MA209 Variational Principles

June 3, 2013

The course covers the basics of the calculus of variations, and derives the Euler-Lagrange equations for minimising functionals of the type $I(y) = \int f(x, y, y')dx$. It then gives examples of this in physics, namely optics and mechanics. It furthermore considers constrained motion and the method of Lagrange multipliers. Required is a basic understanding of differentiation many dimensions, together with a knowledge of how to solve ODEs.

Contents

These notes are based on the 2011 `MA209 Variational Principals` course, taught by J.H.Rawnsley, typeset by Matthew Egginton. No guarantee is given that they are accurate or applicable, but hopefully they will assist your study. Please report any errors, factual or typographical, to m.egginton@warwick.ac.uk

1 Review of Calculus

1.1 Functions of One Variable

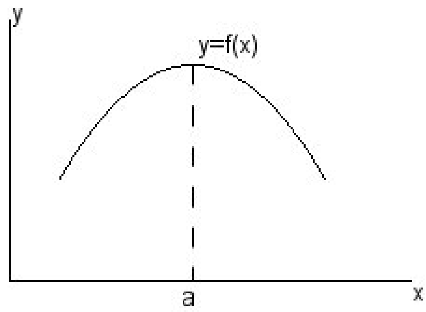

Figure 1: Graph showing a maximum at $x = a$

Suppose that $x = a$ is a maximum of f. Then the graph of f bears some resemblance to that in figure 1. Suppose that f is differentiable and that $f'(a) \neq 0$. Then we either have that $f'(a) > 0$ or $f'(a) < 0$. Consider the former. $f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$ and if $h > 0$ we have that $f(a+h) - f(a) > 0$ for h small and so $f(a+h) > f(a)$, but as $f(a)$ is a maximum, $f'(a) > 0$ must be impossible. A similar argument shows that $f'(a) < 0$ is impossible. Hence our original assumption is false, and so $f'(a) = 0$.

However, there are functions f with $f'(a) = 0$ at values of a which aren't extrema, for example $f(x) = x^3$. We call a point a where $f'(a) = 0$ a **critical point** of the function f and we have shown that the set of extrema is a subset of the set of critical points. This is also true for the set of local extrema.

Example 1.1 *Let $f(x) = ax^2 + bx + c$ with $a \neq 0$. Then $f'(x) = 2ax + b$ with $\frac{-b}{2a}$ the only critical point. $f(x) - f\left(\frac{-b}{2a}\right) = ax^2 + bx + c - a\left(\frac{-b}{2a}\right)^2 - b\left(\frac{-b}{2a}\right) - c = ax^2 + bx - \frac{b^2}{4a} + \frac{b^2}{2a} = a\left(x + \frac{b}{2a}\right)^2$ and so for $a > 0$ is a minimum and $a < 0$ a maximum.*

In general, this won't be so pretty, but for "nice" functions with Taylor series we have $f(a+h) - f(a) = hf'(a) + \frac{h^2}{2!}f''(a) + \ldots$ and so if $f''(a) \neq 0$ we can decide if $f''(a) > 0$ whence we have a local minimum and if $f''(a) < 0$ we have a local maximum.

1.2 Functions of Several Variables

We will look at the two variable case. Consider $f(x_1, x_2)$ a differentiable function with extremum (a_1, a_2). Pick functions $x_1(t)$ and $x_2(t)$ such that $x_1(0) = a_1$ and $x_2(0) = a_2$. Set $g(t) = (x_1(t), x_2(t))$. Then g takes some values of f and at $t = 0$ is an extremum of f and hence g. Thence $g'(0) = 0$ and so if (a_1, a_2) is an extremum then we have that

$$\frac{d}{dt}(f(x_1(t), x_2(t)))\bigg|_{t=0} = 0 \tag{1}$$

for any pair of functions $(x_1(t), x_2(t))$ passing through (a_1, a_2). Thus by the chain rule we have that

$$\frac{\partial f}{\partial x_1}(a_1, a_2)\frac{dx_1}{dt}(0) + \frac{\partial f}{\partial x_2}(a_1, a_2)\frac{dx_2}{dt}(0) = 0$$

As this is true for arbitrary functions, we must have that $\frac{\partial f}{\partial x_1}(a_1, a_2) = 0 = \frac{\partial f}{\partial x_2}(a_1, a_2)$. Note that we could have picked functions with independent derivatives at $t = 0$ specifically.

For n variables, $f(\underline{x})$ real valued and with an extremum at $\underline{x} = \underline{a}$, we pick a function $g_{\underline{v}}(t) = f(\underline{a} + t\underline{v})$, where $\underline{v}$ is an arbitrary vector. Then this will have an extremum at $t = 0$ so $g'_{\underline{v}}(0) = 0$ for all $\underline{v}$ and so $\nabla f(\underline{a}) \cdot \underline{v} = 0$ for all $\underline{v}$ so $\nabla f(\underline{a}) = 0$. If $t = 0$ is a local maximum of $g_{\underline{v}}$ then $\underline{a}$ is a local maximum of f. Then $g''_{\underline{v}}(0) = \sum_{ij} v_i v_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{a}) = \text{Hess}_f(\underline{a}) < 0$ Then all eigenvalues of $\text{Hess}_f(\underline{a})$ must be negative. If they have mixed signs or are zero then we can deduce nothing.

Example 1.2 *Suppose that $f(x, y) = ax^2 + bxy + cy^2$. Then $\nabla f = (2ax + by, bx + 2cy) = (0, 0)$ for an extrema. Thus if $4ac - b^2 \neq 0$ then $x = 0 = y$ is the only critical point.*

$$\sum_{ij} v_i v_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{a}) = v_1^2 2a + 2bv_1 v_2 + 2cv_2^2$$

and so if $a \neq 0$ we get this equal to $2\left(a\left(v_1 + \frac{b}{2a}v_2\right)^2 + \left(c - \frac{b^2}{4a}\right)v_2^2\right)$ and so we have a maximum or minimum when $a\left(c - \frac{b^2}{4a}\right) > 0$ and $4ac > b^2$.

2 Variational Problems

In order to motivate the study of Variational Principles we give some examples of famous problems in the subject.

