreasing' in the generalized sense defined above (i.e. along any path). The same results hold with strict inequalities.*

Proof. Left as an exercise. (This is not totally immediate.) $\square$

Corollary 2.11. *If $f\in C^1(\mathbb{R}^n)$ with stationary point $\mathbf{x}$ then $\mathbf{x}$ is a global minimizer for f if f is convex.*

Proof. Follows directly from the second part. $\square$

Recalling the above, we saw $\boldsymbol{\nabla} f(\mathbf{x})=\mathbf{b}$ could possibly be solved by minimizing $f(\mathbf{x})-\mathbf{b}\cdot\mathbf{x}$.

Corollary 2.12. *If $f \in C^1$ is strictly convex, then*

$$\nabla f(\mathbf{x}) = \mathbf{b}$$

has at most one solution.

Proof. If there were two solutions, $\nabla f(\mathbf{x}) - \nabla f(\mathbf{y}) = \mathbf{0}$, a clear contradiction to the third part of the proposition. $\square$

All of the above can be trivially repeated for concave functions.

Lemma 2.13. *If $f \in C^2(\mathbb{R}^n)$ then*

(i) f is convex $\iff \frac{\partial^2 f}{\partial x_i \partial x_j} \geq 0$ for all $\mathbf{x}$

(ii) If $\frac{\partial^2 f}{\partial x_i \partial x_j} > 0$ for all $\mathbf{x} \implies f$ is strictly convex.

Remark. The implication in the latter statement cannot be reversed; consider $f(x, y) = x^4 + y^4$.

Proof.

(i) $\Longleftarrow$: First, we write

$$\begin{aligned} \nabla f(\mathbf{x}) - \nabla f(\mathbf{y}) &= [\nabla f(\mathbf{u})]_{\mathbf{y}}^{\mathbf{x}} \\ &= [\nabla f(\mathbf{y} + t(\mathbf{x} - \mathbf{y}))]_0^1 \\ &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} \nabla f(\mathbf{y} + t(\mathbf{x} - \mathbf{y}))\, \mathrm{d}t \end{aligned}$$

where we have used the fundamental theorem of calculus in the last step. Then, using the chain rule and the fact that the matrix of partial derivatives $\partial_{ij}^2 f \geq 0$, we have

$$\begin{aligned} [\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})] \cdot (\mathbf{x} - \mathbf{y}) &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} \nabla f(\mathbf{y} + t(\mathbf{x} - \mathbf{y})) \cdot (\mathbf{x} - \mathbf{y})\, \mathrm{d}t \\ &= \sum_i \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial}{\partial x_i} f(\mathbf{y} + t(\mathbf{x} - \mathbf{y}))(x_i - y_i)\, \mathrm{d}t \\ &= \sum_i \int_0^1 \frac{\partial^2}{\partial x_j \partial x_i} f(\mathbf{y} + t(\mathbf{x} - \mathbf{y}))(x_i - y_i)(x_j - y_j)\, \mathrm{d}t \\ &\geq 0 \end{aligned}$$

(i) $\Longrightarrow$: Exercise.

(ii) $\Longrightarrow$: Exercise.

$\square$

Example 2.14. Show that the entropy of the probability distribution $\mathbf{P} = (P_1, \cdots, P_n)$ given by

$$S(P_1, \cdots, P_n) = -\sum P_i \ln P_i$$

is concave, where $0 \le P_i \le 1$ for all i, and $\sum P_i = 1$.

Consider $(1-t)\,\mathbf{p} + t\mathbf{q}$, where $\mathbf{p}, \mathbf{q}$ are both probability distributions on $\{1, 2, \cdots, n\}$ and $t \in [0,1]$. This is also a probability distribution on the given set, since $(1-t)\,p_i + tq_i \in [0,1]$ and

$$\sum_i [(1-t)\,p_i + tq_i] = (1-t) + t = 1$$

Hence S is defined on a convex set.

Then, calculating the Hessian at $\mathbf{p}$,

$$\frac{\partial^2 S}{\partial P_i \partial P_j} = \begin{pmatrix} \frac{\partial^2 S}{\partial P_1^s} & & \\ & \ddots & \\ & & \frac{\partial^2 S}{\partial P_n^s} \end{pmatrix} = \begin{pmatrix} -\frac{1}{P_1} & & \\ & \ddots & \\ & & -\frac{1}{P_n} \end{pmatrix}$$

noting $\frac{\partial}{\partial p}(-p \ln p) = -1 - \ln p$ and $\frac{\partial^2}{\partial p^2}(-p \ln p) = -\frac{1}{p} < 0$, so clearly all eigenvalues are negative, and the map is concave.

2.4 Constraints and Lagrange Multipliers

A common, fairly simple, problem arising in the field of variational principles but which demands a more advanced method than that taught at A-level is maximization subject to a constraint. We write the most simple case as a requirement to maximize some function $f(x,y)$ subject to the constraint

$$C = \{(x,y) : g(x,y) = 0\}$$

Example 2.15. Maximize $f(x,y) = x + y$ where the point (x,y) lies on the unit circle. We define $g(x,y) = x^2 + y^2 - 1$. This clearly has the maximum value $\frac{2}{\sqrt{2}} = \sqrt{2}$ at $x = y = \frac{1}{\sqrt{2}}$. Clearly, the derivative $\boldsymbol{\nabla} f = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is not zero. But notice $\boldsymbol{\nabla} f$ is perpendicular to the constraint line (i.e. the circle) at this point.

Similarly, if we attempted to maximize, say, $f(x,y) = y^2$, we would get maxima at $\begin{pmatrix} 0 \\ \pm 1 \end{pmatrix}$ and then $\boldsymbol{\nabla} f = \begin{pmatrix} 0 \\ 2y \end{pmatrix}$ is also perpendicular to the constraint line as this point. The minima, at $\begin{pmatrix} \pm 1 \\ 0 \end{pmatrix}$ would give $\boldsymbol{\nabla} f = \mathbf{0}$ as $f = 0$ here which is an unconstrained minimum - this is also trivially 'perpendicular' to the constraint line.

To see why this holds, parametrize C as $(x(t), y(t)) = (\cos t, \sin t)$. Then $\phi(t) = f(x(t), y(t)) = (\sin t)^2$.

At ϕ's maxima, then, $t = \frac{\pi}{2}, \frac{3\pi}{2}$, we must have $\frac{\mathrm{d}\phi}{\mathrm{d}t} = 0$. Applying the chain rule,

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = \boldsymbol{\nabla} f \cdot \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix} = 0$$

But $\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix}$ is precisely the tangent to C, so $\frac{\mathrm{d}\phi}{\mathrm{d}t} = 0$ is precisely equivalent to $\boldsymbol{\nabla} f$ being normal to C.

Two ways of interpreting this follow:

(i) If $\boldsymbol{\nabla} f$ is not perpendicular to the constraint line at $\mathbf{x}$, then there is a nearby point $\mathbf{x} + \delta\mathbf{x}$ in the constrained region, where $\delta\mathbf{x}$ has some positive component in the direction of $\boldsymbol{\nabla} f$, so $\boldsymbol{\nabla} f \cdot \delta\mathbf{x} > 0$. Then $f(\mathbf{x} + \delta\mathbf{x}) = f(\mathbf{x}) + \boldsymbol{\nabla} f \cdot \delta\mathbf{x} + O\left(\|\delta\mathbf{x}\|^2\right) > f(\mathbf{x})$ for all sufficiently small $\delta\mathbf{x}$, so $f(\mathbf{x})$ is not a local maximum. A similar argument applies to minima.

(ii) If one draws the constraint line $g = 0$ and then adds contours $f = \text{constant}$ gradually decreasing the constant from $+\infty$ to $-\infty$, the maximum value of f on the constraint line will first be achieved when the contour just touches (i.e. is tangent to) the line $g = 0$. But this is exactly equivalent to $\boldsymbol{\nabla} f$ being perpendicular to the line $g = 0$ at this point.

This in fact allows us to deduce a *first-order necessary condition for a stationary point.*

Theorem 2.16 (First-Order Necessary Condition)**.** *Let $f, g \in C^2(\mathbb{R}^n)$ and $\boldsymbol{\nabla} g(\mathbf{x}) \neq \mathbf{0}$ for all $\mathbf{x}$. Let the constraint set $C = \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) = 0\}$, which we assume admits some parametrization.*

Then if $f|_C$ has a maximum or minimum at $\mathbf{x}_0$,

$$\boldsymbol{\nabla}\left[f(\mathbf{x}) - \lambda g(\mathbf{x})\right]_{\mathbf{x}_0} = 0$$

for some λ.

Remark. The restriction that $\boldsymbol{\nabla} g(\mathbf{x}) \neq \mathbf{0}$ actually implies, via the inverse function theorem, that the set C is locally a hypersurface, and hence that it can be parametrized as will be assumed below. (In fact, we technically only require that the local extremum $\mathbf{x}_0$ is a *regular point* of the constraint.)

Proof. Give the constraint set C the parametrization $\mathbf{x} = \mathbf{v}(t_1, \cdots, t_s)$, then if we find a point

$$\mathbf{x}_0 = \mathbf{v}\left(t_1^0, \cdots, t_s^0\right)$$

where the C^1 function has (without loss of generality) a maximum on C, so that

$$f(\mathbf{x}_0) = \max_{\mathbf{x} \in C} f(\mathbf{x})$$

and then $\phi(t_1,\cdots,t_s)=f(\mathbf{v}(t_1,\cdots,t_s))$ has an unconstrained maximum at these $\left(t_1^0,\cdots,t_s^0\right)$. Hence

$$\begin{aligned} \left[\frac{\partial\phi}{\partial t_j}\right]_{\left(t_1^0,\cdots,t_s^0\right)} &= 0 \\ \nabla f(\mathbf{x}_0)\cdot\left[\frac{\partial\mathbf{v}}{\partial t_j}\right]_{\left(t_1^0,\cdots,t_s^0\right)} &= 0 \end{aligned}$$

This is precisely the statement that the gradient of f is orthogonal to all the tangent vectors of the constraint set (which is a hypersurface).

We can choose to think of this as stating that the derivative $\nabla f(\mathbf{x}_0)$ is parallel to $\nabla g(\mathbf{x}_0)$. In this case, we can find some λ such that

$$\begin{aligned} \nabla f(\mathbf{x}_0) &= \lambda\nabla g(\mathbf{x}_0) \\ \nabla f(\mathbf{x}_0)-\lambda\nabla g(\mathbf{x}_0) &= \mathbf{0} \\ \nabla h(\mathbf{x},\lambda) &= \mathbf{0} \end{aligned}$$

Here, λ is the Lagrange multiplier, and $h(\mathbf{x},\lambda)=f(\mathbf{x})-\lambda g(\mathbf{x})$ is the augmented (Lagrange) function - this new function h has stationary points at the constrained extrema of f. □

Remark. Note that $h=f$ everywhere on the constraint set.

We give another example, choosing one which may be solved with other methods for clarity.

Example 2.17. Find the rectangle inscribed in the unit circle with the largest possible area. Note that a rectangle inscribed in a circle is entirely specified by a single point on the circumference and a rotation. So we can without loss of generality consider a rectangle specified by a point (x,y), as shown in Figure 2.2.

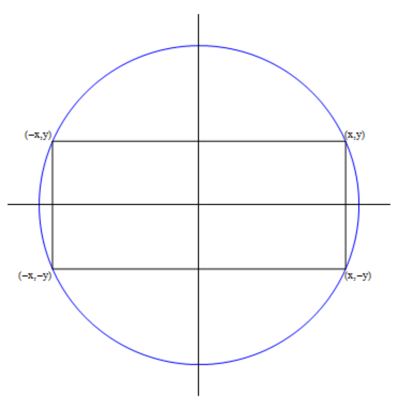

Figure 2.2: An example of a rectangle specified by the point (x,y).

We maximize (without loss of generality) the signed area $A=4xy$ respecting the constraint

$x^2 + y^2 - 1 = 0$. This has the augmented function

$$\begin{aligned} h(x, y, \lambda) &= A - \lambda g \\ &= 4xy - \lambda\left(x^2 + y^2 - 1\right) \\ \boldsymbol{\nabla} h &= \mathbf{0} \end{aligned}$$

Then

$$\begin{aligned} \frac{\partial h}{\partial x} &= 4y - 2\lambda x = 0 \\ \frac{\partial h}{\partial y} &= 4x - 2\lambda y = 0 \\ \frac{\partial h}{\partial \lambda} &= x^2 + y^2 - 1 = 0 \end{aligned}$$

with the last equation giving the constraint equation (as it always must). Then it follows that $y = \frac{1}{2}\lambda x$ and $x = \frac{1}{2}\lambda y$, so $\lambda = 2$.

Then we have $4y - 4x = 0$, so using the last equation we get $x = y = \pm\frac{1}{\sqrt{2}}$. (Note that $x = 0$, $y = 1$ and vice versa gives a minimum of the constrained f.)

Here is another example, using the idea of entropy from 2.14:

Example 2.18. Find the finite probability distribution with the highest entropy.

We wish to maximize $S(\mathbf{p}) = -\sum p_i \ln p_i$, subject to the constraint $\sum p_i = 1$ (although we also require $p \in [0, 1]$):

$$\begin{aligned} h &= -\sum p_i \ln p_i - \lambda\left(\sum p_i - 1\right) \\ \frac{\partial h}{\partial p_i} &= -\ln p_i - 1 - \lambda \\ &= 0 \end{aligned}$$

from which it follows that any stationary point (with arbitrary p_i) is located at $p_1 = p_2 = \cdots = p_n$. This gives the solution $p_i = \frac{1}{n}$.

Since S is convex, we may expect this to be a maximum, and in fact it is - but this is not in general necessarily true.

We can consider necessary conditions for maxima and minima too:

Theorem 2.19 (Second-Order Necessary and Sufficient Conditions)**.** *If the restricted function $f|_C$ has an extremum at $\mathbf{x}_0$, and $f, g \in C^2$*

(i) *the Hessian*

$$H_{ij} = \left[\frac{\partial^2 h}{\partial x_i \partial x_j}\right]_{\mathbf{x}_0} = \left[\frac{\partial^2 (f - \lambda g)}{\partial x_i \partial x_j}\right]_{\mathbf{x}_0}$$

is negative semi-definite on the tangent space[4] *at a maximum and positive semi-definite on the tangent space* $(H_{ij} \geq 0)$ *at a minimum;*

(ii) *if* H_{ij} *is negative definite* on the tangent space*, then* $\mathbf{x}_0$ *is a strict local maximum, and if* H_{ij} *is positive definite* on the tangent space*, then* $\mathbf{x}_0$ *is a strict local minimum.*

Proof. Note:

$$\begin{aligned}
\frac{\partial^2 \phi}{\partial t_j \partial t_i} &= \frac{\partial}{\partial t_j}\left(\boldsymbol{\nabla} f\left(\mathbf{v}\left(\mathbf{t}\right)\right) \cdot \frac{\partial \mathbf{v}}{\partial t_i}\right) \\
&= \boldsymbol{\nabla} f\left(\mathbf{v}\left(\mathbf{t}\right)\right) \cdot \frac{\partial^2 \mathbf{v}}{\partial t_j \partial t_i} + \frac{\partial}{\partial t_j}\left(\frac{\partial f}{\partial x_k}\right)\frac{\partial v_k}{\partial t_i} \\
&= \boldsymbol{\nabla} f\left(\mathbf{v}\left(\mathbf{t}\right)\right) \cdot \frac{\partial^2 \mathbf{v}}{\partial t_j \partial t_i} + \frac{\partial^2 f}{\partial x_l \partial x_k}\frac{\partial v_l}{\partial t_j}\frac{\partial v_k}{\partial t_i}
\end{aligned}$$

and assuming we are at a stationary point,

$$\frac{\partial^2 \phi}{\partial t_j \partial t_i} = \lambda \boldsymbol{\nabla} g \cdot \frac{\partial^2 \mathbf{v}}{\partial t_j \partial t_i} + \frac{\partial^2 f}{\partial x_l \partial x_k}\frac{\partial v_l}{\partial t_j}\frac{\partial v_k}{\partial t_i}$$

But $\mathbf{v}\left(t\right) \in C$, so we can differentiate the constraint to get 0:

$$\begin{aligned}
g\left(\mathbf{v}\left(\mathbf{t}\right)\right) &= 0 \\
\frac{\partial g\left(\mathbf{v}\left(t\right)\right)}{\partial x_k}\frac{\partial v_k}{\partial t_i} &= 0 \\
\frac{\partial g\left(\mathbf{v}\left(t\right)\right)}{\partial x_k}\frac{\partial^2 v_k}{\partial t_j \partial t_i} + \frac{\partial^2 g}{\partial x_l \partial x_k}\frac{\partial v_l}{\partial t_j}\frac{\partial v_k}{\partial t_i} &= 0
\end{aligned}$$

Hence we can write (for stationary points)

$$\begin{aligned}
\frac{\partial^2 \phi}{\partial t_j \partial t_i} &= \left(-\lambda \frac{\partial^2 g}{\partial x_l \partial x_k} + \frac{\partial^2 f}{\partial x_l \partial x_k}\right)\frac{\partial v_l}{\partial t_j}\frac{\partial v_k}{\partial t_i} \\
&= \frac{\partial^2 h}{\partial x_l \partial x_k}\frac{\partial v_l}{\partial t_j}\frac{\partial v_k}{\partial t_i}
\end{aligned}$$

The results then follow on application of the standard second-order tests for the function $\phi\left(t\right)$ - the left-hand side is the Hessian for ϕ in all of its parameters, and the RHS is the Hessian of h acting on vectors from the tangent space, the space of vectors of the form $\frac{\partial \mathbf{v}}{\partial t_i}$. $\square$

It is important to note that this result is different from the unconstrained version, precisely because of the restriction to the constraint-specified subspace, as one might expect.

[4] That is, $\mathbf{y}^T H \mathbf{y} \leq 0$ for vectors in the set $\{\mathbf{y} : \boldsymbol{\nabla} g\left(\mathbf{x}_0\right) \cdot \mathbf{y} = 0\}$, which is called the *tangent space* because all vectors in it are tangents to the constraint set. As an aside, we can note that tangent spaces can in fact be generalized to some other ('nice') metric spaces to begin the study of differential geometry.

Remark. Recall that we can formally test if a matrix is positive definite (and so on) by finding its eigenvalues - and its eigenvectors if we need to know what space they are acting on - and then checking that the relevant eigenvalues are strictly positive (and so on). A shortcut for real, symmetric matrices (or more generally Hermitian matrices) is *Sylvester's criterion*, which considers the signature formed by finding the signs of the determinants of the *principal minors* - the top-left 1×1 matrix, 2×2 matrix, and so on, up to the matrix itself. If the sequence is $++\cdots+$ then the matrix is positive definite; if it is $-+-+\cdots$ then it is negative definite.

Example 2.20. Recall the example of maximizing $x+y$ subject to $x^2+y^2=1$. Here, $h = x+y-\lambda\left(x^2+y^2-1\right)$ and so

$$\begin{aligned}\frac{\partial h}{\partial x} &= 1-2\lambda x\\ \frac{\partial h}{\partial y} &= 1-2\lambda y\\ \frac{\partial^2 h}{\partial x_i \partial x_j} &= \begin{pmatrix} -2\lambda & 0 \\ 0 & -2\lambda \end{pmatrix}\end{aligned}$$

The stationary points are at

$$(x,y,\lambda) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$$

which clearly demonstrates that the first is a maximum and the latter a minimum, since the matrix is diagonal and hence has its eigenvalues as the diagonal entries.

In general, of course, the matrix is not diagonal:

Example 2.21. Recall maximizing $A = 4xy$ subject to $x^2+y^2=1$. In this case, we have $h = 4xy - \lambda\left(x^2+y^2-1\right)$ and then

$$\begin{aligned}\frac{\partial h}{\partial x} &= 4y-2\lambda x\\ \frac{\partial h}{\partial y} &= 4x-2\lambda y\\ \frac{\partial^2 h}{\partial x_i \partial x_j} &= \begin{pmatrix} -2\lambda & 4 \\ 4 & -2\lambda \end{pmatrix}\end{aligned}$$

The stationary points we found to be

$$(x,y,\lambda) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 2\right), \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 2\right), (0,1,0), (1,0,0)$$

Then for the first two points, we have a matrix

$$\begin{pmatrix} -4 & 4 \\ 4 & -4 \end{pmatrix}$$

which has eigenvalues -8 for the eigenvector $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and 0 for $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. So the matrix is overall negative semi-definite. We could go to higher order in the latter direction - but there is no need. This zero is in the direction along which the value of the constraint equation changes - we could say this eigenvector lies outside of the relevant tangent space. Any change in x and y must be in the first direction, with eigenvalue -8. Hence the matrix is negative definite on the tangent space, and the point is a local maximum.

Note that in the example of maximizing entropy, we do in fact have $\frac{\partial^2 g}{\partial x_i \partial x_j} = 0$ because the constraint equation g has no 'mixed' terms, so the Hessian is the same as that for f, and the fact that f is concave implies that all stationary points are (local) maxima.

Corollary 2.22. *If the constraint equation g satisfies*

$$\frac{\partial^2 g}{\partial x_i \partial x_j} = 0$$

then stationary points of the constrained function f are of the same nature as would be determined by inspecting the Hessian for f.

2.5 Legendre Transforms

Transforms form a class of tools very frequently used, particularly by physicists, in order to recast a problem or piece of information in a new domain. There are various reasons for doing this, the main ones typically being that the new version of the problem is much easier to solve, or the new encoding of the information gives some intuitive (possibly physical) insight into its nature. For example, a *Fourier transform* can move from describing a signal shape (amplitude as a function of time, $f(t)$) to describing the component sinusoidal waves (amplitude as a function of the component frequency, $\hat{f}(\nu)$) - the same underlying set of information is encoded by both entities, $f(t)$ and $\hat{f}(\nu)$ (ignoring complications due to functions whose Fourier transform does not converge and so on), but it is represented differently. The Fourier transform, therefore, is useful when we are not particularly interested in the amplitude of the wave at any particular point, but are instead interested in the frequency with which components in the wave are oscillating.

The transform we are going to investigate here, however, is called the *Legendre transform*, and it is not concerned with a decomposition in the same way as the Fourier transform is. Instead, is is useful when it is for some reason preferable to think about the derivative of f than the variable x - again, note that the current independent variable x is considered to be of less interest than the new alternative independent variable $\mathrm{d}f/\mathrm{d}x$. Of course, if we want to have a one-to-one correspondence between the derivative and x, we need some special condition on f, which it seems natural would be concavity or convexity, as we need a (strictly) monotone derivative. It turns out that the way we define this map restricts this to a specific class of strictly convex functions.

Remark. We are talking functions of one variable here, but the Legendre transform can be easily generalized to higher dimensions, as we will see below.

2.5.1 Definition and discussion

There are several ways of approaching the definition of the Legendre transform, but the standard way is via the following (not very obvious) equation:

> **Definition 2.23.** Given a function $f : \mathbb{R} \to \mathbb{R}$, we define its *Legendre transform* $f^\star$ by
>
> $$\boxed{f^\star (p) = \sup_x [px - f(x)]}$$
>
> wherever this supremum exists.

So how do we arrive at this construction? There is not an immediate satisfactory explanation, but in this section we will describe a vague approach to deriving it. Do not worry if the discussion seems unclear, since it simply is.

Imagine for simplicity we have a strictly convex, twice-differentiable function $f(x)$ - in fact, we will eventually need $f''(x) > 0$ everywhere. Then the derivative is a strictly increasing function of x, which we will write

$$p(x) \equiv \frac{\mathrm{d}f}{\mathrm{d}x}$$

Then in this case, we can already parameterize f by p, because this differentiable, monotone function has an inverse, $x(p)$. We can write $g(p) = f(x(p))$, which encodes all information in f, but is parameterized by p, the derivative of f. But in practice, this is not the definition we use. There are several ways of justifying the different definition we use, the most natural of which is that this transform lacks any inherent symmetry.

If we apply the same process to $g(p)$, if this is all still valid, we find that

$$\begin{aligned} g'(p) &= \frac{\mathrm{d}}{\mathrm{d}p} f(x(p)) \\ &= x'(p) \cdot f'(x(p)) \\ &= x'(p) \cdot p \end{aligned}$$

which is not very elegant (even if we apply the inverse function theorem to $x'(p)$), and certainly does not return us to anything like a representation involving x. So consider a new function

$$\begin{aligned} h(p) &= x(p)\,p - f(x(p)) \\ h'(p) &= x'(p)\,p + x(p) - g'(p) \\ &= x(p) \end{aligned}$$

which seems a much nicer result, since applying this process again to $h(p)$ we find that if $q(p) =$

$h'(p) = x(p)$ then

$$\begin{aligned} h(p(q)) &= h(p(x)) \\ &= xp(x) - f(x) \\ p(q)\,q - h(p(q)) &= p(x)\,x - [xp(x) - f(x)] \\ &= f(x) \end{aligned}$$

so that the Legendre transform is its own inverse in this case! (We have not confirmed this is valid in this case either; we will do that below.)

Remark. This property, where it holds, makes the Legendre transformation an *involution*. The symmetry is particularly manifest when we write

$$f(x) + f^\star(p) = xp$$

where it is understood that $x = x(p)$ or $p = p(x)$, since x and p are not independent (they are conjugate variables under the Legendre transform).

To see how to make the final step from $f^\star(p) = x(p)\,p - f(x(p))$ to $f^\star(p) = \sup_x [xp - f(x)]$, simply note that at the point $x(p)$,

$$\frac{\mathrm{d}}{\mathrm{d}x}[xp - f(x)] = p - f'(x) = p - p = 0$$

so the term under the sup has a stationary point - in fact, this is the unique stationary point if $f'(x) = p$ has only one solution, as is the case for strictly convex functions. Further, this stationary point must be a maximum, since $xp - f(x)$ is a concave function of x for fixed p: its second derivative is just $-f''(x) < 0$. So the sup is achieved at this point.

The advantage of phrasing the definition in terms of this supremum is chiefly that it allows an easy definition for arbitrary f, and that it can be modified in certain ways which we will not address here. We close this introductory section by stating the generalization to higher dimensions:

Definition 2.24. Given a function $f : \mathbb{R}^n \to \mathbb{R}$, we define its *Legendre transform* $f^\star$ by

$$f^\star(\mathbf{p}) = \sup_{\mathbf{x}} [\mathbf{p} \cdot \mathbf{x} - f(\mathbf{x})]$$

wherever this supremum exists.

2.5.2 Examples and properties

Example 2.25. Consider the function $y = f(x) = ax^2$ where $a > 0$. Its Legendre transform is

$$f^\star(p) = \sup_x \left[px - ax^2\right]$$

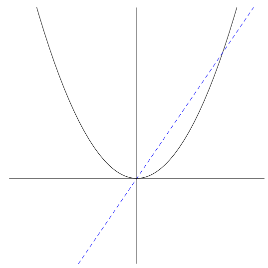

Figure 2.3: Transforming $f(x) = ax^2$

The term to be maximized is simply a quadratic (and in fact it is convex), and since $a > 0$ it has a well-defined maximum at

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[px - ax^2\right] = p - 2ax = 0$$

so that $x = p/2a$. (Note that this is the same as finding the largest distance by which the line $y = px$ lies above the quadratic $y = f(x)$, as shown in Figure 2.3.) It follows that the Legendre transform $f^\star$ is given by

$$f^\star(p) = \frac{p^2}{2a} - a \cdot \frac{p^2}{4a^2} = \frac{p^2}{4a}$$

which is another quadratic (and hence also convex). We can verify that

$$f^{\star\star}(y) = \sup_p \left[yp - \frac{p^2}{4a}\right] = ay^2$$

so $f^{\star\star} \equiv f$, as we expected.

We will prove this property a more formally than the above in Theorem 2.28 after a few examples of what can happen when f is not strictly convex:

Example 2.26.

(i) $f(x) = ax^2$ with $a < 0$, a convex function.

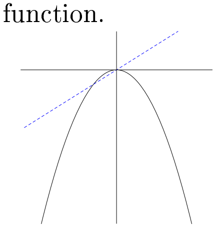

Figure 2.4: Transforming $f(x) = -|a|\, x^2$

In this case, $f^\star(p) = \sup_x \left[px - ax^2\right]$ is not defined for any p since the term in the brackets grows arbitrarily large as $x \to \pm\infty$. Hence the domain of $f^\star = \emptyset$.

(ii) $f(x) = 0$. Here, $\sup_x [px]$ exists if and only if $p = 0$, so $f^\star$ has the domain $\{0\}$.

(iii) More generally, if $f(x) = ax + b$ is any line, then $\sup_x [px - ax - b]$ is defined if and only if $p = a$; then $f^\star(a) = -b$.

This last example is probably the most revealing, in that it suggests a geometrical interpretation of $f^{\star}(p)$ as being -1 multiplied by the y-intercept of the tangent to the graph having the slope p. In fact, we will use this in the proof of Theorem 2.28.

But first, we prove the following proposition hinted at by our initial discussion:

Proposition 2.27. *$f^{\star}(p)$ is convex on its domain.*

Proof. We need the domain to be a convex set for this to be possible.

For any $t \in (0,1)$ and any x we have

$$t(p_1 x - f(x)) + (1-t)(p_2 x - f(x)) = (tp_1 + (1-t)p_2)x - f(x)$$

where the left-hand side is bounded above by

$$t \sup_x (p_1 x - f(x)) + (1-t) \sup_x (p_2 x - f(x)) = tf^{\star}(p_1) + (1-t) f^{\star}(p_2)$$

So we have

$$tf^{\star}(p_1) + (1-t) f^{\star}(p_2) \geq (tp_1 + (1-t)p_2)x - f(x)$$

and therefore if p_1 and p_2 lie in the domain, so does $tp_1 + (1-t)p_2$ because the right-hand side is bounded above. Hence the domain is a convex set.

Further, we can now take suprema to get

$$tf^{\star}(p_1) + (1-t) f^{\star}(p_2) \geq f^{\star}(tp_1 + (1-t)p_2)$$

which established the convexity of $f^{\star}$ on this set. □

We are now ready to prove the following result:

Theorem 2.28. *If $f \in C^2(\mathbb{R})$ with $f''(x) \geq c > 0$ - that is, f strictly convex with a non-zero lower bound on its second derivative - then $f^{\star\star} = f$.*

Proof. By the result in Corollary 2.12, the strict convexity of f implies that $f'(x) = p$ is satisfied by at most one x. Clearly, $f^{\star}(p)$ is defined for all p, because the expression $px - f(x)$ is concave with second derivative $-f''(x) \leq c < 0$, and is therefore bounded above.

It follows that we can define a function $X(p)$ defined uniquely by $f'(X(p)) = p$. So

$$\begin{aligned} f^{\star}(p) &= \sup [px - f(x)] \\ &= pX(p) - f(X(p)) \end{aligned}$$

Now we turn to the geometrical interpretation of $f^{\star}$. Consider, for some fixed p, the *unique* tangent line to $y = f(x)$ which has slope p. Its equation is

$$\begin{aligned} y - f(X(p)) &= p[x - X(p)] \\ y &= px - [pX(p) - f(X(p))] \\ &= px - f^\star(p) \end{aligned}$$

Recall that convex functions always lie *above* their tangent lines - so $f(z) \geq pz - f^\star(p)$ for any z, and equality is obtained at the point $z = X(p)$. But p is also arbitrary. Thus for a fixed z, $f(z) \geq pz - f^\star(p)$ for any p, and equality is obtained at the point $p = f'(z)$.