1. Suppose that y is a function such that $y(x_1) = y_1$ and $y(x_2) = y_2$. We want to find y with the shortest length. The length $L(y)$ is given by

$$L(y) = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx$$

   We say that L is a "functional" of the function y

2. **Brachistochrone**. Suppose that we have a bead of mass m sliding down a frictionless wire under gravity along a curve from (x_1, y_1) to (x_2, y_2). Let $T(y)$ be the time taken to go from (x_1, y_1) to (x_2, y_2) along the curve y. We want to find a minimum of this. If the time is t_1 at (x_1, y_1) and t_2 at (x_2, y_2), and we denote by s the arclength parametrisation, then

$$T(y) = t_2 - t_1 = \int_{t_1}^{t_2} dt = \int_{s_1}^{s_2} \frac{ds}{\frac{ds}{dt}} = \int_{s_1}^{s_2} \frac{ds}{v} = \int_{x_1}^{x_2} \frac{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}}{v} dx$$

   We can find the velocity v from conservation of energy. We know that $E = \frac{1}{2}mv^2 + mg(y(x) - y_1) = \frac{1}{2}mv_1^2 + 0$ if the initial speed is v_1. If we set $v_1 = 0$ then $v = \sqrt{2g(y_1 - y(x))}$ and so

$$T(y) = \int_{x_1}^{x_2} \frac{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}}{\sqrt{2g(y_1 - y(x))}} dx$$

3. **Least area of revolution** Take a curve y with $y(x_1) = y_1$ and $y(x_2) = y_2$ and rotate it about the x-axis. One then gets a surface of revolution around the x-axis. We want to find the curve for which the surface area is as small as possible. The surface area is equal to

$$A(y) = 2\pi \int_{x_1}^{x_2} y\sqrt{1 + (y')^2} dx$$

3 Derivation of the Euler Lagrange Equations

3.1 The one variable - one derivative case

The problems in section 2 involve minimising functionals built from a function of one variable by integration of the function and its derivatives with values of the function specified at the ends of the range of integration. These are typically called fixed endpoint problems. In general, the class of problems of this kind have a functional of the form

$$I(y) = \int_{x_1}^{x_2} f(x, y(x), y'(x))dx \tag{2}$$

for $y(x)$ with $y(x_1) = y_1$ and $y(x_2) = y_2$. In future I will write y for $y(x)$ and y' for $y'(x)$ to simplify the notation.

How do we find extrema of $I(y)$? We proceed in a similar manner to finding conditions for functions at extrema. We consider a one parameter family of functions y_t with y_0 the extremising function. Clearly they all have the same fixed endpoints. Then if $g(t) = I(y_t)$ we have $g'(0) = 0$ or $\frac{d}{dt}I(y_t)\,|_{t=0} = 0$. If $y_t = y_0 + tv$ then $v(x_1) = 0 = v(x_2)$ Hence

$$\frac{d}{dt}I(y_0 + tv)\,|_{t=0} = 0 \tag{3}$$

for v as defined above. The solutions to this equation are called critical points of $I(y)$.

Example 3.1 *Consider*

$$I(y) = \int_0^1 [xy^2 + (y')^2]dx$$

We then have

$$I(y_0 + tv) = \int_0^1 [x(y_0 + tv)^2 + (y_0' + tv')^2]dx$$

and so

$$\frac{d}{dt}I(y_0 + tv)\bigg|_{t=0} = \int_0^1 [x(y_0 + tv)^2 + (y_0' + tv')^2]dx\bigg|_{t=0} = \int_0^1 (2xy_0v + 2y'v')dx$$

y_0 *is a critical point if the integral is* 0 *for all* v *with the conditions as above.*

In the general case $I(y_0+tv) = \int_{x_1}^{x_2} f(x, y_0 + tv, y_0' + tv')dx$ and so if we proceed formally we get

$$\begin{aligned}\frac{d}{dt}I(y_0 + tv)\bigg|_{t=0} &= \frac{d}{dt}\int_{x_1}^{x_2} f(x, y_0 + tv, y_0' + tv')dx\bigg|_{t=0} \\ &= \int_{x_1}^{x_2} \frac{\partial f}{\partial y}(x, y_0, y_0')v + \frac{\partial f}{\partial y'}(x, y_0, y_0')v'dx\end{aligned}$$

If y_0 is a critical point and $v(x)$ is any suitable function with $v(x_1) = 0 = v(x_2)$ then we have from equation (3)

$$\begin{aligned}0 &= \int_{x_1}^{x_2} \frac{\partial f}{\partial y}(x, y_0, y_0')v + \frac{\partial f}{\partial y'}(x, y_0, y_0')v'dx \\ &= \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right] vdx + \frac{\partial f}{\partial y'}v\bigg|_{x_1}^{x_2} \\ &= \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right] vdx\end{aligned}$$

and hence we want to solve

$$\int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right] v dx = 0$$

for suitable v.

We now make rigorous sense of this, and so we need f and its partial derivatives up to order two and y_0'' to be continuous. Then $\frac{\partial f}{\partial y} - \frac{d}{dt}\left(\frac{\partial f}{\partial y'}\right)$ is continuous. We also need $y_0 + tv$ to be a family of functions in a suitable space and so v must have two continuous derivatives.

Theorem 3.1 (The Fundamental Theorem of the Calculus of Variations) *If $u(x)$ is continuous on $[x_1, x_2]$ and*

$$\int_{x_1}^{x_2} u(x)v(x)dx = 0$$

for all $v(x)$ with two continuous derivatives and $v(x_1) = 0 = v(x_2)$ then $u(x) = 0$ for all $x \in [x_1, x_2]$.

Proof We use a contradiction argument. Suppose there is some point $x_0 \in (x_1, x_2)$ with $u(x_0) \neq 0$. Without loss of generality we can assume that $u(x_0) > 0$. IF not, consider the function $-u$. Then $u(x)$ is non zero on some interval around x_0 (positive here even), as u is continuous. Call this interval (x_1', x_2'). Suppose we have $v(x)$ with two continuous derivatives and $v(x) = 0$ where $x \notin [x_1', x_2']$. Then

$$0 = \int_{x_1}^{x_2} u(x)v(x)dx = \int_{x_1'}^{x_2'} u(x)v(x)dx$$

If furthermore $v(x) > 0$ for any $x \in (x_1', x_2')$ then we have that

$$\int_{x_1'}^{x_2'} u(x)v(x)dx > 0$$

This is a contradiction; hence there is no point x_0 where $u(x_0) \neq 0$ and so $u(x) = 0$ for all $x \in [x_1, x_2]$. Thus the proof is reduced to a construction of such a function $v(x)$. A suitable function would be

$$v(x) = \begin{cases} 0 & x \notin (x_1', x_2') \\ (x - x_1')^3(x - x_2')^3 & x \in (x_1', x_2') \end{cases}$$

Q.E.D.

Remark If functionals have more derivatives then this argument could be modified for those. We simply take one higher power than the derivatives.

Aside If we need infinitely many derivatives, we can use $e^{\frac{-1}{x^2}}$ as it has infinitely many derivatives at $x = 0$ and they are all equal to zero.