But then we are done, because $f^{\star\star}$ is defined, at some point z, by

$$\begin{aligned} f^{\star\star}(z) &= \sup_p [zp - f^\star(p)] \\ &= f(z) \end{aligned}$$

□

Remark. Note $f^\star(p)$ is precisely the negative of the y-intercept; also, by the above proposition, $f^\star(p)$ is a globally defined, convex function of p.

One interesting corollary of the above is that $f^\star(p)$ is C^1 (taking the supremum of functions does not in general preserve even continuity, let alone differentiability). This can be seen from from the geometrical nature of $f^\star(p)$.

You may be curious about the extra condition we used, that $f''(x) \geq c > 0$ for some constant c. This ensures that the solution goes to infinity at least as rapidly as any straight line as $x \to \pm\infty$, so that the supremum is always well-defined. Without this or a similar condition, we can easily find strictly convex functions with no Legendre transform:

Example 2.29. If $f(x) = e^x$ then

$$\sup_x [px - f(x)] = \sup_x [px - e^x]$$

is undefined if $p < 0$. This is because $px \to \infty$ as $x \to -\infty$ but $e^x \to 0$.

The argument from this theorem has the following corollary:

Corollary 2.30. *If f is convex (at least, with $f''(x) > c$ condition), it is the supremum of a family of affine functions (straight lines).*

This can also be expressed by saying it is the *envelope* of such a family - a shape which is tangent to all elements of the collection.

One way of expressing the symmetry of the Legendre transform is to say that (generally convex) functions f and g are *dual* (in the sense of Young) when they are Legendre transformations of each other. Then

$$g(\mathbf{p}) = \sup_{\mathbf{x}} [\mathbf{p} \cdot \mathbf{x} - f(\mathbf{x})] \geq \mathbf{p} \cdot \mathbf{x} - f(\mathbf{x})$$

for any $\mathbf{x}$. From this we can deduce the (*generalized*, in the case where we use vectors $\mathbf{p}$ and $\mathbf{x}$) *Young's inequality*:

$$f(\mathbf{x}) + g(\mathbf{p}) \geq \mathbf{p} \cdot \mathbf{x}$$

for any $\mathbf{x}$ and $\mathbf{p}$.

2.5.3 Physical applications

The first example we look at is of extreme importance in theoretical physics, and is of particular relevance to the formulation of quantum mechanics and quantum field theory.

Example 2.31. In simple cases of classical physical problems, we are used to working with forces, accelerations, velocities and positions. However, this formulation of physical laws is ultimately deeply tied to the coordinate system we choose, and does not generalize to quantum theory, and is not convenient for dealing with either special or general relativistic physics. Instead, we usually work with one of two alternative mathematical setups, called the *Lagrangian* and *Hamiltonian* formulations. The Lagrangian $\mathcal{L}$ is defined by

$$\begin{aligned} \mathcal{L} &= T - V \\ &= \text{kinetic energy} - \text{potential energy} \end{aligned}$$

For the classical case, we can write $T = T(\dot{\mathbf{x}}) = \frac{1}{2}m\dot{\mathbf{x}} \cdot \dot{\mathbf{x}}$ for the kinetic energy, and $V = V(\mathbf{x})$ for the potential energy. Hence we have

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

The Lagrangian, more generally, can be a function of any *generalized* coordinates q_i, including angles, or field strength, or so on - here, we shall just use position $\mathbf{x}$ and its derivative for simplicity.

We shall see the relevance of the Lagrangian to physics in section 3.6.2, when we see the example of an *action principle* from which equations of motion can be deduced. What concerns us here is the relationship between this formalism and the Hamiltonian one. The Legendre transform with respect to $\dot{\mathbf{x}}$ of the Lagrangian is

$$L^{\star}(\mathbf{x}, \mathbf{p}) = \sup_{\dot{\mathbf{x}}} [\mathbf{p} \cdot \dot{\mathbf{x}} - L(\mathbf{x}, \dot{\mathbf{x}})]$$

To calculate this, note that the supremum is achieved at the point when all following the partial derivatives with respect to $\dot{\mathbf{x}}$ vanish:

$$\frac{\partial}{\partial \dot{x}_j} [\mathbf{p} \cdot \dot{\mathbf{x}} - L(\mathbf{x}, \dot{\mathbf{x}})] = p_j - m\dot{x}_j = 0$$

Hence $\mathbf{p} = m\dot{\mathbf{x}}$ (that is, the classical momentum) and the transform is given by

$$\begin{aligned} L^{\star}(\mathbf{x},\mathbf{p}) &= \mathbf{p}\cdot\frac{\mathbf{p}}{m} - \left[\frac{1}{2m}\mathbf{p}\cdot\mathbf{p} - V(\mathbf{x})\right] \\ &= \frac{1}{2m}\mathbf{p}\cdot\mathbf{p} + V(\mathbf{x}) \end{aligned}$$

We can then define the Hamiltonian to be the Legendre transformation of the Lagrangian:

$$\begin{aligned} \mathcal{H}(\mathbf{x},\mathbf{p}) &= L^{\star}(\mathbf{x},\mathbf{p}) \\ &= \frac{1}{2m}\mathbf{p}\cdot\mathbf{p} + V(\mathbf{x}) \\ &= T + V \\ &= \text{kinetic energy} + \text{potential energy} \end{aligned}$$

Note that we write the Hamiltonian as a function of a generalized coordinate and its so-called *conjugate momentum* p_i. It is easy to show that Newton's equations fall out naturally from the Hamiltonian in the form of the rules

$$\dot{x}_j = \frac{\partial H}{\partial p_j} \quad \text{and} \quad \dot{p}_j = -\frac{\partial H}{\partial x_j}$$

and doing this is left as an exercise.

Remark. These last two equations are called the *Hamilton equations*, and they can in fact be derived from Lagrange's equations (though they are sometimes viewed as more fundamental).

The second example is a key application in thermodynamics.

Example 2.32. In thermodynamics, we often assume that we have a gas (with a fixed number of particles N) governed by its internal energy

$$U = U(S,V)$$

where V is the volume it occupies, and S is its entropy[5]. It is helpful to think of the gas as occupying a perfectly sealed piston with adjustable volume.

There are several formulae associated with this formulation of thermodynamics - the underlying definitions are

$$\begin{aligned} \text{Heat flow } = \mathrm{d}q &= T\mathrm{d}S \\ \text{Energy change } = \mathrm{d}U &= \text{Heat flow} - \text{Mechanical work done on piston} \\ &= T\mathrm{d}S - p\mathrm{d}V \\ &= \left.\frac{\partial U}{\partial S}\right|_V \mathrm{d}S + \left.\frac{\partial U}{\partial V}\right|_S \mathrm{d}V \end{aligned}$$

where we have

$$\begin{aligned} T &= \left.\frac{\partial U}{\partial S}\right|_V \\ p &= -\left.\frac{\partial U}{\partial V}\right|_S \end{aligned}$$

From these, we can derive one of the so-called Maxwell relations:

$$\left.\frac{\partial T}{\partial V}\right|_S = -\left.\frac{\partial p}{\partial S}\right|_V$$

Now if the system is immersed in a constant temperature reservoir instead, the system is best described not by internal energy U but by the so-called *(Helmholtz) free energy*

$$F = F(T, V) = \inf_S \left[U(S, V) - ST\right]$$

which is the negative Legendre transform with respect to entropy of the fundamental energy $U(S, V)$.

The infimum is attained where the partial derivative with respect to S of the expression in brackets is zero - that is, at the S such that

$$T = \left.\frac{\partial U}{\partial S}\right|_V$$

This defines $S = S(T, V)$ and hence we can substitute back to find

$$\begin{aligned} F(T, V) &= U(S(T, V), V) - TS(T, V) \\ \mathrm{d}F &= \mathrm{d}U - T\mathrm{d}S - S\mathrm{d}T \\ &= (T\mathrm{d}S - p\mathrm{d}V) - T\mathrm{d}S - S\mathrm{d}T \\ &= -p\mathrm{d}V - S\mathrm{d}T \end{aligned}$$

Hence in the Helmholtz description,

$$\begin{aligned} p &= -\left.\frac{\partial F}{\partial V}\right|_T \\ S &= -\left.\frac{\partial F}{\partial T}\right|_V \end{aligned}$$

The associated Maxwell relation is

$$\left.\frac{\partial p}{\partial T}\right|_V = \left.\frac{\partial S}{\partial V}\right|_T$$

Note that $S = S(T, V)$ is determined *implicitly* by

$$T = \left.\frac{\partial U}{\partial S}\right|_V$$

which determines S uniquely where

$$\left.\frac{\partial^2 U}{\partial S^2}\right|_V > 0$$

But note that the constant volume heat capacity c_V, the heat needed to raise the temperature by one unit at the fixed volume V, is given by

$$c_V = T\left.\frac{\partial S}{\partial T}\right|_V = \frac{T}{\left.\frac{\partial T}{\partial S}\right|_V} = \frac{T}{\left.\frac{\partial^2 U}{\partial S^2}\right|_V}$$

so U is convex with respect to $S \iff$ we need heat input to raise the temperature, establishing the validity of our result.

In general, the Legendre transform is used to change between thermodynamic potentials.

[5]Entropy is one of the hardest quantities to give an intuitive, mathematical definition for. Classically, we define it as a property of a system moving between thermodynamic equilibria: in any process where energy ΔE is surrendered, and its entropy falls by ΔS, at least $T_0 \Delta S$ of the energy passed on will be passed directly to the environment (which is at the temperature T_0) without being used. In statistical thermodynamics, it is a measure of how uncertain the state of the gas particles is after the macroscopic properties (like temperature, pressure and volume) have been taken into account: $S = -k_B \sum_i p_i \log p_i$ where k_B is the Boltzmann constant, and the sum is over all states which the system has a probability p_i of lying in. These are equivalent notions. In an idealized (reversible) change, we can consider entropy as corresponding to *heat* loss, hence the relationship $\mathrm{d}q = T\mathrm{d}S$.

3 Calculus of Variations

In this section, we are going to address the other type of problem we discussed in the introduction: finding not simply a point, but a function which maximizes or minimizes some property. In order to do this, we need to have some way of assigning a single value to a function. Maps in this general class are called *functionals*:

> **Definition 3.1.** A *functional* is a map $V \to \mathbb{R}$ or $\mathbb{C}$ where V is a space of functions.

We will work only with the real case here.

3.1 Examples and Functional Derivatives

Examples of this type of map abound; two classes of example follow:

Example 3.2.

(i) $V = C(\mathbb{R})$, the space of continuous functions of $\mathbb{R} \to \mathbb{R}$. We might consider the 'Dirac functional' operating at x_0 by the map

$$\delta_{x_0} : f \mapsto f(x_0) \in \mathbb{R}$$

(ii) $V = \{f \in C^\infty : f(x + 2\pi) = f(x) \quad \forall x\}$, the space of smooth, 2π-periodic functions, like $\sin x$. All functions in V are integrable, because they are continuous, so we can define

$$I_0[f] = \int_0^{2\pi} [f(x)]^2 \, \mathrm{d}x$$

In fact, since all derivatives of smooth functions are continuous, we can define further functionals like

$$I_1[f] = \int_0^{2\pi} \left([f(x)]^2 + [f'(x)]^2\right) \mathrm{d}x$$

Now when we went about finding extrema of a function $h(\mathbf{x})$ in a finite dimensional vector space previously, we hit upon the idea of checking that all *directional derivatives* were 0, so that the point $\mathbf{x}$ was a stationary point. This meant picking a vector $\mathbf{v}$ in the space, and seeing that the restricted function $h_{\mathbf{v}}(t) = h(\mathbf{x}_0 + t\mathbf{v})$ was stationary at $t = 0$. Can we generalize this?

The answer is yes, in the most natural way possible - remember that a function space can also be a vector space, albeit one of infinite dimension, so long as it obeys the basic axioms. If we have a functional $I[f]$, then we want to investigate $I[f + t\phi]$ where $\phi(x)$ is the direction along which we take the derivative - it is a vector in our space, which makes it another function. What this amounts to is considering small *variations* made to f, and seeing what happens as $|t|$ grows from 0. If one of the gradients

$$\frac{\mathrm{d}}{\mathrm{d}t} I[f + t\phi]$$

is not zero, then f cannot be a local minimum or maximum for I.

Example 3.3. Consider $I_0[f]$ defined above. Then we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} I_0[f+t\phi] &= \frac{\mathrm{d}}{\mathrm{d}t}\int_0^{2\pi} [f(x)+t\phi(x)]^2 \,\mathrm{d}x \\ &= \int_0^{2\pi} \frac{\mathrm{d}}{\mathrm{d}t}[f(x)+t\phi(x)]^2 \,\mathrm{d}x \\ &= \int_0^{2\pi} 2\phi(x)[f(x)+t\phi(x)] \,\mathrm{d}x\end{aligned}$$

where we have used the fact that smooth integrands allow differentiation through the integral - note that $\phi(x)$ must be smooth, since the variation functions we are considering lie inside the vector space. Then the derivative at $t=0$ is

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} I_0[f+t\phi] = \int_0^{2\pi} 2\phi f \,\mathrm{d}x$$

We write

$$D_\phi I_0[f] = \int_0^{2\pi} 2\phi f \,\mathrm{d}x$$

This quantity must be zero for all ϕ satisfying the conditions of the vector space, just as in the case of finite-dimensional vector spaces, when $\mathbf{v}\cdot\nabla h(\mathbf{x})$ had to be zero for all $\mathbf{v}$ in the space for $\mathbf{x}$ to be the location of a minimum or maximum. In that case, it was easy to deduce from this that the gradient $\nabla h = \mathbf{0}$ at extrema, which meant that we could just check the simple condition that $\mathbf{x}$ was a *stationary point* as the first stage in locating these extrema. We need a way of expressing some kind of 'gradient' for our functional.

Example 3.4. To obtain a generalization of the gradient $\nabla h(\mathbf{x})$, we must first define an inner product (that is, a generalization of the 'dot' product) on our specific vector space of functions - we can do this by

$$\langle f,g\rangle = \int_0^{2\pi} f(x)\,g(x)\,\mathrm{d}x$$

where, for complex-valued functions, we would generalize this to

$$\langle f,g\rangle = \int_0^{2\pi} \overline{f(x)}g(x)\,\mathrm{d}x$$

(or the real part thereof).

Then we can write this very concisely as

$$D_\phi I_0[f] = \langle 2f,\phi\rangle$$

and we can therefore replace the idea of the gradient $\boldsymbol{\nabla} h(\mathbf{x})$ with the idea of a *functional derivative*, denoted

$$\boxed{\frac{\delta I_0}{\delta f} = 2f}$$

In general, we define the functional derivative in exactly this manner, though the definition of the inner product $\langle \cdot, \cdot \rangle$ may vary according to which space we are working in:

Definition 3.5. For a functional $I[f]$,

$$D_\phi I[f] \equiv \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} I[f + t\phi] \equiv \left\langle \frac{\delta I}{\delta f}, \phi \right\rangle$$

where such a function $\delta I / \delta f$ exists - it is called the *functional derivative* of $I[f]$.

Remark. Any inner product space over $\mathbb{R}$ or $\mathbb{C}$ is a *metric space* - if the space is complete (so any Cauchy sequence of elements converge to a point in the space), it is a *Hilbert space*. Hilbert spaces have the property that any linear, continuous L map $\mathbf{x} \mapsto L(\mathbf{x})$ from the space to $\mathbb{R}$ or $\mathbb{C}$ has a corresponding constant $\mathbf{y}$ (technically from the *dual space*) such that $\langle \mathbf{y}, \mathbf{x} \rangle = L(\mathbf{x})$. Since $D_\phi I[f]$ is clearly a linear functional of ϕ.

The inner product definitions we will work with will be of the form

$$\left\langle \frac{\delta I}{\delta f}, \phi \right\rangle \equiv \int \frac{\delta I}{\delta f} \phi \mathrm{d}x$$

where the integral is carried out over some suitable range.

It should not be surprise that not all functionals have such a representation:

Example 3.6. Consider δ_{x_0}, the Dirac functional which extracts the value of a function at x_0:

$$\begin{aligned}
D_\phi \delta_{x_0}[f] &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \delta_{x_0}[f + t\phi] \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (f(x_0) + t\phi(x_0)) \\
&= \phi(x_0) \\
&= \delta_{x_0}[\phi]
\end{aligned}$$

Now formally, we cannot write

$$\int \frac{\delta(\delta_{x_0})}{\delta f} \phi \, \mathrm{d}x = \phi(x_0)$$

for any true function $\delta(\delta_{x_0}) / \delta f$ - in particular, no function in the same function space that the integral inner product is defined on. However, if we adopt the notation of the Dirac delta function, so that $\int \delta(x-x_0)\, g(x)\, \mathrm{d}x = g(x_0)$ for suitable intervals of intergration, then we can write

$$\frac{\delta(\delta_{x_0})}{\delta f} \equiv \delta(x-x_0)$$

It may seem like functionals consisting of integrals of derivatives of f, like $I_1[f] = \int \left[f^2 + (f')^2\right] \mathrm{d}x$, should not have such a representation is this form, since the directional derivative would appear to necessarily involve derivatives of ϕ. However, importantly, this is not in fact the case:

Example 3.7. Recall the functional

$$I_1[f] = \int_0^{2\pi} \left([f(x)]^2 + [f'(x)]^2\right) \mathrm{d}x$$

defined on the space of smooth, 2π-periodic functions. Then we have

$$\begin{aligned}
D_\phi I_1[f] &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} I_1[f+t\phi] \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \int_0^{2\pi} \left([f+t\phi]^2 + [f'+t\phi']^2\right) \mathrm{d}x \\
&= \left[\int_0^{2\pi} (2f\phi + 2t\phi + 2f'\phi' + 2t\phi')\, \mathrm{d}x\right]_{t=0} \\
&= \int_0^{2\pi} (2f\phi + 2f'\phi')\, \mathrm{d}x
\end{aligned}$$

which currently does involve ϕ'. However, we can eliminate this by integration by parts:

$$\begin{aligned}
\int_0^{2\pi} f'\phi' \mathrm{d}x &= [f'\phi]_0^{2\pi} - \int_0^{2\pi} f''\phi \mathrm{d}x \\
&= -\int_0^{2\pi} f''\phi \mathrm{d}x
\end{aligned}$$

because, by periodicity, $f'(2\pi)\,\phi(2\pi) = f'(0)\,\phi(0)$, so the boundary terms vanish. This gives us the following expression for the directional derivative:

$$\begin{aligned}
D_\phi I_1[f] &= \int_0^{2\pi} (2f\phi - 2f''\phi)\, \mathrm{d}x \\
&= \int_0^{2\pi} (2f - 2f'')\, \phi \mathrm{d}x
\end{aligned}$$

and it follows that

$$\frac{\delta I_1}{\delta f} = -2f'' + 2f$$

This is typical of how we work with functionals involving integration of derivatives: we eliminate derivatives of ϕ via integration by parts, using boundary conditions or periodicity in order to restore the purely integral form form of the operator - this generally increases the order of the resulting functional derivative; that is, we obtain higher order derivatives of f, as in this case, when we obtained a term in f''.

3.2 Euler-Lagrange Equations

In fact, we can derive a much more general rule for integral operators like this. (Note that we now use the standard notation of $y(x)$ for solution curves, reserving $f(x, y, y')$ for the integrand.)

> **Lemma 3.8.** *Let $V = \left\{y(x) \in C^2[a,b] : y(a) = \alpha, y(b) = \beta\right\}$ be a space of twice-differentiable functions on $[a,b]$ with fixed endpoints, and let the functional $I : V \to \mathbb{R}$ be given by*
>
> $$I[y] = \int_a^b f\left(x, y, \frac{\mathrm{d}y}{\mathrm{d}x}\right)\mathrm{d}x = \int_a^b f(x, y, y')\,\mathrm{d}x$$
>
> *where $f(x, y, y')$ has continuous first partial derivatives with respect to each of its three arguments. Then the functional derivative is given by*
>
> $$\frac{\delta I}{\delta y} = \frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right)$$

Proof. This is a straightforward application of the same approach that we saw above, with the slight change that our variation functions $\phi(x)$, which still being in $C^2[a,b]$, must have

$$\phi(a) = \phi(b) = 0$$

so that the function $y + t\phi \in V$. Then we have

$$\begin{aligned} D_\phi I[y] &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} I[y + t\phi] \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \int_a^b f(x, y + t\phi, y' + t\phi')\,\mathrm{d}x \end{aligned}$$

Now the properties of f we required mean that we can exchange differentiation and integration[6], so that by the chain rule we have

$$D_\phi I[y] = \int_a^b \left[\phi\frac{\partial f}{\partial y}(x, y, y') + \phi'\frac{\partial f}{\partial y'}(x, y, y')\right]\mathrm{d}x$$

Then integrating the last term by parts, we have

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

where the boundary terms now vanish because $\phi(a) = \phi(b) = 0$. So finally,

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

which we can also write as

$$\frac{\delta I}{\delta y} = f_y - f_{y'x} - f_{y'y}y' - f_{y'y'}y''$$

□

So the question is: how can we use these functional derivatives? By analogy with the finite-dimensional case, where $\boldsymbol{\nabla} h(\mathbf{x}) = \mathbf{0}$, it seems that the fact that the inner product of the functional derivative with any suitable ϕ in some vector space (which may or may not be the same as that which f lies in) vanishes could mean that $\delta I/\delta y = 0$, giving us a differential equation for y. In order to establish such a rule, we need a lemma very much the following:

Lemma 3.9 (Fundamental lemma of the calculus of variations)**.** *If*

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

for all smooth functions $\phi(x)$ with

$$\phi(x) = 0 \quad \textit{for} \quad x \notin [c,d] \in (a,b)$$

and f is continuous, $f \in C[a,b]$, then $f \equiv 0$ in the interval $[a,b]$.

The idea here is that if we can find a general smooth function like that shown in Figure 3.1, which vanishes outside some arbitrary subinterval of $[a,b]$, and is strictly positive inside it, then by moving and scaling this shape, we can show that f cannot be non-zero. This is because, by continuity, it would follow that there was some interval where $f > 0$ or $f < 0$, and then multiplying this by our carefully chosen function ϕ, we would get a strictly positive or negative result.

Proof. Assume that there is some x_0 such that $f(x_0) = \theta \neq 0$. Take $\phi > 0$ without loss of generality, noting that otherwise we can simply consider $-f$.

Then by the continuity of f, there is some $\epsilon > 0$ such that $|f(x) - \theta| < \theta/2$ for all x with $|x - x_0| < \epsilon$, so that

$$f(x) \geq \frac{\theta}{2} \quad \text{for } x \in (x_0 - \epsilon, x_0 + \epsilon)$$

[6] The derivative of the integrand is continuous, by assumption, and the interval $[a,b]$ is closed and bounded and therefore compact - then since continuous functions on compact sets are uniformly continuous, the derivative of the integrand is uniformly continuous. Therefore, we can interchange the two limiting operations. (This is the Leibniz integral rule.)

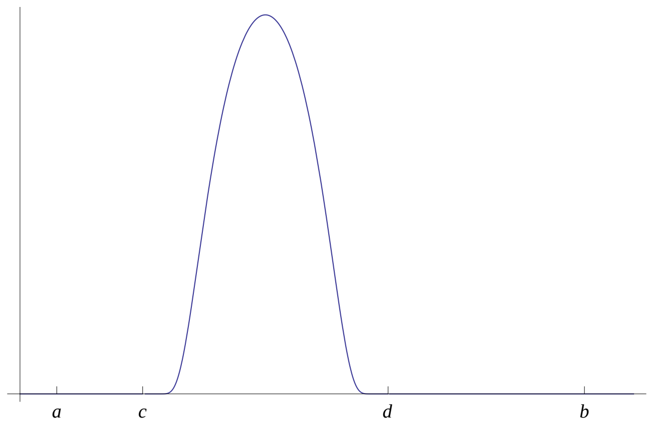

Figure 3.1: A bump function

Now consider the function

$$\psi(x) = \begin{cases} e^{-1/\left(x^2-1\right)^2} & x^2 < 1 \\ 0 & x^2 \geq 1 \end{cases}$$

which is motivated by recalling that the function $\upsilon(x) = e^{-1/x^2}$ has all derivatives tending to 0 as $x \to 0$, and then multiplying $\nu(x-1)\,\nu(x+1)$ to form a function which is positive in $(-1, 1)$ but which can be smoothly joined to the function which is constantly 0 at $x = \pm 1$. It can therefore be shown that this piecewise function is smooth by checking[7] that

$$\lim_{x \to 0} \nu^{(n)}(x) = 0$$

for all n, as it follows that the same holds for $\psi(x)$:

$$\lim_{x \to \pm 1, x^2 < 1} \psi^{(n)}(x) = 0$$

Now consider the function

$$\phi(x) = \psi\left(\frac{x - x_0}{\epsilon}\right) = \begin{cases} 0 & (x - x_0)^2 \geq \epsilon^2 \\ \text{strictly positive} & (x - x_0)^2 < \epsilon^2 \end{cases}$$

It is clear that $\phi(x)$ satisfies the conditions in the statement of the lemma, and hence we just note that

$$\int \phi f \, \mathrm{d}x = \int_{x_0 - \epsilon}^{x_0 + \epsilon} \phi f \, \mathrm{d}x \geq \frac{\theta}{2} \int \phi > 0$$

which is a contradiction. Hence $f(x_0) = 0$. $\square$

[7]This can be done by simply calculating the derivatives of ν, and using the fact that $\lim_{z \to \infty} z^N e^{-z} = 0$ for any N.

Remark. We can restate the conditions on ϕ by defining the *support* of a function

$$\operatorname{supp}\phi = \operatorname{cl}\{x : \phi(x) \neq 0\}$$

as the *closure* of the set of points where ϕ is non-zero. Then we say ϕ is *properly supported* in $[a, b]$ or (a, b) if

$$\operatorname{supp}\phi \subset [c, d] \subset (a, b)$$

for some c and d. Hence the lemma requires that $\int f\phi \,\mathrm{d}x$ over this interval vanishes for all smooth functions ϕ which are properly supported in (a, b).

Note also that with the strictly *weaker* requirement that $f \in C^k$ and that the integral vanish for all C^k functions $\phi(x)$ with $\phi(a) = \phi(b) = 0$, we could simply take $\phi = -(x-a)(x-b) f$, which satisfies all the necessary conditions, so that

$$\int_a^b \phi f \,\mathrm{d}x = \int_a^b -(x-a)(x-b) f^2 \mathrm{d}x = 0$$

and since the integrand is non-negative, it must be identically zero. Thus $f \equiv 0$ in (a, b), and if $k \geq 1$, $f \equiv 0$ in $[a, b]$.

With this lemma, we are now ready to address all the problems we have seen before, according to the following method:

Solution. The *indirect method* for finding a minimizer (without loss of generality) goes as follows:

- **(i)** *Assume that a minimizer exists.* In our case, assume that there is a a minimizing function f exists for the functional $I[f]$ which is of the above form.
- **(ii)** *Obtain a necessary condition for such a minimizer.* Here, we now have a differential equation for y, since the fundamental lemma implies that $\delta I/\delta y \equiv 0$, or

$$\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right) = 0$$

- **(iii)** *Show that there exists a solution satisfying this condition.* For us, this means solving the above differential equation for a function $y(x)$.
- **(iv)** *Show that the solution found is actually a minimizer.* In general, it is often clear whether or not the solution gives a minimum value for $I[f]$.

We can now apply the above method to some of the problems we originally wanted to study, using differential equations of the form we have deduced:

Definition 3.10. The *Euler-Lagrange equation* associated with a functional of the form $I[y] = \int_a^b f(x, y, y')\,\mathrm{d}x$ obeying the conditions described in Lemma 3.8 is

$$\boxed{\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right) = 0}$$

It is a clear consequence of the results of Lemma 3.8 and Lemma 3.9 that any admissible stationary point of the functional $I[y]$ must satisfy this differential equation. Hence, the above equation is a *necessary condition* for an extremal function. It is not, however, sufficient, which is why we must check if a solution to the Euler-Lagrange equation is actually the required function.

Example 3.11. Recall Problem 1.1, that of finding the shortest curve joining two points (a, α) and (b, β) in Euclidean space.

If we assume that the curve can be parametrized as $y = y(x)$, as a twice-differentiable function, then we can use the Euler-Lagrange equations on

$$I[y] = \int_a^b \sqrt{1 + y'^2} \mathrm{d}x$$

Since $f = \sqrt{1 + y'^2}$ we have

$$\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right) = -\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{y'}{\sqrt{1 + y'^2}}\right) = 0$$

This is in fact easy to solve; since f does not depend explicitly on y, we can integrate this once to find that $f_{y'} = \text{constant}$, and in fact it follows that $y' = \text{constant}$, so that our solution is

$$\begin{aligned} y_0 &= cx + d \\ &= \frac{(\beta - \alpha)}{(b - a)}(x - a) + \alpha \end{aligned}$$

Now we must show that this solution is in fact minimizing. A certain property of $f = f(y')$ makes this easy: it is convex. You can check that $f''(y') = 1/(1 + y')^{3/2} > 0$. It follows that

$$f(y') > f(y_0') + f_{y'}(y_0')[y' - y_0']$$

whenever $y' \neq y_0'$. Then if $y \neq y_0$,

$$\begin{aligned} I[y] = \int_a^b f(y')\,\mathrm{d}x &> \int_a^b [f(y_0') + f_{y'}(y_0')[y' - y_0']]\,\mathrm{d}x \\ &\geq I[y_0] + (\text{const.})\int_a^b [y' - y_0']\,\mathrm{d}x \\ &= I[y_0] \end{aligned}$$

since y and y_0 have the same endpoints. So any other curve has a strictly larger length.

Remark. We will see in section 3.5 that properties like f being independent from x or y lead to conservation laws.

Here is a slightly more complicated example of converting a problem into a solvable format:

Example 3.12. An industrial pump uses electricity at a rate $r(u) = 10+u$ units per kilotonne when pumping water at u kilotonnes per hour; in this remote location, the cost of wholesale electricity varies significantly over the day, with a price of

$$c(t) = 169 - (t-12)^2$$

tenths of a penny per unit after t hours (an average of 12.1p). What is the lowest cost that can be achieved if, over the course of a day, it must pump 100 kilotonnes?