Theorem 3.2 *If f is a function of three variables with all partial derivatives up to order two continuous then any critical point y of $I(y) = \int_{x_1}^{x_2} f(x, y(x), y'(x))dx$ on the set of functions with two continuous derivatives and satisfying endpoint conditions $y(x_1) = y_1$ and $y(x_2) = y_2$ has*

$$\boxed{\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0 \ \forall x \in [x_1, x_2]} \tag{4}$$

Proof We showed above that

$$\int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right] v dx = 0$$

for all v with two continuous derivatives. The expression in the square brackets is continuous and so by the fundamental theorem (theorem 3.1) must be zero $\forall x \in [x_1, x_2]$ *Q.E.D.*

Definition 3.1 *If a functional* $I(y) = \int_{x_1}^{x_2} f(x, y(x), y'(x))dx$ *then* f *is called the* **Lagrangian** *of* I *and* $\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0$ *is called the* **Euler-Lagrange equation** *of* I

Remark The E-L equation is a second order ODE for $y(x)$ with endpoint conditions.

3.2 Solutions of some examples

Example 3.2 *Find the E-L equation for* $I(y) = \frac{1}{2}\int_0^{\pi} (y^2 - (y')^2)dx$. *We have that* $\frac{\partial f}{\partial y} = y$ *and* $\frac{\partial f}{\partial y'} = -y'$ *and so the E-L equation is* $y - \frac{d}{dx}(y') = 0$ *giving* $y'' + y = 0$

We now solve the examples in section 2.

1. We have from before that

$$L(y) = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx$$

and so $\frac{\partial f}{\partial y} = 0$ and $\frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{1+(y')^2}}$. The E-L equation then gives

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

and so $\frac{y'}{\sqrt{1+(y')^2}}$ is constant, hence $y' = m$ giving the line $y = mx + a$

Remark Any case where $\frac{\partial f}{\partial y} = 0$ will have an immediate integral of the E-L equation $-\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0$ as $\frac{\partial f}{\partial y'} = constant$. We call this a **first integral** of the E-L equation.

Before looking at the other two examples, we note that x does not appear explicitly so we ask if there is a first integral. Observe that

$$\begin{aligned}\frac{d}{dx}\left(y'\frac{\partial f}{\partial y'} - f\right) &= y''\frac{\partial f}{\partial y'} + y'\frac{d}{dx}\frac{\partial f}{\partial y'} - \left(\frac{\partial f}{\partial x} + y'\frac{\partial f}{\partial y} + y''\frac{\partial f}{\partial y'}\right) \\ &= y'\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) - \frac{\partial f}{\partial x} - y'\frac{\partial f}{\partial y}\end{aligned}$$

and if y is a solution of the E-L equations we have that

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

and so if f is independent of x then $y'\frac{\partial f}{\partial y'} - f$ is a constant. This is called the **first integral** for the case of a Lagrangian independent of x.

2. **Brachistochrone** We have $f(x, y, y') = \frac{\sqrt{1+(y')^2}}{\sqrt{(y_1-y)}}$ if we ignore the constants. There is no x dependence here and so $y'\frac{\partial f}{\partial y'} - f$ is a constant.

$$y'\frac{\partial f}{\partial y'} - f = \frac{y'y'}{\sqrt{1+(y')^2}\sqrt{(y_1-y)}} - \frac{\sqrt{1+(y')^2}}{\sqrt{(y_1-y)}} = A$$

giving $\frac{-1}{\sqrt{1+(y')^2}\sqrt{(y_1-y)}} = A$ and hence $(1+(y')^2)(y_1-y) = \frac{1}{A^2}$ and we thus get

$$y' = \pm\sqrt{\frac{1}{A^2(y_1-y)} - 1}$$

If we now make the substitution $A^2(y_1 - y) = \sin^2\frac{\theta}{2}$ then we get that $-A^2y' = \sin\frac{\theta}{2}\cos\frac{\theta}{2}\theta'$ and we get that

$$-\frac{1}{A^2}\sin\frac{\theta}{2}\cos\frac{\theta}{2}\theta' = \pm\sqrt{\frac{1-\sin^2\frac{\theta}{2}}{\sin^2\frac{\theta}{2}}} = \pm\frac{\cos\frac{\theta}{2}}{\sin\frac{\theta}{2}}$$

and so $-\frac{1}{A^2}\sin^2\frac{\theta}{2}\theta' = \pm 1$ giving $-\frac{1}{2A^2}(1-\cos\theta)\theta' = \pm 1$ and then integrating gives

$$-\frac{1}{2A^2}(\theta - \sin\theta) = B \pm x$$

which implicitly determines $\theta(x)$ and so $y(x)$

This curve is called a **cycloid**. Figure 2 shows such a curve.

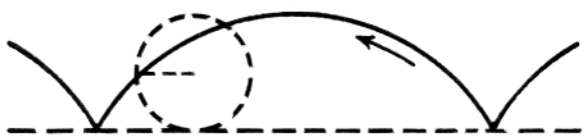

Figure 2: A cycloid

3. $f(x, y, y') = 2\pi y\sqrt{1+(y')^2}$ and observe that we have no x dependence again. Thus we look at the first integral:

$$y'\frac{\partial f}{\partial y'} - f = \frac{(y')^2 y}{\sqrt{1+(y')^2}} - y\sqrt{1+(y')^2} = A$$

and so we get that

$$\frac{-y}{\sqrt{1+(y')^2}} = A$$

and so $y' = \pm\sqrt{\frac{y^2}{A^2} - 1}$. If we then make the substitution $\frac{y}{A} = \cosh z$ we get that $\frac{y'}{A} = \sin z z'$ and hence the equation to solve becomes

$$A\sinh z z' = \pm\sqrt{\cosh^2 z - 1} = \pm\sinh z$$

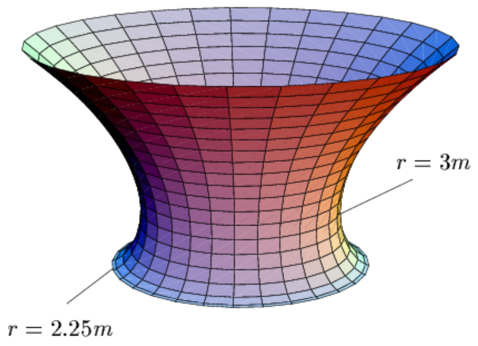

Figure 3: The shape of surface which minimises the surface of revolution

and thus we get that $z' = \pm\frac{1}{A}$ and so $z = B \pm \frac{x}{A}$ and so

$$y = A\cosh\left(B \pm \frac{x}{A}\right)$$

and so it looks like figure 3

We now try to fit this shape of solution to the endpoint conditions. Without loss of generality we will assume that $y = A\cosh\left(B' + \frac{x}{A}\right)$, and we want a solution with $y(x_1) = y_1$ and $y(x_2) = y_2$. Using the first of these we get that $B' = \cosh^{-1}\left(\frac{y_1}{A}\right) - \frac{x_1}{A}$ and then $y = y_1\cosh\left(\frac{x-x_1}{A}\right) + \sqrt{y_1^2 - A^2}\sinh\left(\frac{x-x_1}{A}\right)$ and using the second condition gives a pretty nasty equation (I leave to the reader to work it out). To see if solutions exist we plot the graph of $y(x)$ for various values of A. Thus from this graph you can see that if (x_2, y_2) is to the right of the dotted line then there is no solution. Also note that if (x_2, y_2) is above the dotted line then there are two solutions. Also these solutions may not be extrema, as a broken line may well minimise the problem.