Letting $V(t)$ be the volume pumped after t hours, the total cost is

$$\begin{aligned} I[V] &= \int c(t)\, r(u)\, \mathrm{d}V \\ &= \int c(t)\, r(u) \frac{\mathrm{d}V}{\mathrm{d}t} \mathrm{d}t \\ &= \int c(t)\, r(u)\, u\, \mathrm{d}t \end{aligned}$$

In this problem the independent variable is t, the function we are working with is V, and u is its first derivative - so in terms of the usual notation we would have had $t \to x$, $V \to y$ and $u \to y'$. This has the Euler-Lagrange equation

$$\frac{\partial f}{\partial V} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f}{\partial u}\right) = 0 - \frac{\mathrm{d}}{\mathrm{d}t}\left(c(t)\left[r(u) + r'(u)\, u\right]\right) = 0$$

which immediately gives us

$$c(t)\left[r(u) + r'(u)\, u\right] = \text{constant}$$

since the function is once again independent of V. Explicitly,

$$\begin{aligned} \left[169 - (t-12)^2\right][10 + u + u] &= A \\ u &= \frac{A}{2\left[169 - (t-12)^2\right]} - 5 \\ &= \frac{A/2}{[13 + (t-12)][13 - (t-12)]} - 5 \end{aligned}$$

We can integrate this to get

$$V(t) = B \cdot \operatorname{arctanh}\left(\frac{t-12}{13}\right) - 5t + C$$

and the initial conditions give

$$\begin{aligned} V(0) &= B \cdot \operatorname{arctanh}\left(-\frac{12}{13}\right) + C = 0 \\ V(24) &= B \cdot \operatorname{arctanh}\left(\frac{12}{13}\right) - 120 + C = 100 \end{aligned}$$

so that

$$\begin{aligned} C &= 110 \\ B &= \frac{110}{\operatorname{arctanh}\left(\frac{12}{13}\right)} \end{aligned}$$

You may check that the solution curve

$$\boxed{V(t) = 110\frac{\operatorname{arctanh}\left(\frac{t-12}{13}\right)}{\operatorname{arctanh}\left(\frac{12}{13}\right)} - 5t + 110}$$

is in fact valid (i.e. $u = V'(t) \geq 0$ at all times). It is depicted in Figure 3.2.

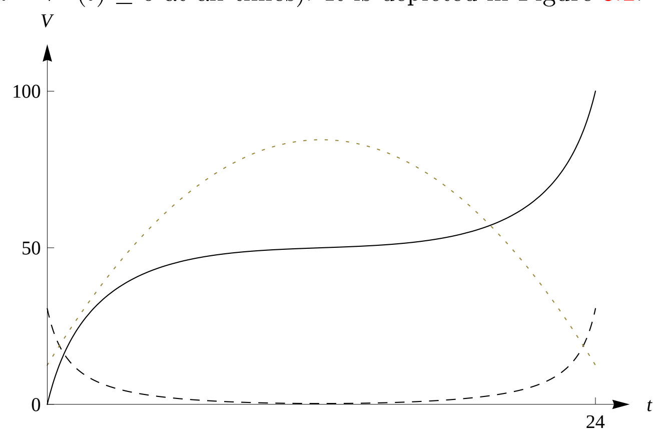

Figure 3.2: The optimum volume over time; the pump rate is shown as a dashed line, and the cost per unit electricity shown as a dotted line.

It is left as an exercise to plug this back into the original functional $I[V]$ to obtain

$$I[V] = 24200\left(-3 + \frac{13}{\log 5}\right)$$

tenths of a penny, a cost of £122.87.

You can compare this to the solution which does work at a constant rate, $V_1(t) = 100t/24$. The rate of electricity usage is $r(u) = 10 + u \approx 14.17$ units per kilotonne. Multiplying this by 100 gives the number of units used, and multiplying this by the time-averaged cost of a unit (around 12.1p) gives a cost of approximately £171.41. This represents a saving of around 29%.

We will not determine whether or not this is a global minimum here; however, this is a reasonably tractable problem which you may like to attempt as an exercise.

The above lemmas deal only with the case of fixed endpoints; but in fact the above lemmas can be used equally well to apply to general periodic problems of the type we saw above. A more interesting example than these is given by the following:

Example 3.13. Let $g(x) = \sin(nx)$. Minimize

$$I[u] = \int_{-\pi}^{\pi} \left[\frac{1}{2}\left((u')^2 + u^2\right) - gu\right] \mathrm{d}x$$

amongst all smooth 2π-periodic functions u,

$$u \in C^{\infty}_{\text{per}}([-\pi, \pi])$$

We have a functional of the above form, with

$$\begin{aligned} f(x, u, u') &= \frac{1}{2}\left[(u')^2 + u^2\right] - gu \\ \frac{\partial f}{\partial u} &= u - g \\ \frac{\partial f}{\partial u'} &= u' \end{aligned}$$

We can immediately deduce that the Euler-Lagrange equation is

$$\begin{aligned} \frac{\partial f}{\partial u} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial u'}\right) &= u - g - \frac{\mathrm{d}u'}{\mathrm{d}x} \\ &= u - g - u'' \\ &= 0 \end{aligned}$$

since boundary terms in the proof of Lemma 3.8 still vanish, and we can trivially modify the ϕ used in the proof of Lemma 3.9 to be 2π-periodic - then since any variation on a candidate curve can be written as $u = u_0 + t\phi$ where ϕ is 2π-periodic with arbitrary endpoints, the same arguments hold.

So any minimizing function u_0 satisfies

$$\boxed{u_0'' - u_0 + \sin(nx) = 0}$$

We can calculate the general solution of this equation as

$$u_0 = A\cosh x + B\sinh x + \frac{\sin(nx)}{1 + n^2}$$

(which could also be written in terms of e^x and e^{-x}). But recall that we must have $u \in C^{\infty}_{\text{per}}([-\pi, \pi])$ - that is, all solutions u must be periodic. Clearly, no non-trivial linear combination of $\cosh x$ and $\sinh x$ can be periodic (you can prove this as a quick exercise; see the note at the end of this example for another method), so in fact the only stationary point of the functional is

$$u_0 = \frac{\sin(nx)}{1 + n^2}$$

Now all that remains is to show that this is indeed a minimum. We can do this very directly, as follows:

$$\begin{aligned} I[u_0+\phi] &= \int_{-\pi}^{\pi}\left[\frac{1}{2}\left((u_0'+\phi')^2+(u_0+\phi)^2\right)-g(u_0+\phi)\right]\mathrm{d}x \\ &= I[u_0]+\int_{-\pi}^{\pi}[u_0'\phi'+u_0\phi-g\phi]\,\mathrm{d}x+\int_{-\pi}^{\pi}\frac{1}{2}\left(\phi'^2+\phi^2\right)\mathrm{d}x \end{aligned}$$

Now the first of these two integrals is in fact identically zero, by our choice of u_0, as can be shown by integrating the first term by parts: we end up integrating $(-u_0''+u-g)\,\phi=0$, and the boundary terms vanish. Therefore,

$$I[u_0+\phi]=I[u_0]+\int_{-\pi}^{\pi}\frac{1}{2}\left(\phi'^2+\phi^2\right)\mathrm{d}x$$

But the second integral is obviously non-negative, so we have immediately $I[u_0+\phi]\geq I[u_0]$. In fact, for all functions $\phi\in C^{\infty}_{\text{per}}([-\pi,\pi])$ with $\phi\not\equiv 0$, the last term is strictly positive, so that $I[u_0+\phi]>I[u_0]$. Therefore,

$$u_0=\frac{\sin(nx)}{1+n^2}$$

is a strict global minimizer for I.

An alternative and more general way to show that u_0 is the only smooth, 2π-periodic solution of the equation of the above differential equation is to consider another solution v, and form $w=u_0-v$. Then clearly w is a 2π-periodic function satisfying the now homogeneous equation $w''-w=0$. Again, we can simply assert that there is no periodic solution to this; or we can consider

$$0=\int_{-\pi}^{\pi}w(-w''+w)\,\mathrm{d}x=-[ww']_{-\pi}^{\pi}+\int_{-\pi}^{\pi}\frac{1}{2}\left(w'^2+w^2\right)\mathrm{d}x$$

Here, the boundary terms vanish by the periodicity of w, so the (non-negative) integrand on the right-hand side must be identically zero: $w\equiv w'\equiv 0$. Hence $v=u_0$.

(The complex analyst may like to prove this particular result via an application of Liouville's theorem, which states that a bounded function which is complex differentiable everywhere in $\mathbb{C}$ - an entire function - is constant.)

Remark. You may be curious about the existence of a '*direct method*' for proving the existence of solutions to these problems, generalizing the idea of "continuous functions on closed, bounded intervals attain their bounds". The *direct method in the calculus of variations* does exactly this. We will not discuss this in detail here, because it is essentially an exercise in topology. For proving the existence of a minimizer, the essential idea is to first show that the functional is bounded below, and hence that there must be functions (u_n) which tend to the infimum of the functional's value; then, we show that there is some subsequence which converges to, $u_{n_k}\to u_0$, with respect to some topology on the function space V; and finally, we show that J is sufficiently continuous with respect to this topology, so that it follows that $J(u_0)$ is a minimum value.

3.3 Multi-Dimensional Euler-Lagrange Equations

The above arguments carry over very well to the case of higher-dimensional integrals, where we have functionals of the form

$$I[u] = \int_\Omega f(\mathbf{x}, u, \boldsymbol{\nabla} u)\,\mathrm{d}V$$

where $u = u(\mathbf{x})$ is a function defined on some domain $\Omega \subset \mathbb{R}^n$, and

$$\boldsymbol{\nabla} u = \left(\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \cdots, \frac{\partial u}{\partial x_n}\right)$$

We are still integrating over a region Ω of values of the independent variable $\mathbf{x} = (x_1, x_2, \cdots, x_n)$, with a volume element $\mathrm{d}V$ instead of a line element $\mathrm{d}x$.

Remark. Note that f can actually be an arbitrary function of the x_i and the $\partial u/\partial x_i$ by taking dot (inner) products:

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

The key generalization which needs to be made is that what was integration by parts in one dimension becomes an application of Green's identities in higher dimensions. This is best introduced with an example (with origins in physics):

Example 3.14. Consider the functional

$$I[u] = \int_\Omega \left(\frac{1}{2}|\boldsymbol{\nabla} u|^2 - g(\mathbf{x})\,u\right)\mathrm{d}V$$

where Ω is some domain in $\mathbb{R}^n$. Then the directional derivative

$$\begin{aligned}
D_\phi I[u] &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} I[u + t\phi] \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \int_\Omega \left(\frac{1}{2}|\boldsymbol{\nabla} u + t\boldsymbol{\nabla}\phi|^2 - g(\mathbf{x})(u + t\phi)\right)\mathrm{d}V \\
&= \int_\Omega (\boldsymbol{\nabla} u \cdot \boldsymbol{\nabla}\phi - g\phi)\,\mathrm{d}V \\
&= \int_\Omega \frac{\delta I}{\delta u}\phi\,\mathrm{d}V
\end{aligned}$$

so we need to transform the term in $\boldsymbol{\nabla}\phi$ into a term in ϕ. To achieve this, recall Green's first identity, which gives

$$\int_\Omega \boldsymbol{\nabla} u \cdot \boldsymbol{\nabla}\phi\,\mathrm{d}\mathbf{x} = \int_{\partial\Omega} \phi\boldsymbol{\nabla} u \cdot \mathrm{d}\mathbf{S} - \int_\Omega \phi\boldsymbol{\nabla}^2 u\,\mathrm{d}V]$$

Assuming that there are fixed boundary conditions, we have $\phi = 0$ on the boundary $\partial\Omega$, and hence

$$D_\phi I[u] = \int_\Omega \left(-\phi\boldsymbol{\nabla}^2 u - g\phi\right)\mathrm{d}V$$

so that the directional derivative is

$$\frac{\delta I}{\delta u} = -\boldsymbol{\nabla}^2 u - g$$

Therefore, carrying over the results from the previous section, we can infer that this is must zero everywhere for u to be an extremal function: hence *Poisson's equation* arises: u is the solution to

$$\boldsymbol{\nabla}^2 u = -g$$

We can give this a physical interpretation: the field u obeying Poisson's equation, for the gravitational potential of a mass distribution proportional to $g(\mathbf{x})$, or the electrostatic potential of a charge distribution proportional to $g(\mathbf{x})$, will be that which minimizes the associated amount of energy given by

$$I[u] = \int_\Omega \left(\frac{1}{2} |\boldsymbol{\nabla} u|^2 - g(\mathbf{x}) u \right) \mathrm{d}V$$

More generally, we can have a functional of the form

$$I[u] = \int_\Omega f(\mathbf{x}, u, \boldsymbol{\nabla} u) \, \mathrm{d}V$$

and the associated Euler-Lagrange equation is given by

$$\boxed{\frac{\partial f}{\partial u} - \sum_{j=1}^{n} \frac{\partial}{\partial x_j} \left[\frac{\partial f}{\partial p_j} (\mathbf{x}, u, \boldsymbol{\nabla} u) \right] = 0}$$

and $p_j = \partial u / \partial x_j$.

(Hamilton's principle states that a system always evolves along a path which makes its action stationary - most familiar microscopic scale physical laws can be expressed as the Euler-Lagrange equations of a suitable integral functional.)

Proof. We proceed as before:

$$\begin{aligned}
D_\phi I[u] &= \left. \frac{\mathrm{d}}{\mathrm{d}t} \right|_{t=0} I[u + t\phi] \\
&= \left. \frac{\mathrm{d}}{\mathrm{d}t} \right|_{t=0} \int_\Omega f(\mathbf{x}, u + t\phi, \boldsymbol{\nabla} u + t \boldsymbol{\nabla} \phi) \, \mathrm{d}V \\
&= \int_\Omega \left[\phi \frac{\partial f}{\partial u} + \sum_{j=1}^{n} \frac{\partial f}{\partial p_j} \frac{\partial \phi}{\partial x_j} \right] \mathrm{d}V \\
&= \int_\Omega \left[\phi \frac{\partial f}{\partial u} + \boldsymbol{\nabla} \phi \cdot \frac{\partial f}{\partial \mathbf{p}} \right] \mathrm{d}V \\
&= \int_\Omega \left[\phi \frac{\partial f}{\partial u} - \phi \boldsymbol{\nabla} \cdot \frac{\partial f}{\partial \mathbf{p}} \right] \mathrm{d}V \\
&= \int_\Omega \left[\frac{\partial f}{\partial u} - \boldsymbol{\nabla} \cdot \frac{\partial f}{\partial \mathbf{p}} \right] \phi \, \mathrm{d}V
\end{aligned}$$

where we have adopted the notation $\frac{\partial f}{\partial \mathbf{p}} = \left(\cdots, \frac{\partial f}{\partial p_j}, \cdots \right)$, and used the fact that ϕ is 0 at the boundary.

It follows, by application of the same sort of methods which we used in the one-dimensional case, that

$$\boxed{\frac{\partial f}{\partial u} - \boldsymbol{\nabla} \cdot \frac{\partial f}{\partial \mathbf{p}} \equiv 0}$$

□

We can apply this directly to the following example of an action for a field:

Example 3.15. The action

$$\int_{\mathbb{R}^2} \frac{1}{2} \left[\left(\frac{\partial u}{\partial t} \right)^2 - \left(\frac{\partial x}{\partial t} \right)^2 \right] \mathrm{d}x \mathrm{d}t$$

associated with one spatial dimension gives $\mathbf{x} = (t, x)$ and $\mathbf{p} = (u_t, u_x)$, so

$$f = \frac{1}{2} \left(u_t^2 - u_x^2 \right)$$

which has the Euler-Lagrange equation

$$-\frac{\partial}{\partial t} (u_t) - \frac{\partial}{\partial x} (-u_x) = -u_{tt} + u_{xx} = 0$$

which is the *wave equation.*

Remark. In fact, this action is very much like a component of that determining the evolution of the electromagnetic $\mathbf{E}$ and $\mathbf{B}$ fields, which also exhibit this wave-like behaviour in the form of light.

3.4 Constrained Euler-Lagrange Equations

A natural question to ask is whether our techniques for finding extremal points of functionals can be generalized to include constraints, as we could in finite-dimensional vector spaces via the introduction of Lagrange multipliers. The answer is yes, thanks to the way that the properties of vector spaces are highly independent of the dimension. We will leave aside the details of showing that this is rigorously valid, instead illustrating by example how we go about constructing the 'augmented functional' and solving for the extremal function.