Remark $y_0 + tv$ is called a **variation** of y_0, hence the name Calculus of Variations

3.3 Extension of the Theory

3.3.1 More Derivatives

Suppose that

$$I(y) = \int_{x_1}^{x_2} f(x, y, y', ..., y^{(n)})dx$$

We try the same method as before, considering $I(y + tv)$ for y an extremum. Set $g(t) = T(y + tv)$ and then this has an extremum at $t = 0$ so $g'(0) = 0$ and thus $\frac{d}{dt}I(y + tv)\big|_{t=0} = 0$ and so

$$\frac{d}{dt}I(y+tv)\bigg|_{t=0} = \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y}v + \frac{\partial f}{\partial y'}v' + ... + \frac{\partial f}{\partial y^{(n)}}v^{(n)}\right]dx = 0$$

If we assume that $v(x_1) = 0 = v(x_2)$ and all partial derivatives up to $v^{(n-1)}$ are zero at x_1 and x_2 then we get that

$$\frac{d}{dt}I(y+tv)\,|_{t=0} = \int_{x_1}^{x_2}\left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) + ... + (-1)^n\frac{d^n}{dx^n}\left(\frac{\partial f}{\partial y^{(n)}}\right)\right]vdx = 0$$

For the argument to be complete we need f to have $(n+1)$ continuous derivatives and y to have $2n$ continuous derivatives. Then the term in square brackets is continuous and we need the version of the fundamental theorem for v with $2n$ continuous derivatives. Then for y an extremum it satisfies

$$\boxed{\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) + ... + (-1)^n \frac{d^n}{dx^n}\left(\frac{\partial f}{\partial y^{(n)}}\right) = 0} \tag{5}$$

This is again called the **Euler Lagrange equation** for the functional. There is no existence or uniqueness theorem in this case again.

Example 3.3 *Suppose* $I(y) = \int_0^{\frac{\pi}{2}} ((y'')^2 - y^2)dx$ *with* $y(0) = 0 = y'(0)$ *and* $y(\frac{\pi}{2}) = 1$ *and* $y'(\frac{\pi}{2}) = 0$. *The E-L equation gives* $-2y + \frac{d^2}{dx^2}(2y'') = 0$ *and so* $y^{(4)} - y = 0$ *and this has a general solution of* $y = A\cos x + B\sin x + Ce^x + De^{-x}$ *and solving for the endpoint conditions gives the four equations* $0 = A + C + D$, $0 = B + C - D$, $1 = B + Ce^{\frac{\pi}{2}} + De^{-\frac{\pi}{2}}$ *and* $0 = -A + Ce^{\frac{\pi}{2}} - De^{-\frac{\pi}{2}}$ *and these can be solved.*

3.3.2 Several dependent functions

Problems involving curves may not be expressible as $y = y(x)$ and so instead we could write the curve in parametric form, i.e. for the length problem we could write $L(x,y) = \int_{t_1}^{t_2} \sqrt{(x')^2 + (y')^2}dt$. In general these have the form

$$I(x,y) = \int_{t_1}^{t_2} f(t, x(t), y(t), x'(t), y'(t))dt$$

and we use a one parameter variation $(x + hu, y + hv)$. Then (x,y) is an extremum of I means that

$$\frac{d}{dh} I(x + hu, y + hv)\bigg|_{h=0} = 0$$

Note that u and v must vanish at the endpoints to preserve the endpoint conditions.

If we first take $v(x) = 0\ \forall x \in [t_1, t_2]$ then $\frac{d}{dh} I(x + hu, y)\big|_{h=0} = 0$ and so

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

Similarly if $u(x) = 0\ \forall x \in [t_1, t_2]$ then $\frac{d}{dh} I(x, y + hv)\big|_{h=0} = 0$ and so

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

In other words both x and y satisfy the Euler Lagrange equation for one variable.

One can also derive these two equations as we did before: using the chain rule on the necessary condition, then integrating by parts. Then taking $v = 0$ and then $u = 0$ we can apply the fundamental theorem in both cases, giving the result above.

It should be clear that this works for any number of independent variables, so long as they can be varied independently. If

$$I(x_1, ..., x_n) = \int_{t_1}^{t_2} f(t, x_1(t), ..., x_n(t), x_1'(t), ..., x_n'(t))dt$$

then I has n simultaneous E-L equations

$$\boxed{\frac{\partial f}{\partial x_i} - \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}_i}\right) = 0\ \forall i = 1, ..., n} \tag{6}$$

Example 3.4 *Suppose that* $L(x,y) = \int_0^1 \sqrt{(x')^2 + (y')^2}dt$ *with* $x(0) = x_1$ *and* $x(1) = x_2$ *as well as* $y(0) = y_1$ *and* $y(1) = y_2$*. This has two E-L equations:*

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

and so both $\frac{x'}{\sqrt{(x')^2+(y')^2}}$ *and* $\frac{y'}{\sqrt{(x')^2+(y')^2}}$ *are constants. Thus* $\frac{1}{\sqrt{(x')^2+(y')^2}}(x', y') = (A, B)$ *is a constant unit vector. Hence* $(x(t), y(t))$ *is a curve with a constant direction. If* $c(t) = \sqrt{(x')^2 + (y')^2}$ *then* $(x, y) = d(t)(A, B) + (C, D)$ *where* $d' = c$

Remark Observe that although it is written in term of two variables, the problem is degenerate. It has infinitely many solutions given by different possible functions $d(t)$.