3.4.1 Single constraint

Example 3.16. Recall Problem 1.2, that of maximizing the area beneath a curve,

$$I[y] = \int_{-a}^{a} y(x) \, \mathrm{d}x$$

where we have a fixed length,

$$J[y] = \int_{-a}^{a} \sqrt{1 + y'^2} \, \mathrm{d}x = L$$

We define the augmented functional by

$$\begin{aligned}\Phi[y,\lambda] &= I[y]+\lambda(J[y]-L)\\ &= \int_{-a}^{a}\left(y+\lambda\sqrt{1+y'^2}-\lambda\frac{L}{2a}\right)\mathrm{d}x\end{aligned}$$

where the constant term is actually going to be irrelevant.

This has the functional derivative, with respect to y, of

$$\frac{\delta\Phi}{\delta y}=\frac{\partial f}{\partial y}-\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right)=1-\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\lambda y'}{\sqrt{1+y'^2}}\right)$$

It follows, integrating once, that

$$x-\frac{\lambda y'}{\sqrt{1+y'^2}}=c$$

or

$$\begin{aligned}\frac{y'}{\sqrt{1+y'^2}} &= \frac{x-c}{\lambda}\\ y'^2 &= \frac{[(x-c)/\lambda]^2}{1-[(x-c)/\lambda]^2}\end{aligned}$$

Taking square roots and integrating, therefore, we have

$$\int \mathrm{d}y=\pm\int\frac{[(x-c)/\lambda]}{\sqrt{1-[(x-c)/\lambda]^2}}\mathrm{d}x$$

At this point, it is useful to make the substitution $x=c+\lambda\sin\theta$ which gives $y=y_0\pm\lambda\cos\theta$. This implies

$$(x-c)^2+(y-y_0)^2=\lambda^2$$

where the constants may be adjusted to fit the initial conditions and the constraint - it is clear, however, that the solution is a *circle*.

Remark. This formulation of the problem forbids shapes which double back on themselves and so on. The solution to this is to work with curves parametrized by a new variable: we write $\mathbf{x}(t)=(x(t),y(t))$, in a way which can be obviously generalized to more dependent variables, and $\mathbf{x}(t)\in\mathbb{R}^n$. Then we get integral functionals of the form $\int f(t,\mathbf{x}(t),\dot{\mathbf{x}}(t))\,\mathrm{d}t$. These generate a *family* of Euler-Lagrange equations:

$$\frac{\partial f}{\partial x_k}-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f}{\partial \dot{x}_k}\right)=0\quad\text{for } j=1,2,\cdots,n$$

As an aside, we show how these give the same solution:

Example 3.17. Find the closed curve $\mathbf{x}(t) \in \mathbb{R}^2$ with maximal area

$$A = \frac{1}{2}\int (x\dot{y} - y\dot{x})\,\mathrm{d}t$$

given the fixed length

$$L = \int \left(\dot{x}^2 + \dot{y}^2\right)^{\frac{1}{2}}\mathrm{d}t$$

This leads to

$$\Phi[\mathbf{x}, t] = \int \left[\frac{1}{2}(x\dot{y} - y\dot{x}) + \lambda\left(\dot{x}^2 + \dot{y}^2\right)^{\frac{1}{2}}\right]\mathrm{d}t$$

The Euler-Lagrange equations are

$$\begin{aligned}
\frac{\partial f}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f}{\partial \dot{x}}\right) &= \frac{1}{2}\dot{y} - \frac{\mathrm{d}}{\mathrm{d}t}\left(-\frac{1}{2}y + \frac{\lambda\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}}\right) \\
&= \dot{y} - \frac{\lambda\dot{y}\,(\dot{y}\ddot{x} - \dot{x}\ddot{y})}{(\dot{x}^2 + \dot{y}^2)^{\frac{3}{2}}} \\
&= \frac{\dot{y}\left(\left(\dot{x}^2 + \dot{y}^2\right)^{\frac{3}{2}} - \lambda(\dot{y}\ddot{x} - \dot{x}\ddot{y})\right)}{(\dot{x}^2 + \dot{y}^2)^{\frac{3}{2}}} \\
&= 0
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f}{\partial \dot{y}}\right) &= -\frac{1}{2}\dot{x} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}x + \frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}}\right) \\
&= \dot{x} - \frac{\dot{x}\,(\dot{x}\ddot{y} - \dot{y}\ddot{x})}{(\dot{x}^2 + \dot{y}^2)^{\frac{3}{2}}} \\
&= \frac{\dot{x}\left(-\left(\dot{x}^2 + \dot{y}^2\right)^{\frac{3}{2}} - \lambda(\dot{x}\ddot{y} - \dot{y}\ddot{x})\right)}{(\dot{x}^2 + \dot{y}^2)^{\frac{3}{2}}} \\
&= 0
\end{aligned}$$

Now we can dismiss solutions where $\dot{x} = 0$ and $\dot{y} = 0$ except at isolated points on geometrical grounds. This gives us

$$\begin{aligned}
\left(\dot{x}^2 + \dot{y}^2\right)^{\frac{3}{2}} - \lambda(\dot{y}\ddot{x} - \dot{x}\ddot{y}) &= 0 \\
-\left(\dot{x}^2 + \dot{y}^2\right)^{\frac{3}{2}} - \lambda(\dot{x}\ddot{y} - \dot{y}\ddot{x}) &= 0
\end{aligned}$$

which are obviously equivalent to each other. Then we have

$$\frac{(\dot{y}\ddot{x} - \dot{x}\ddot{y})}{(\dot{x}^2 + \dot{y}^2)^{\frac{3}{2}}} = \lambda^{-1}$$

Now we can integrate this once, by multiplying through by $\dot{y}$ and then noting the left-hand side is exactly the expression which arose above:

$$\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}} = \lambda^{-1}\left(y-y_0\right)$$

Similarly,

$$\frac{\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}} = -\lambda^{-1}\left(x-x_0\right)$$

Squaring and adding these two equations, we have

$$\begin{aligned} 1 &= \lambda^{-2}\left(y-y_0\right)^2+\lambda^{-2}\left(x-x_0\right)^2 \\ \lambda^2 &= \left(x-x_0\right)^2+\left(y-y_0\right)^2 \end{aligned}$$

which is still a circle!

Remark. Alternatively, we could have noted that

$$\frac{\left(\dot{y}\ddot{x}-\dot{x}\ddot{y}\right)}{\left(\dot{x}^2+\dot{y}^2\right)^{\frac{3}{2}}} = \lambda^{-1}$$

specifies that the curvature of the curve $\mathbf{x}\left(t\right)$ is a constant, exactly λ^{-1}, and hence that it is a circle with radius λ.

3.4.2 Multiple constraints

If there are only finitely many constraints, we generalize exactly as in the finite case.

If we have a family of constraints $J_\alpha\left[y\right]=0$, $\alpha=1,\cdots,N$, then we construct the functional

$$\Phi = I\left[y\right]+\sum_\alpha \lambda_\alpha J_\alpha\left[y\right]$$

However, if there are an continuous infinity of constraints, then we need to construct a 'Lagrange multiplier function'. This is best illustrated with an example from physics.

Example 3.18. In fluid mechanics, the velocity field $\mathbf{v}:\mathbb{R}^3\to\mathbb{R}^3$ is subject to an infinite number of constraints for incompressible flow,

$$\boldsymbol{\nabla}\cdot\mathbf{v}\left(\mathbf{x}\right)=0 \quad \forall\mathbf{x}$$

The question is then to minimize

$$I\left[\mathbf{v}\right]=\int\left(\frac{1}{2}\left|\boldsymbol{\nabla}\mathbf{v}\right|^2-\mathbf{v}\cdot\mathbf{f}\right)\mathrm{d}V$$

subject to $\boldsymbol{\nabla}\cdot\mathbf{v}\left(\mathbf{x}\right)=0$. There is some unusual notation here: we define the gradient of a vector to be the tensor

$$\boldsymbol{\nabla}\mathbf{v}=\left(\frac{\partial v^i}{\partial x_j}\right)_{i,j=1,2,3}$$

where v^i are the three components of the $\mathbf{v}$ (using the raised index to define a contravariant component, since we are defining the tensor formally - you may think of them as v_i for practical purposes here). The expression $|\boldsymbol{\nabla}\mathbf{v}|^2$ is used here to denote the sum of the squares of the components of this tensor:

$$|\boldsymbol{\nabla}\mathbf{v}|^2 = \sum_{i,j=1,2,3} \left(\frac{\partial v^i}{\partial x_j}\right)^2 = \sum_{i=1}^{3} |\boldsymbol{\nabla} v_i|^2$$

We construct the augmented function

$$\Phi[\mathbf{v}, \lambda] = \int \left(\frac{1}{2}|\boldsymbol{\nabla}\mathbf{v}|^2 - \mathbf{v}\cdot\mathbf{f} - \lambda(\mathbf{x})\,\boldsymbol{\nabla}\cdot\mathbf{v}\right) \mathrm{d}V$$

where there is still 'one Lagrange multiplier per constraint', in that $\lambda(\mathbf{x})$ is defined for all $\mathbf{x} \in \mathbb{R}^3$.

Now we can apply one of Green's identities to the term $\lambda(\mathbf{x})\,\boldsymbol{\nabla}\cdot\mathbf{v}$ to see

$$\int_\Omega \lambda\boldsymbol{\nabla}\cdot\mathbf{v}\,\mathrm{d}\mathbf{x} = -\int_\Omega [\boldsymbol{\nabla}\lambda]\cdot\mathbf{v}\,\mathrm{d}\mathbf{x} + \int_{\partial\Omega} \lambda\mathbf{v}\cdot\mathrm{d}\mathbf{S}$$

Hence if we assume that we have $\mathbf{0}$ boundary conditions for $\mathbf{v}$, for large $|\mathbf{x}|$ for example,

$$\Phi[\mathbf{v}, \lambda] = \int \left(\frac{1}{2}|\boldsymbol{\nabla}\mathbf{v}|^2 - \mathbf{v}\cdot\mathbf{f} + [\boldsymbol{\nabla}\lambda(\mathbf{x})]\cdot\mathbf{v}\right) \mathrm{d}V$$

Now taking the directional derivative of this, we see

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi[\mathbf{v} + t\mathbf{w}, \lambda] = \int (\boldsymbol{\nabla}\mathbf{v} : \boldsymbol{\nabla}\mathbf{w} - \mathbf{f}\cdot\mathbf{w} + [\boldsymbol{\nabla}\lambda(\mathbf{x})]\cdot\mathbf{w})\,\mathrm{d}V$$

where we use the colon to denote summing the products of corresponding cells - this arises because

$$\begin{aligned}
\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \frac{1}{2}|\boldsymbol{\nabla}\mathbf{v} + t\boldsymbol{\nabla}\boldsymbol{w}|^2 &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \sum_{i,j} \frac{1}{2}\left(\frac{\partial v^i}{\partial x_j} + t\frac{\partial w^i}{\partial x_j}\right)^2 \\
&= \sum_{i,j} \frac{\partial w^i}{\partial x_j}\frac{\partial v^i}{\partial x_j} \\
&= \boldsymbol{\nabla}\mathbf{v} : \boldsymbol{\nabla}\mathbf{w}
\end{aligned}$$

We now rewrite this first term, using the Green's identity again, as

$$\int \left(\sum_i \boldsymbol{\nabla} v_i \cdot \boldsymbol{\nabla} w_i\right) \mathrm{d}\mathbf{x} = \int \left(\sum_i -w_i\boldsymbol{\nabla}^2 v_i\right) \mathrm{d}V$$

Hence

$$D_w\Phi = \int \left(-\boldsymbol{\nabla}^2\mathbf{v} - \mathbf{f} + \boldsymbol{\nabla}\lambda\right)\cdot\mathbf{w}\,\mathrm{d}V$$

and thus the Euler-Lagrange equation (really a family of three equations, but we can place them in one system) can be written as

$$-\boldsymbol{\nabla}^2\mathbf{v} + \boldsymbol{\nabla}\lambda = \mathbf{f}$$

But since $\nabla \cdot \mathbf{v} = 0$, if we take the divergence of this equation we obtain

$$\nabla^2 \lambda = \nabla \cdot \mathbf{f}$$

So the field $\mathbf{v}$, with $\nabla \cdot \mathbf{v} = 0$, is stationary for I if

$$\begin{aligned} -\nabla^2 \mathbf{v} + \nabla \lambda &= \mathbf{f} \\ \nabla^2 \lambda &= \nabla \cdot \mathbf{f} \end{aligned}$$

Remark. These can be compared to the Navier-Stokes equation, for a time-independent (static) field, without the non-linear terms $O\left(|\mathbf{v}|^2\right)$. This gives the above equations, where $\lambda(\mathbf{x})$ is the *pressure*.

3.5 Conservation Laws and Noether's Theorem

As already noted, some forms of f lead to special versions of the Euler-Lagrange equation. These special cases are in fact of value to us, as we shall see in section 2.5.3. For now, however, we will simply note the two key results.

Let y be a solution of

$$f_y - \frac{\mathrm{d}}{\mathrm{d}x}(f_{y'}) = 0$$

Theorem 3.19.

(i) *If $f = f(x, y')$ has no y-dependence, then*

$$f_{y'} = \textit{constant}$$

(ii) *If $f = f(y, y')$ has no x-dependence, then*

$$y' f_{y'} - f = \textit{constant}$$

These two results state *conserved quantities* when the functional is independent of some property of the system: we call this invariance a *symmetry*. The laws above are *conservation laws*.

Proof.

(i) This part is trivial: $f_y = 0$ so

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}(f_{y'}) &= 0 \\ f_{y'} &= \text{constant} \end{aligned}$$

(ii) This is more involved, but easily deduced from working backwards, and using the chain rule:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\left[y' f_{y'} - f\right] &= y'' f_{y'} + y' \frac{\mathrm{d}}{\mathrm{d}x}\left(f_{y'}\right) - \frac{\mathrm{d}f}{\mathrm{d}x} \\
&= y'' f_{y'} + y' \frac{\mathrm{d}}{\mathrm{d}x}\left(f_{y'}\right) - y' f_y - y'' f_{y'} \\
&= y' \left[\frac{\mathrm{d}}{\mathrm{d}x}\left(f_{y'}\right) - f_y\right] \\
&= 0
\end{aligned}
$$

□

This shows that in general, when the integrand has some symmetry (which corresponds to independence from some type of coordinate), the solution has a symmetry also. This is particularly relevant in physics; as mentioned above, this will be briefly discussed in section 2.5.3.

Note that the argument made in the latter conservation law actually gives rise to an alternative expression of the Euler-Lagrange equation:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\left[y' f_{y'} - f\right] &= 0 - f_x \\
\frac{\mathrm{d}}{\mathrm{d}x}\left[y' f_{y'} - f\right] + f_x &= 0
\end{aligned}
$$

which is the *Beltrami identity.* As we shall see below, it is significant that the term which is differentiated is actually the Legendre transform of f with respect to y', if f is a convex function of y' - this is because we define the Hamiltonian as the Legendre transform of the Lagrangian, so Lagrangians which do not depend on the independent coordinate (namely the time t, confusingly represented by x above) give rise to constant Hamiltonians, $\mathrm{d}H/\mathrm{d}t = 0$ - so energy conservation is a consequence of the time-invariance of physical laws.

3.6 Scientific Applications

Laws of nature can often be expressed in terms of the minimization (or more generally making stationary) of some quantity.