If there is no explicit t dependence, i.e. $\frac{\partial f}{\partial t} = 0$, then consider

$$F(t) = x_1'\frac{\partial f}{\partial x_1'} + ... + x_n'\frac{\partial f}{\partial x_n'} - f$$

Then

$$\begin{aligned}\frac{dF}{dt} &= x_1''\frac{\partial f}{\partial x_1'} + x_1'\frac{d}{dt}\left(\frac{\partial f}{\partial x_1'}\right) + ... \\ &\quad + x_n''\frac{\partial f}{\partial x_n'} + x_n'\frac{d}{dt}\left(\frac{\partial f}{\partial x_1'}\right) - \frac{\partial f}{\partial t} - x_1'\frac{\partial f}{\partial x_1} - \\ &\quad ... - x_n'\frac{\partial f}{\partial x_n} - x_1''\frac{\partial f}{\partial x_1'} - ... - x_n''\frac{\partial f}{\partial x_n'} \\ &= 0\end{aligned}$$

if there is no explicit time dependence and $x_1, ..., x_n$ satisfy the E-L equations. Hence F is constant and this is another First Integral.

4 Relationship with Optics and Fermat's Principle

We look here at rays of light in the plane moving with speed $c(x, y)$

4.1 Fermat's Principle

Theorem 4.1 (Fermat's Principle) *Light Travels along a path between two points* (x_1, y_1) *and* (x_2, y_2) *so as to take the least time to get from* (x_1, y_1) *to* (x_2, y_2)

$c(x, y)$ is the speed at (x, y), and if we travel along a path the speed will be the rate of change of arclength along the path. Thus if we measure arclength s from an initial position, then $\frac{ds}{dt} = c(x, y)$. If the path is a graph of a function $y(x)$ then from a path from (x_1, y_1) to (x_2, y_2), where we are at (x_1, y_1) at time t_1 and arclength s_1 and at (x_2, y_2) at time t_2 and arclength s_2, we get that

$$T(y) = t_2 - t_1 = \int_{t_1}^{t_2} dt = \int_{s_1}^{s_2} \frac{ds}{\frac{ds}{dt}} = \int_{s_1}^{s_2} \frac{ds}{c} = \int_{x_1}^{x_2} \frac{\sqrt{1 + (y')^2}}{c(x, y)}dx$$

The actual path followed by a light ray will be a minimum of $T(y)$.

Example 4.1 Light in a homogeneous medium *Here we assume that c is a constant. We have that*

$$T(y) = \frac{1}{c}\int_{x_1}^{x_2} \sqrt{1+(y')^2}dx = \frac{1}{c}L(y)$$

and hence in a homogeneous medium light travels in straight lines since these are critical points of the length functional.

Example 4.2 The Law of Refraction *Suppose that we have two homogeneous media with speeds c_1 and c_2 and have a straight line interface and a ray of light from the first to the second. We know that we will have a broken line, but what is the change in direction at the interface. We look at broken straight line paths passing through the point$(x_0, 0)$ on the x-axis. The time taken is $\tau(x_0)$ and this is equal to*

$$\tau(x_0) = \frac{\sqrt{(x_0-x_1)^2+y_1^2}}{c_1} + \frac{\sqrt{(x_2-x_0)^2+y_2^2}}{c_2}$$

The actual path will be a minimum with respect to x_0 and so at that point where the path crosses the x-axis we have $\frac{d\tau}{dx_0} = 0$. Now $\frac{d\tau}{dx_0} = \frac{x_0-x_1}{c_1\sqrt{(x_0-x_1)^2+y_1^2}} - \frac{x_2-x_0}{c_2\sqrt{(x_2-x_0)^2+y_2^2}} = 0$ whence $\frac{\sin\theta_1}{c_1} - \frac{\sin\theta_2}{c_2} = 0$ or

$$\boxed{\frac{\sin\theta_1}{\sin\theta_2} = \frac{c_1}{c_2}} \tag{7}$$

This is known as **Snell's Law**.

Suppose that c is only a function of y, i.e. that $c(x,y) = c(y)$. We divide into strips parallel to the x-axis. In each strip, the path is approximated by a straight line segment. Then the slope in the strip will be approximately $\frac{dy}{dx}$. We then have that $\cot\theta = \frac{dy}{dx} = y'$ and then $\sin\theta = \frac{1}{\sqrt{1+(y')^2}}$. It is $\cot\theta$ here because θ is the angle the ray makes with the y direction. According to Snell's Law $\frac{\sin\theta}{c}$ is a constant and so $\frac{1}{c(y)}\frac{1}{\sqrt{1+(y')^2}}$ is a constant. This equation gives $\sqrt{1+(y')^2} = \frac{1}{Kc(y)}$ and so $y' = \pm\sqrt{\frac{1}{K^2c^2(y)} - 1}$. Then dividing by the square root term and integrating with respect to x gives

$$\int \frac{dy}{\sqrt{\frac{1}{K^2c^2(y)} - 1}} = A \pm x$$

This gives an equation for x as a function of y and by solving, or using a substitution, we get an explicit solution.

We now rework the above using the Calculus of Variations. In this case we have a functional independent of x as

$$T(y) = \int_{x_1}^{x_2} \frac{\sqrt{1+(y')^2}}{c(y)}dx$$

and then this has a first integral of

$$y'\frac{y'}{c(y)\sqrt{1+(y')^2}} - \frac{\sqrt{1+(y')^2}}{c(y)} = K$$

and this gives $\frac{-1}{c(y)\sqrt{1+(y')^2}} = -K$ which we deduced from Snell's law before. Hence the first integral of Fermat's Principle is Snell's Law.

4.2 Optical Analogy

If a problem in the Calculus of Variations leads to a functional of the same form as that coming from Fermat's Principle and the optical problem is already solved then the same solution applies to the variational problem. It then has the solution, when independent of x in the functional, given by

$$A \pm x = \int \frac{dy}{\sqrt{\frac{1}{K^2c^2(y)} - 1}}$$

This was how Bernoulli first solved the Brachistochrone problem, where we have

$$\int_{x_1}^{x_2} \frac{\sqrt{1+(y')^2}}{\sqrt{2g(y_1-y)}}dx$$

as our functional. If we take $c(y) = \sqrt{2g(y_1-y)}$ then we can write down the integral formula for the solution.

When x appears explicitly we have to go to the full E-L equations.

5 Hamilton's Principle

Suppose that $\underline{x}(t) = (x(t), y(t), z(t))$ describes the motion of a point particle in three dimensions where t is the time variable. We define $\dot{\underline{x}} := \frac{d\underline{x}}{dt}$ and call it the velocity $\underline{v}$. Furthermore we define $\ddot{\underline{x}} := \frac{d^2\underline{x}}{dt^2}$ and call it the acceleration. $v := |\underline{v}| := \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} = \sqrt{\underline{v}\cdot\underline{v}}$ is called the speed. The motion is governed by the mass $m > 0$. The kinetic energy is $\frac{1}{2}mv^2 = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$. If we have many particles then the total kinetic energy is $T = \sum_i \frac{1}{2}m_1 v_i^2$. If $q_1, ..., q_n$ is a different set of coordinates of which $x_1, y_1, z_1, x_2, y_2, z_2, ...$ are functions then we get T as a function of $q_1, ..., q_n, \dot{q_1}, ..., \dot{q_n}$ by substitution.

Definition 5.1 *A* **conservative system** *is where the forces acting $\underline{F}$ can be given in terms of a function V such that $\underline{F} = -\nabla V$. V is called the* **potential energy** *and is a function of $q_1, ..., q_n$ independent coordinates.*

Definition 5.2 *The* **Lagrangian** *of the system is*

$$L(q_1, ..., q_n, \dot{q_1}, ..., \dot{q_n}) := T - V$$

Example 5.1 *Suppose that a particle of mass m is moving in a circle in the x-y plane with gravity acting in the negative y direction. Then the potential is given by $V := mgy = mgR\sin\theta$ and the kinetic energy is $T = \frac{1}{2}mR^2\dot{\theta}^2$ and so the Lagrangian is $L(\theta, \dot{\theta}) = \frac{1}{2}mR^2\dot{\theta}^2 - mgR\sin\theta$*

Theorem 5.1 (Hamilton's Principle) *The path followed by a system described by a Lagrangian $L = T - V$ in getting from an initial position P_1 at time t_1 to a final position P_2 at time t_2 is a critical point of the functional*

$$\boxed{I = \int_{t_1}^{t_2} L dt}$$

amongst all possible paths from P_1 to P_2 at the relevant times.

Hence the actual path satisfies the E-L equations for L, namely

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q_i}}\right) = 0 \text{ for } i = 1, ..., n \tag{8}$$

Example 5.2 *Suppose we have a particle on a circle of radius R and is acted upon by gravity (see example 5.1). Then we have $L(\theta,\dot{\theta}) = \frac{1}{2}mR^2\dot{\theta}^2 - mgR\sin\theta$ and so the E-L equations for this gives*

$$-mgR\cos\theta - \frac{d}{dt}(mR^2\dot{\theta}) = 0 \implies \ddot{\theta} + \frac{g}{R}\cos\theta = 0$$

and this is called the pendulum equation

Example 5.3 *Suppose that we have a particle of mass m moving in $\mathbb{R}^3$ with a force $\underline{F} = -\nabla V$. Then $L = \frac{1}{2}m(\dot{x}^2+\dot{y}^2+\dot{z}^2) - V(x,y,z)$ and the E-L equations give*

$$\left.\begin{array}{l} \frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = 0 \implies -\frac{\partial V}{\partial x} - m\ddot{x} = 0 \\ \frac{\partial L}{\partial y} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right) = 0 \implies -\frac{\partial V}{\partial y} - m\ddot{y} = 0 \\ \frac{\partial L}{\partial z} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{z}}\right) = 0 \implies -\frac{\partial V}{\partial z} - m\ddot{z} = 0 \end{array}\right\} \implies \underline{F} - m\ddot{\underline{x}} = 0$$

i.e. Newton's Second Law. Thus Hamilton's principle is in accord with Newton's second Law.

Observe that L is independent of the time variable, and so we always have a first integral of the form

$$\dot{q_1}\frac{\partial L}{\partial \dot{q_1}} + ... + \dot{q_n}\frac{\partial L}{\partial \dot{q_n}} - L = constant$$

Observe that the kinetic energy is quadratic in the derivatives, and will be so for any system. Thus

$$T(q_1,...,q_n,\dot{q_1},...,\dot{q_n}) = \sum_{i=1}^{n}\sum_{j=1}^{n}\dot{q_i}\dot{q_j}T_{ij}(q_1,...,q_n)$$

And hence we get the identity

$$T(q_1,...,q_n,a\dot{q_1},...,a\dot{q_n}) = a^2T(q_1,...,q_n,\dot{q_1},...,\dot{q_n}) \tag{9}$$

which is called Euler's Formula. It should be clear that $\frac{\partial L}{\partial \dot{q_i}} = \frac{\partial T}{\partial \dot{q_i}}$ as V is independent of the $\dot{q_1},...,\dot{q_n}$, the first integral becomes

$$\dot{q_1}\frac{\partial T}{\partial \dot{q_1}} + ... + \dot{q_n}\frac{\partial T}{\partial \dot{q_n}} - L = constant$$

and hence, by differentiating (9) with respect to a and setting $a = 0$, we get that $T + V = constant$ and this is called conservation of energy.

6 Constraints and Lagrange Multipliers

6.1 Finite Dimensions

6.1.1 Two dimensions

A typical example is to find extrema of $f(x,y)$ on the set $\{(x,y) \in \mathbb{R}^2 | g(x,y) = 0\}$. The implicit function theorem tells us which variable in an equation can be solved for in terms of the others.

If $\frac{\partial g}{\partial x}(x_0,y_0) \neq 0$ for some point then there is a function $\eta(x)$ defined for x near x_0 with $\eta(x_0) = y_0$, η differentiable, such that all solutions (x,y) of $g(x,y) = 0$ near (x_0,y_0) have the form $(x,\eta(x))$.

We say the constraint is regular if at every solution at least one of the partial derivatives is non-zero.

If (x_0, y_0) is an extremum of f on $\{(x,y)|g(x,y) = 0\}$ then let $y = \eta(x)$ be a solution near (x_0, y_0) of the constraint, and then substitute this in f to give $f(x, \eta(x))$ and this has x_0 as an extremum. Therefore

$$\frac{d}{dx} f(x, \eta(x))\Big|_{x=x_0} = 0 \tag{10}$$

We also have the fact that $g(x, \eta(x)) = 0$ for all x for which η is defined. Equation (10), by the chain rule, yields

$$\frac{\partial f}{\partial x}(x_0, y_0) + \frac{\partial f}{\partial y}(x_0, y_0)\frac{d\eta}{dx}(x_0) = 0$$

and we also have that

$$\frac{d}{dx} g(x, \eta(x)) = 0 \implies \frac{\partial g}{\partial x}(x, y) + \frac{\partial g}{\partial y}(x, y)\frac{d\eta}{dx}(x) = 0$$

for all x near x_0, and if one evaluates this at $x = x_0$ we get that

$$\frac{d\eta}{dx}(x_0) = -\frac{\frac{\partial g}{\partial x}(x_0, y_0)}{\frac{\partial g}{\partial y}(x_0, y_0)}$$

as the denominator is non zero by assumption. From these we get that:

$$\frac{\partial f}{\partial x}(x_0, y_0) - \frac{\partial f}{\partial y}(x_0, y_0)\frac{\frac{\partial g}{\partial x}(x_0, y_0)}{\frac{\partial g}{\partial y}(x_0, y_0)} = 0$$

and if we define $\lambda = \frac{\frac{\partial f}{\partial y}(x_0,y_0)}{\frac{\partial g}{\partial y}(x_0,y_0)}$ then this becomes

$$\frac{\partial f}{\partial x}(x_0, y_0) - \lambda\frac{\partial g}{\partial x}(x_0, y_0) = 0 \implies \frac{\partial}{\partial x}(f - \lambda g)\big|_{(x_0,y_0)} = 0$$

Also, by definition of λ we get that $\frac{\partial}{\partial y}(f - \lambda g)\big|_{(x_0,y_0)} = 0$ and therefore $f - \lambda g$ has a critical point at (x_0, y_0).