3.6.1 Fermat's Principle

Light, in the absence of changing density or an interacting field, is well-known to travel in straight lines. It is also well-known that when light reflects off a boundary, the angles of incidence and reflection, measured from the normal of the surface at the point of reflection, are both equal: $\theta_i = \theta_r$. One way of expressing the first fact is with the postulate that light always takes the shortest path possible between two points (assuming it moves between them at all). This does not quite square with the second fact though: if the light moves between (x_1, y) and (x_2, y), then reflecting off a horizontal boundary is obviously not the shortest route. However, assuming that light travels in straight lines unless it interacts with a medium, note that the route of reflecting off the boundary at the point with

x-coordinate $(x_1 + x_2)/2$ actually is a *local* minimum for the possible distance taken: given paths $(x_1, y) \rightsquigarrow (x_2, y)$ striking the boundary at one point $(a, 0)$, the time taken is

$$T(a) = \frac{1}{c}\left(\left[(x_1 - a)^2 + y^2\right]^{1/2} + \left[(x_2 - a)^2 + y^2\right]^{1/2}\right)$$

and

$$T'(a) = \frac{x_1 - a}{\left[(x_1 - a)^2 + y^2\right]^{1/2}} + \frac{x_2 - a}{\left[(x_2 - a)^2 + y^2\right]^{1/2}}$$

which is zero precisely when $\sin\theta_i = \sin\theta_r$, as can be shown with a suitable substitution, noting that $x_1 < a < x_2$.

This principle, *Fermat's principle*, is in fact perfectly general.

Example 3.20. Consider light, in two dimensions, passing through an inhomogeneous medium where the speed of light is a function of y: $c = c(y)$. Then given a path which can be parametrized by $y = y(x)$, the time taken for light to follow that path is

$$\begin{aligned} T &= \int_a^b \frac{\sqrt{1 + y'^2}}{c(y)} \mathrm{d}x \\ &= \int_a^b f(y, y')\,\mathrm{d}x \end{aligned}$$

Therefore a sufficiently differentiable minimizing path would necessarily satisfy

$$f_y - \frac{\mathrm{d}}{\mathrm{d}x}(f_{y'}) = 0$$

which can then be solved for the path taken by the light.

3.6.2 Lagrangian mechanics

One striking class of applications of finding extremal functions for a functional comes from the ability to encode virtually every fundamental classical physics laws in terms of a correctly chosen *action*.

In elementary mechanics, for a particle moving in a time-independent potential $V(\mathbf{x})$ we have a force $\mathbf{F} = -\boldsymbol{\nabla} V(\mathbf{x})$, and an equation of motion

$$m\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} = -\boldsymbol{\nabla} V(\mathbf{x})$$

We can derive this as the Euler-Lagrange equation for the action

$$\begin{aligned} S[\mathbf{x}] &= \int \left[\frac{1}{2}m|\dot{\mathbf{x}}|^2 - V(\mathbf{x})\right] \mathrm{d}t \\ &= \int \mathcal{L}(\mathbf{x}, \dot{\mathbf{x}})\,\mathrm{d}t \end{aligned}$$

where $\mathcal{L}$ is the *Lagrangian* we defined in section 2.28, equal to the kinetic energy minus the potential energy. This is because

$$\frac{\partial f}{\partial x_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f}{\partial \dot{x}_i}\right) = -\frac{\partial V}{\partial x_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(m\dot{x}_i\right) = 0$$

so

$$m\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} = -\boldsymbol{\nabla} V$$

Recall from section 3.5 that if the integrand, here the Lagrangian $\mathcal{L}$, has some symmetry, then this should be reflected by a conservation law for some property of the physical system:

- If V is independent of $\mathbf{x}$ (which is equivalent to saying that there is no force $\mathbf{F}$), then the conserved quantity is clearly

  $$\frac{\partial f}{\partial \dot{x}_i} = m\dot{x}_i$$

  which is otherwise known as *momentum*. This is one way of thinking about Newton's first law.

- Since V is independent of time, we necessarily have

  $$\begin{aligned} \dot{\mathbf{x}} \cdot \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{x}}} - \mathcal{L} &= \dot{\mathbf{x}} \cdot m\dot{\mathbf{x}} - \left(\frac{1}{2}m\left|\dot{\mathbf{x}}\right|^2 - V\left(\mathbf{x}\right)\right) \\ &= \frac{1}{2}m\left|\dot{\mathbf{x}}\right|^2 + V\left(\mathbf{x}\right) \\ &= \text{constant} \end{aligned}$$

  which is the statement of the *conservation of energy*.

If we wrote $\mathcal{L}$ to include rotational coordinates, we could also deduce the conservation of angular momentum from the independence of the potential on these angles.

A generalization of the second observation comes from recalling that the Hamiltonian H is defined as the Legendre transform[8] of $\mathcal{L}$ with respect to $\dot{\mathbf{x}}$:

$$\mathcal{H} = \mathbf{p} \cdot \dot{\mathbf{x}} - \mathcal{L}\left(\mathbf{x}, \dot{\mathbf{x}}\right)$$

where $\mathbf{p}$ is the conjugate momentum, given by

$$\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{x}}}$$

That is,

$$\mathcal{H} = \dot{\mathbf{x}} \cdot \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{x}}} - \mathcal{L} = \text{constant}$$

Remark. Forces which are derived as approximations, like friction, which arises only at the macroscopic scale as the average effect of complicated electromagnetic interactions, cannot usually be described in this way. However, as we have seen above, even the evolution of fields can be expressed in this way. All of Maxwell's equations can be readily deduced from a single action.

One of the most important applications of Lagrangian mechanics, and the action principle, is to quantum mechanics.

[8] Note the definition given is actually only that of the Legendre transform if L is convex in $\dot{\mathbf{x}}$.

3.6.3 Geodesics

Definition 3.21. A *geodesic* is a locally length-minimizing curve - a curve of least length, or more generally a stationary point for the length

$$L = \int \mathrm{d}s$$

where s is the arclength.

Remark. By '*locally* length minimizing', we mean that any sufficiently small variation on the line will increase its length.

In the Euclidean plane, we have already seen (in Example 3.11) that a path between (a, α) and (b, β) minimizing the length is a straight line. In fact, this is the only stationary point; at the time, we only saw this for curves which can be parametrized as $y = y(x)$, so that the length is

$$L = \int_a^b \sqrt{1 + y'^2}\mathrm{d}x$$

We have since shown how a curve can be parametrized by a new coordinate to derive a more general result, in Example 3.17. This is the approach we will adopt here.

Example 3.22. Consider the length of a parametrized curve $\mathbf{x}(t)$,

$$L[\mathbf{x}] = \int_a^b \|\dot{\mathbf{x}}\| \,\mathrm{d}t$$

We can easily analyze this because the integrand is independent of the components of $\mathbf{x}$ - the Euler-Lagrange equation can immediately be integrated once, so

$$\frac{\partial f}{\partial \dot{x}_j} = \frac{\dot{x}_j}{\left(\sum \dot{x}_j^2\right)^{1/2}} = \text{constant}$$

and hence

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{\mathbf{x}}}{\|\dot{\mathbf{x}}\|}\right) = 0$$

so that the direction taken is constant, though 'speed' may vary.

In fact, since

$$\begin{aligned} L &= \int \left[\sum_j \left(\frac{\mathrm{d}x_j}{\mathrm{d}t}\right)^2\right]^{1/2} \mathrm{d}t \\ &= \int \left[\sum_j \left(\frac{\mathrm{d}x_j}{\mathrm{d}\tau}\right)^2\right]^{1/2} \mathrm{d}\tau \end{aligned}$$

for any change of variables $\tau = \tau(t)$ where $\tau'(t) > 0$, we can rescale the parameter arbitrarily.

This means it is simplest to choose the parameter τ such that the speed

$$\left[\sum_j \left(\frac{\mathrm{d}x_j}{\mathrm{d}\tau}\right)\right]^{1/2} = \text{constant}$$

which means τ is proportional to arclength.

Using this parametrization, geodesics are curves which make stationary

$$I[\mathbf{x}] = \int \frac{1}{2} \|\dot{\mathbf{x}}\|^2 \,\mathrm{d}t = \int \{\text{kinetic energy}\}\,\mathrm{d}t$$

which is actually an equivalent definition for the multiple of arclength parametrization

The Euler-Lagrange equation becomes

$$\boxed{\ddot{\mathbf{x}} = \mathbf{0}, \quad \text{i.e.} \quad \dot{\mathbf{x}} = \text{constant}}$$

Note that this is all exactly the same as the mechanics example, with $m = 1$ and $V = 0$. This shows some of the generality of the method. The above equation states that geodesics are the paths followed by non-accelerating particles.

This method can be generalized to very different spaces in order to find geodesics on them. For an example, we will see two ways of finding the geodesic curves on a cylinder

$$C = \left\{(x, y, z) : x^2 + y^2 = R^2, -\infty < z < \infty\right\}$$

Example 3.23. Firstly, recall that we can convert from cylindrical coordinates to Cartesian coordinates via

$$x = R\cos\theta \quad y = R\sin\theta \quad z = z$$

where in this case, R is fixed, and θ and z are the two variable coordinates. We will therefore parametrize our path by t, so that $\theta = \theta(t)$ and $z = z(t)$.

Now the 'speed squared' is given by

$$\|\dot{\mathbf{x}}\|^2 = \left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 = \left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2$$

But we can write this in terms of the cylindrical coordinates:

$$\begin{aligned} \|\dot{\mathbf{x}}\|^2 &= \left(-R\sin(\theta)\dot{\theta}\right)^2 + \left(R\cos(\theta)\dot{\theta}\right)^2 + \dot{z}^2 \\ &= R^2\dot{\theta}^2 + z^2 \end{aligned}$$

Now according to the same theory we developed for a free Cartesian space, a geodesic curve on C is a curve

$$\mathbf{x}(t) = (R\cos\theta(t), R\sin\theta(t), z(t))$$

which makes stationary

$$I[\mathbf{x}] = \frac{1}{2}\int \left(R^2\dot{\theta}^2 + \dot{z}^2\right) \mathrm{d}t$$

This gives Euler-Lagrange equations for both variables:

$$\begin{aligned} \frac{\delta I}{\delta \theta} &= 0 - \frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{\theta}R^2\right) = 0 \\ \frac{\delta I}{\delta z} &= 0 - \frac{\mathrm{d}}{\mathrm{d}t}(\dot{z}) = 0 \end{aligned}$$

which respectively give

$$\begin{aligned} \dot{\theta} &= \text{const} \\ \dot{z} &= \text{const} \end{aligned}$$

Hence the geodesic curves are *helicoidal* curves, which rotate about the cylinder at a constant rate, whilst moving upwards at a constant rate. (Or at least, when we constrain the particle to move at a constant speed, the rate of change of the angle and vertical ascent are both constant.) Note that whilst all helicoidal curves are stationary points, they are not all minima.

The second solution we give treats this as a constraint problem, rather than as a parametrization problem. (Recall we initially derived the ideas behind Lagrange multipliers from a parametrization of the constrained domain.)

Example 3.24. The relevant constraint is

$$g(x, y) = x^2 + y^2 - R^2 = 0$$

In fact, because this constraint must apply at every point on the path, this corresponds to an infinite set of constraints: hence we need an infinite number of multipliers, which can be denoted by $\lambda(t)$.

Then we form

$$\begin{aligned} \Phi[\mathbf{x}, \lambda] &= \int \left(\frac{1}{2}\|\dot{\mathbf{x}}\|^2 - \lambda(t)\left[x^2 + y^2 - R^2\right]\right) \mathrm{d}t \\ &= \int \left(\frac{1}{2}\left[\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right] - \lambda(t)\left[x^2 + y^2 - R^2\right]\right) \mathrm{d}t \end{aligned}$$

The associated equations are then

$$\begin{aligned} -2\lambda x - \frac{\mathrm{d}}{\mathrm{d}t}(\dot{x}) = 0 &\iff \ddot{x} + 2\lambda x = 0 \\ -2\lambda y - \frac{\mathrm{d}}{\mathrm{d}t}(\dot{y}) = 0 &\iff \ddot{y} + 2\lambda y = 0 \\ 0 - \frac{\mathrm{d}}{\mathrm{d}t}(\dot{z}) = 0 &\iff \dot{z} = \text{constant} \end{aligned}$$

Now if $\lambda(t) \geq 0$ is constant, then x and y are both harmonic oscillators, though they are not independent, since we have the constraint $x^2 + y^2 - R^2 = 0$. Without assuming anything about λ, we can differentiate this twice, giving

$$\begin{aligned} x\dot{x} + y\dot{y} &= 0 \\ x\ddot{x} + y\ddot{y} + \dot{x}^2 + \dot{y}^2 &= 0 \end{aligned}$$

Then to find $\lambda(t)$, we can use the Euler-Lagrange equations: $\ddot{x} = -2\lambda x$ and $\ddot{y} = -2\lambda y$ so

$$\begin{aligned} -2\lambda x^2 - 2\lambda y^2 + \dot{x}^2 + \dot{y}^2 &= 0 \\ 2\lambda R^2 &= \dot{x}^2 + \dot{y}^2 \end{aligned}$$

which gives us

$$\lambda(t) = \frac{1}{2R^2}\left(\dot{x}^2 + \dot{y}^2\right) \geq 0$$

Now writing $\omega(t) = \sqrt{2\lambda} \in \mathbb{R}$ we have

$$\ddot{x} + \omega^2 x = 0, \quad \ddot{y} + \omega^2 y = 0$$

Now we can use this to write

$$\begin{aligned} x\dot{x} + y\dot{y} &= 0 \\ \frac{1}{\omega^2}\left[\ddot{x}\dot{x} + \ddot{y}\dot{y}\right] &= 0 \\ \dot{x}^2 + \dot{y}^2 &= \text{constant} \end{aligned}$$

where in the last step, we multiplied by $2\omega^2$ and integrated. It follows that λ and hence ω are both constant. Hence the solutions may be written as

$$\begin{aligned} x &= R\cos(\omega t + \alpha) \\ y &= \pm R\sin(\omega t + \alpha) \\ z &= at + b \end{aligned}$$

These methods for finding geodesics can be readily generalized to more abstract spaces (manifolds), so long as one takes care to define all the terms correctly.

3.6.4 Brachistochrone problem

One of the classic problems in the calculus of variations is to find the curve $y(x)$ such that a bead moving under gravity along a frictionless wire described by $y(x)$ takes the shortest possible time to fall from rest at $(0,0)$ to (X,Y). This is called the *brachistochrone problem*, from the Greek for 'shortest time'.