Similarly if $\frac{\partial g}{\partial y}(x_0, y_0) \neq 0$ and (x_0, y_0) is an extremum of f on $\{(x,y)|g(x,y) = 0\}$ then there is a constant $\lambda' = \frac{\frac{\partial f}{\partial x}(x_0,y_0)}{\frac{\partial g}{\partial x}(x_0,y_0)}$ such that $f - \lambda' g$ has a critical point at (x_0, y_0).

We have thus proved:

Theorem 6.1 (Lagrange Multiplier) *If g is a regular constraint with $\nabla g \neq 0$ for all (x, y) with $g(x, y) = 0$ then any extremum (x_0, y_0) of $f(x, y)$ on the set $\{(x, y)|g(x, y) = 0\}$ has an associated real number λ such that $f - \lambda g$ has a critical point at (x_0, y_0).*

We call λ the Lagrange multiplier for (x_0, y_0). The unknowns are now (x_0, y_0) and λ. The condition of $f - \lambda g$ having a critical point at (x_0, y_0) is $\nabla(f - \lambda g)(x_0, y_0) = 0$ and we also have the condition of $g(x_0, y_0) = 0$.

Example 6.1 *Find the extrema of $f(x, y) = ax + by$ on $x^2 + y^2 = 1$. The extrema has a critical point of $f - \lambda g = ax + by - \lambda(x^2 + y^2 - 1)$ and so $0 = a - 2\lambda x$ and $0 = b - 2\lambda y$ giving $a^2 + b^2 = 4\lambda^2$ and so $\lambda = \pm\frac{\sqrt{a^2+b^2}}{2}$ and so $(x, y) = \left(\pm\frac{a}{\sqrt{a^2+b^2}}, \pm\frac{b}{\sqrt{a^2+b^2}}\right)$. Then $f = \pm\frac{a^2+b^2}{\sqrt{a^2+b^2}} = \pm\sqrt{a^2 + b^2}$ and hence there is a maximum at $+\sqrt{a^2 + b^2}$ and a minimum at $-\sqrt{a^2 + b^2}$.*

In general there may be more solutions to $\nabla(f - \lambda g)(x_0, y_0) = 0$ and $g(x_0, y_0) = 0$ than there are extrema (x_0, y_0) of f on $\{(x, y)|g(x, y) = 0\}$.

Definition 6.1 *We call the solutions to the above* **constrained critical points** *of f*

The constrained critical points of f on $g(x, y) = 0$ are unconstrained critical points of $f - \lambda g$ for some λ.

6.1.2 n dimensions

Let f be a function of n variables and look for extrema of $f(x_1, ..., x_n)$ on the set of points where a function $g(x_1, ..., x_n) = 0$. Suppose that $\underline{x} = (x_1, ..., x_n)$ is an extreme point and pick two vectors $\underline{u}$ and $\underline{v}$ and consider a function of two variables $F_{\underline{u},\underline{v}}(h,k) := f(\underline{x}+h\underline{u}+k\underline{v})$ subject to the constraints $G_{\underline{u},\underline{v}}(h,k) := g(\underline{x}+h\underline{u}+k\underline{v}) = 0$. Then $(h,k) = (0,0)$ is an extremum of $F_{\underline{u},\underline{v}}$ subject to $G_{\underline{u},\underline{v}}(h,k) = 0$. Hence there is a Lagrange multiplier $\lambda_{\underline{u},\underline{v}}$ such that $F_{\underline{u},\underline{v}} - \lambda_{\underline{u},\underline{v}} G_{\underline{u},\underline{v}}$ has a critical point at $(h,k) = (0,0)$. Therefore we have that

$$\frac{\partial}{\partial h}(F_{\underline{u},\underline{v}} - \lambda_{\underline{u},\underline{v}} G_{\underline{u},\underline{v}}) = 0$$

and that

$$\frac{\partial}{\partial k}(F_{\underline{u},\underline{v}} - \lambda_{\underline{u},\underline{v}} G_{\underline{u},\underline{v}}) = 0$$

This then gives us that

$$\underline{u} \cdot \nabla(f - \lambda_{\underline{u},\underline{v}} g)(\underline{x}) = 0$$

and that

$$\underline{v} \cdot \nabla(f - \lambda_{\underline{u},\underline{v}} g)(\underline{x}) = 0$$

from the fact that $\frac{\partial F_{\underline{u},\underline{v}}}{\partial h} = \frac{\partial}{\partial h}(f(\underline{x}+h\underline{u}+k\underline{v})) = \underline{u} \cdot f(\underline{x}+h\underline{u}+k\underline{v})$ and similarly for the other partial derivatives.

Then for every pair of vectors $\underline{u}$ and $\underline{v}$ we have that $\nabla(f - \lambda_{\underline{u},\underline{v}} g)(\underline{x})$ is perpendicular to both $\underline{u}$ and $\underline{v}$.

If $\underline{e}_1, ..., \underline{e}_n$ is the standard basis and $\lambda_{ij} = \lambda_{\underline{e}_i,\underline{e}_j}$ then $\nabla f(\underline{x}) - \lambda_{ij} \nabla g(\underline{x})$ is perpendicular to both $\underline{e}_i$ and $\underline{e}_j$ for each i and j. In terms of the partial derivatives this becomes

$$\frac{\partial f}{\partial x_i}(\underline{x}) - \lambda_{ij}\frac{\partial g}{\partial x_i}(\underline{x}) = 0 = \frac{\partial f}{\partial x_j}(\underline{x}) - \lambda_{ij}\frac{\partial g}{\partial x_j}(\underline{x}) \ \forall i,j$$

We aim to find a condition that is independent of j and so have a single Lagrange multiplier for each of the equations. For a regular constraint we need $\nabla g \neq 0$ everywhere on $g(\underline{x}) = 0$. Thus at least one partial derivative of g is non zero, say the i_0th. Then we can write

$$\lambda_{i_0 j} = \frac{\frac{\partial f}{\partial x_{i_0}}(\underline{x})}{\frac{\partial g}{\partial x_{i_0}}(\underline{x})}$$

and this is independent of j. Then if we put $\lambda := \lambda_{i_0 j}$ for any j and then input this into the second equation, we get that

$$\boxed{\frac{\partial f}{\partial x_j}(\underline{x}) - \lambda\frac{\partial g}{\partial x_j}(\underline{x}) = 0 \ \forall j}$$