Measuring y downwards, so that $Y > 0$, the speed of the particle $v = \left(\dot{x}^2 + \dot{y}^2\right)^{1/2}$ must satisfy

$$\frac{1}{2}mv^2 = mgy$$

by conservation of energy, from which it follows that

$$v = \sqrt{2gy}$$

Then the functional describing the time taken is

$$\begin{aligned} I[y] = \int \frac{\mathrm{d}s}{u} &= \int \frac{\left(\dot{x}^2 + \dot{y}^2\right)^{1/2}}{\sqrt{2gy}} \mathrm{d}t \\ &= \frac{1}{\sqrt{2g}} \int_0^X \frac{\left(1 + y'^2\right)^{1/2}}{\sqrt{y}} \mathrm{d}x \end{aligned}$$

The associated Euler-Lagrange equation is clearly a very unpleasant affair if expanded directly. However, we can save ourselves some time using the conservation law for integrands independent of x:

$$y' f_{y'} - f = \frac{y'^2}{\sqrt{y\left(1 + y'^2\right)}} - \frac{\sqrt{\left(1 + y'^2\right)}}{\sqrt{y}} = C$$

This implies that

$$\begin{aligned} y'^2 - \left(1 + y'^2\right) &= C\sqrt{y\left(1 + y'^2\right)} \\ 1 &= C^2 y \left(1 + y'^2\right) \end{aligned}$$

which we can rearrange and attempt to integrate:

$$\int \frac{y^{1/2}\mathrm{d}y}{\left(1 - c^2 y\right)^{1/2}} = \frac{1}{c} \int \mathrm{d}x = \frac{x}{c}$$

Let $u = y^{1/2}$. Then we have $\mathrm{d}y/\mathrm{d}u = 2u$, so

$$\int \frac{y^{1/2}\mathrm{d}y}{\left(1 - c^2 y\right)^{1/2}} = \int \frac{2u^2 \mathrm{d}u}{\left(1 - c^2 u^2\right)^{1/2}}$$

which you might recognize as being most readily solved with a substitution like $u = \frac{1}{c}\sin\frac{\theta}{2}$. In fact, substituting $y = \frac{1}{c^2}\sin^2\frac{\theta}{2}$ into the original formula gives $\mathrm{d}y/\mathrm{d}\theta = \frac{1}{c^2}\sin\frac{\theta}{2}\cos\frac{\theta}{2}$ and hence

$$\begin{aligned} \int \frac{y^{1/2}\mathrm{d}y}{\left(1 - c^2 y\right)^{1/2}} &= \int \frac{\frac{1}{c}\sin\frac{\theta}{2} \cdot \frac{1}{c^2}\sin\frac{\theta}{2}\cos\frac{\theta}{2}\mathrm{d}\theta}{\cos\frac{\theta}{2}} \\ &= c^{-3}\int \sin^2\frac{\theta}{2}\mathrm{d}\theta \\ &= \frac{1}{2c^3}\left(\theta - \sin\theta\right) \end{aligned}$$

Hence we have

$$
\begin{aligned}
y &= \frac{1}{c^2}\sin^2\frac{\theta}{2} = \frac{1}{2c^2}(1-\cos\theta) \\
x &= \frac{1}{2c^2}(\theta - \sin\theta)
\end{aligned}
$$

which is precisely the parametrized equation of a cycloid, the curve traced out by a point on the boundary of a rolling wheel. Note this is independent from the mass or gravitational field. It turns out that there is exactly one cycloid which passes through $(0,0)$ and (X,Y) with $Y \geq 0$ such that there are no maxima on the curve between the two points, and which passes through $(0,0)$ with an infinite gradient.

3.7 The Second Variation

One final natural extension to the ideas we have developed in the calculus of variations is to consider the second term in the Taylor expansion of $I[y]$ - we can develop a way for testing whether a solution of the Euler-Lagrange equations is a (global) minimizer of

$$I[y] = \int_a^b f(x,y,y')\,\mathrm{d}x$$

In general, there are two possible approaches:

- Use properties of the function f: recall we used the convexity of $f(y')$ in Example 3.11, where we showed that straight line was the geodesic in the Euclidean plane.
- Look for the second-order term in the Taylor expansion of $I[y]$, and generalize the condition $f''(\mathbf{x}) > 0$.

Of course, the second method does not necessarily guarantee us a *global* minimum - we can only calculate all minima and find the smallest, possibly taking advantage of the shape of the functional. However, it is worth developing this theory.

Recall that for a function $h \in C^2(\mathbb{R}^n)$, Taylor's theorem tells us that for any $\epsilon > 0$ there is a $\delta > 0$ such that

$$\left| h(\mathbf{x}+\delta\mathbf{x}) - h(\mathbf{x}) - \boldsymbol{\nabla} h(\mathbf{x})\cdot\delta\mathbf{x} - \frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2 h}{\partial x_i \partial x_j}\delta x_i \delta x_j \right| \leq \epsilon \|\delta\mathbf{x}\|^2$$

for all $\|\delta\mathbf{x}\| < \delta$. Then it follows that

(i) if $\boldsymbol{\nabla} h(\mathbf{x}) = \mathbf{0}$, and

$$A_{ij} = \left.\frac{\partial^2 h}{\partial x_i \partial x_j}\right|_{\mathbf{x}}$$

is a *positive definite* matrix, then

$$h(\mathbf{x}+\delta\mathbf{x}) > h(\mathbf{x})$$

for all non-zero $\delta\mathbf{x}$ which are sufficiently small: hence $\mathbf{x}$ is a *strict local minimum.*

(ii) if $\mathbf{x}$ is a *local minimum*, then $\boldsymbol{\nabla} h(\mathbf{x}) = \mathbf{0}$, and A_{ij} is *positive semi-definite.*

To extend this to a functional $I[y]$, let ϕ be a smooth, or more generally C^1, function, with $\phi(a) = \phi(b) = 0$. Here, ϕ corresponds to $\delta\mathbf{x}$. Then assuming as ever that f is sufficiently differentiable,

$$\begin{aligned} f(x, y+\phi, y'+\phi') &= f(x,y,y') + \phi f_y + \phi' f_{y'} \\ &\quad + \frac{1}{2}\left[\phi^2 f_{yy} + 2\phi\phi' f_{yy'} + \phi'^2 f_{y'y'}\right] \\ &\quad + O\left([|\phi| + |\phi'|]^3\right) \end{aligned}$$

where all terms in f on the right-hand side are evaluated at (x, y, y').

Then for all $\epsilon > 0$, there is some $\delta > 0$ such that the remainder is

$$O\left(\epsilon\left(|\phi|^2 + |\phi'|^2\right)\right)$$

whenever

$$\max_{[a,b]}(|\phi| + |\phi'|) < \delta$$

In this case, it is clear that

$$I[y+\phi] = I[y] + D_\phi I[y] + \frac{1}{2}D^2 I[y] + O\left(\epsilon \int_a^b \left(|\phi|^2 + |\phi'|^2\right) \mathrm{d}x\right)$$

where $D_\phi I[y]$ is the *first variation* and $D^2 I[y]$ is the *second variation*:

$$\boxed{D^2 I[y] = \int_a^b \left[\phi^2 f_{yy} + 2\phi\phi' f_{yy'} + \phi'^2 f_{y'y'}\right] \mathrm{d}x}$$

3.7.1 Weak extrema

It is important to note the dependence on $|\phi'|$ of the error term in the above expansion; this is a very different feature to anything we have encountered before. We need to formalize our notions of what precisely a 'small' variation is:

Definition 3.25. Write

$$|\phi|_{C^1} = \max_{[a,b]}(|\phi| + |\phi'|)$$

A curve $y \in C^1$ is a *weak local minimum* for $I[y]$ if $I[y+\phi] \geq I[y]$ for $|\phi|_{C^1}$ sufficiently small. The curve y is a *strict weak local minimum* if the inequality is strict whenever $\phi \not\equiv 0$.

The *weak* terminology refers to the restriction that we have placed on $|\phi'|$ - a weak local minimum might not be a minimum with respect to variations with steep gradients.

However, weakening the definition in this way allows us to state the following theorem:

Theorem 3.26.

(i) *If* $D_\phi I[y] = 0$ *and* $D^2 I[y] \geq c\int_a^b (\phi^2 + \phi'^2)\,\mathrm{d}x$ *for some* $c > 0$*, then* y *is a strict weak local minimum for* I*.*

(ii) *If* y *is a weak local minimum for* I*, then* $D_\phi I[y] = 0$ *and* $D^2 I[y] \geq 0$*.*

Remark. Note that we have had to weaken the condition for a strict weak local minimum also, to require a strictly positive lower bound on $D^2 I[y] / \int_a^b (\phi^2 + \phi'^2)\,\mathrm{d}x$ - heuristically, this is because if we can find variations of a fixed magnitude (with respect to some norm) but for which $D^2 I[y]$ tends to 0, we cannot be certain that higher-order in the terms will not come to dominate the second-order term.

It is important to note that function spaces are infinite dimensional vector spaces, with various norms which are not equivalent. (All Euclidean norms for finite dimensions are essentially the same.) A *(strict) strong local minimum* obeys the same inequalities as above, but for ϕ close to zero with respect to the *supremum* or *infinity* norm on V, $\sup|y(x)|$, rather than a norm like

$$\sum_{k=1}^{r} \sup\left|y^{(k)}(x)\right|$$

or like the one above, where the sum and absolute value supremum are interchanged.

For an example of the application of this theory, consider the following functionals:

Example 3.27. Let

$$I_\pm[y] = \int_0^1 \left(\frac{1}{2}y'^2 \pm 5y^2 + y^3\right)\mathrm{d}x$$

Find and classify a stationary curve for each which satisfies $y(0) = y(1) = 0$.

These have the Euler-Lagrange equation

$$\begin{aligned} \pm 10y + 3y^2 - \frac{\mathrm{d}}{\mathrm{d}x}(y') &= 0 \\ -y'' \pm 10y + 3y^2 &= 0 \end{aligned}$$

One solution to this is simply $y_0 = 0$. Then you may easily check that

$$\begin{aligned} \frac{1}{2}\left[\phi^2 f_{yy} + 2\phi\phi' f_{yy'} + \phi'^2 f_{y'y'}\right] &= \frac{1}{2}\left[\phi^2(\pm 10 + 6y_0) + \phi'^2\right] \\ &= \frac{1}{2}\left[\phi'^2 \pm 10\phi^2\right] \end{aligned}$$

Now for the functional $I_+[y]$, we have a second variation of

$$\int_0^1 \left(\phi'^2 + 10\phi^2\right)\mathrm{d}x$$

which is strictly positive, and satisfies the above condition with $c = 1$: hence for I_+, $y_0 = 0$ is a strict weak local minimum.

For $I_-[y]$, however, we have

$$\int_0^1 \left(\phi'^2 - 10\phi^2\right) \mathrm{d}x$$

and trying $\phi(x) = \sin \pi x$ with $\phi'(x) = \pi \cos \pi x$ we see that this gives

$$\int_0^1 \left(\pi^2 \cos^2 \pi x - 10 \sin^2 \pi x\right) \mathrm{d}x \quad = \quad \frac{\pi^2}{2} - \frac{10}{2} < 0$$

This y_0 is not a weak local minimum for I_-.

Remark. * Sometimes, if investigating general second variations thoroughly, it may be necessary to relate the 'size' of the derivative ϕ' to that of ϕ in some way - otherwise, comparing the magnitude of $D^2 I$ to something like $\int_a^b \left(|\phi|^2 + |\phi'|^2\right) \mathrm{d}x$ may be difficult. This can be done using some special cases of the *Poincaré inequality*, which gives a very general stating that the derivative of a function 'cannot be too small'. In particular, for any function ϕ which is C^1 on some interval $[a,b]$ we obtain:

$$\begin{aligned} \text{if } \phi(a) = 0 : \int_a^b \phi(x)^2 \, \mathrm{d}x &\leq \frac{(b-a)^2}{2} \int_a^b \phi'(x)^2 \, \mathrm{d}x \\ \text{if } \phi(a) = \phi(b) = 0 : \int_a^b \phi(x)^2 \, \mathrm{d}x &\leq \frac{(b-a)^2}{8} \int_a^b \phi'(x)^2 \, \mathrm{d}x \end{aligned}$$

You may like to try to prove these results.

3.7.2 * Sturm-Liouville theory

In general, there is a wide class of problems involving expressions of the form

$$\int_a^b \left(P(x)\,\phi'^2 + Q(x)\,\phi^2\right) \mathrm{d}x$$

where ϕ must satisfy $\phi(a) = \phi(b) = 0$ - in fact, it can be straightforwardly shown by integrating by parts and applying these boundary conditions, that

$$D^2 I[y] = \int_a^b \left[\phi^2 f_{yy} + 2\phi\phi' f_{yy'} + \phi'^2 f_{y'y'}\right] \mathrm{d}x = \int_a^b \left(P(x)\,\phi'^2 + Q(x)\,\phi^2\right) \mathrm{d}x$$

for some suitable choice of the functions P and Q.

The key idea is that we want to consider the effect of all possible variations ϕ - but it far simpler, as we have noted previously, to work with all '*directions*' in which variations can be made: that is, $\frac{\mathrm{d}}{\mathrm{d}t} I[y + t\phi(x)]$ for fixed ϕ. Once more referring to our previous work in finite dimensional situations, where we took partial derivatives along the axes, it is easier to try and investigate behaviour along the vectors of a *basis* for the space of all variations - if the second variation is bounded below[9] by some $c > 0$ 'along all the axes' then the second variation is everywhere positive and bounded below by c.

[9] As mentioned above, this ensures there is not a collection of variations of some fixed magnitude but for which the second variation is arbitrarily small - this could conceivably result, for example, in higher-order terms dominating the expansion for $I[y + t\phi]$.

More accurately, if we find a set of *normalized* basis vectors and the effect of $D^2 I$ along each is to increase the value of I upwards.

There are many suitable bases. This is analogous, for example, to finding the eigenvalues of a linear operator like the Hessian matrix

$$A_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

with a complete eigenbasis - we do not in fact even need to work out the basis explicitly in this case, since e.g. knowing all the eigenvalues of A are positive tells us that the result of moving along any vector in the basis increases the value of f. When we defined a positive definite matrix A, we stated that $\mathbf{v}^T A \mathbf{v} > 0$ for all non-zero vectors $\mathbf{v}$ - in terms of the eigenvalues, it is easy to verify that

$$\lambda_{\min} |\mathbf{v}|^2 \leq \mathbf{v}^T A \mathbf{v} \leq \lambda_{\max} |\mathbf{v}|^2$$

where $\lambda_{\min}, \lambda_{\max}$ are the smallest and largest eigenvalues respectively. We might equivalently write

$$\frac{\mathbf{v}^T A \mathbf{v}}{|\mathbf{v}|^2} \in [\lambda_{\min}, \lambda_{\max}] \;\; \forall \mathbf{v} \neq \mathbf{0} \quad \text{or} \quad \mathbf{v}^T A \mathbf{v} \in [\lambda_{\min}, \lambda_{\max}] \;\; \forall \mathbf{v} : |\mathbf{v}| = 1$$

One important aspect to notice is that, in fact, as we vary the directional vector $\mathbf{v}$, this ratio has stationary points at each eigenvector (a good finite-dimensional optimization exercise) - and the ratio is precisely the eigenvalue at this point. In particular, therefore, the minimum and maximum values of this ratio are exactly $\lambda_{\min}$ and $\lambda_{\max}$, and could be found by considering this as an extremizing problem.

So for our infinite-dimensional variational problem, we could attempt to find stationary points of the ratio of the directional second deritvative $D^2 I [y]$ to $M[\phi] = \int_a^b R(x) \phi(x)^2 \, \mathrm{d}x$, or equivalently (also a useful exercise) where $M[\phi] = C$ is fixed at some arbitrary value. Here, $M[\phi]$ is giving a quantity analogous to the size of the vector displacement $\mathbf{v}$; $R(x)$ is called a *weight function*, and allows for some needed flexibility as discussed in the *Methods* course - it corresponds to tweaking the relative importance of basis vectors in finite dimensions. We will take $R = 1$ here. Hence we wish to investigate stationary points of

$$\frac{D^2 I}{M} = \frac{\int_a^b \left(P(x) \phi'^2 + Q(x) \phi^2\right) \mathrm{d}x}{\int_a^b \phi(x)^2 \, \mathrm{d}x}$$

Associated to this is an Euler-Lagrange equation, with one constraint corresponding to the multiplier λ:

$$L[\phi] \stackrel{\text{def}}{=} -\frac{\mathrm{d}}{\mathrm{d}x} (P\phi') + Q\phi = \lambda\phi$$

This in fact has the special form of a so-called *Sturm-Liouville eigenvalue problem* (the theory of which is developed in the *Methods* course) - we have defined a Sturm-Liouville operator $L[\phi]$.

This type of operator has an infinite sequence of typically discrete *eigenvalues* given by the λ_n in

$$L\phi_n = \lambda_n \phi_n$$

(these functions arise as the family of solutions to the variational problem) and we get a condition similar to that for a minimum if $\lambda_n \geq c > 0$ for all n, since then

$$D^2 I \geq c \int_a^b \phi(x)^2 \, \mathrm{d}x$$

Of course, one cannot immediately relate this to the quantity

$$\int_a^b \left(\phi^2 + \phi'^2\right) \mathrm{d}x$$

so this is not always very useful.