Hence we have a λ such that $\nabla(f - \lambda g)(\underline{x}) = 0$

Example 6.2 *Find the point on the plane $\underline{x} \cdot \underline{n} = p$ closest to a given point $\underline{a}$ not on the plane.*

We aim to minimise the distance from a point $\underline{x}$ to the point $\underline{a}$ such that $\underline{x} \cdot \underline{n} = p$ The Euclidean distance is given by $d(\underline{x},\underline{a}) = |\underline{x}-\underline{a}|$ but we will take the square of this to simplify working out. It should be clear that if the square of the distance has a minimum, then so must the distance itself. Thus we have that we want to minimise $f(\underline{x}) = |\underline{x}-\underline{a}|^2 = \sum_{i=1}^{i=m}(x_i - a_i)^2$ subject to $g(\underline{x}) = \underline{x} \cdot \underline{n} - p$. Now $\nabla f = 2(\underline{x}-\underline{a})$ and $\nabla g = \underline{n}$. At the critical point there is a number λ such that $\nabla f - \lambda g = 0$ and in this case this is $2(\underline{x}-\underline{a}) - \lambda\underline{n} = 0$ and so we get that $\underline{x} = \underline{a} + \frac{\lambda}{2}\underline{n}$ and $(\underline{a} + \frac{\lambda}{2}\underline{n}) \cdot \underline{n} = p$ and so $\frac{\lambda}{2} = p - \underline{a} \cdot \underline{n}$ thus $\underline{x} = \underline{a} + (p - \underline{a} \cdot \underline{n})\underline{n}$.

This is a minimum as any point which is different from the above one will have distance on a hypotenuse of a right angled triangle with one side equal to the length at a critical point.

6.1.3 Examples

The following examples are ones that we aim to solve, and will develop techniques to do so in the next section.

1. **Hanging rope or chain** Suppose we have a rope hanging in equilibrium between two points (x_1, y_1) and (x_2, y_2). What is the shape of the rope? This is called a catenary.

   Suppose the shape is the graph of a function $y = y(x)$. In equilibrium its potential energy will be minimised. Let ρ be the density per unit length of the rope and assume that it is constant. Then the total mass is $J(y) := \rho \int_{x_1}^{x_2} \sqrt{1 + (y')^2} dx =: M$. The potential energy is then given by $I(y) := \rho g \int_{x_1}^{x_2} y\sqrt{1 + (y')^2} dx$. Hence we want to minimise $I(y)$ subject to $J(y)$ being a constant value M.

2. **Isoperimetric problem** Consider a closed curve in the plane. For a given length, we want to find the curve which encloses the greatest area.

   Let the curve C be given by $(x(t), y(t))$ with $x(t_1) = x_0 = x(t_2)$ and $y(t_1) = y_0 = y(t_2)$. Then the length of C is given by $L(C) = \int_{t_1}^{t_2} \sqrt{\dot{x}^2 + \dot{y}^2} dt$ and the area is given by $A(C) = \frac{1}{2} \int_{t_1}^{t_2} (x\dot{y} - y\dot{x}) dt$ We want to minimise $A(C)$ for fixed $L(C)$.

3. **Geodesics on Surfaces** Curves which minimise the distance in a surface are called geodesics. Here we minimise a length functional $L(\underline{x})$ for curves $\underline{x}(t)$ which satisfy $g(\underline{x}(t)) = 0$ for all t.

6.1.4 A functional constrained by a functional

We first look at problems with two functionals (like one and two above), with two parameter variations, and then look at two variation problems.

If $I(y)$ is extremised on the set $g(x)$ with $J(y)$ constant, we look at two parameter variations $y + hu + kv$ where u and v are chosen such that $u(x_1) = u(x_2) = v(x_1) = v(x_2) = 0$ and then $(h, k) = (0, 0)$ is an extremum for $I(y + hu + kv) = J_0$ which is a fixed constant. Define $F_{uv}(h, k) = I(y + hu + kv)$ and $G_{uv}(h, k) = I(y + hu + kv) - J_0$. Then F_{uv} has an extremum at $(h, k) = (0, 0)$ for (h, k) such that $G_{uv}(h, k) = 0$. Hence we have a Lagrange Multiplier λ_{uv} such that $F_{uv} - \lambda_{uv} G_{uv}$ has a critical point at $(0, 0)$. Thus

$$\frac{\partial}{\partial h}(F_{uv}(h, k) - \lambda_{uv} G_{uv}(h, k))\,|_{h,k=0} = 0$$

and also

$$\frac{\partial}{\partial k}(F_{uv}(h, k) - \lambda_{uv} G_{uv}(h, k))\,|_{h,k=0} = 0$$

If $I(y) = \int_{x_1}^{x_2} f(x, y, y') dx$ and $J(y) = \int_{x_1}^{x_2} g(x, y, y') dx$ then the h partial equation gives

$$0 = \int_{x_1}^{x_2} \left(\frac{\partial}{\partial y}(f - \lambda_{uv} g) - \frac{d}{dx} \left(\frac{\partial}{\partial y'}(f - \lambda_{uv} g) \right) \right) u dx$$

and the k partial equation gives

$$0 = \int_{x_1}^{x_2} \left(\frac{\partial}{\partial y}(f - \lambda_{uv} g) - \frac{d}{dx} \left(\frac{\partial}{\partial y'}(f - \lambda_{uv} g) \right) \right) v dx$$

Then we have that

$$0 = \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) \right) u dx - \lambda_{uv} \int_{x_1}^{x_2} \left(\frac{\partial g}{\partial y} - \frac{d}{dx} \left(\frac{\partial g}{\partial y'} \right) \right) u dx$$

The regularity condition gives that the latter integrand and hence integral in the above equation is non zero on the set of $g(x)$ and so $J(y) = J_0$. Then the former integral is non zero and so we can set

$$\lambda_{u_0 v} = \frac{\int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right) u_0 dx}{\int_{x_1}^{x_2} \left(\frac{\partial g}{\partial y} - \frac{d}{dx}\left(\frac{\partial g}{\partial y'}\right)\right) u_0 dx}$$

and note that the right hand side here is independent of v, and so we can write $\lambda_{u_0 v} =: \lambda$. Then for any v vanishing at x_1 and x_2, and for λ defined before, we get that

$$\int_{x_1}^{x_2} \left(\frac{\partial}{\partial y}(f - \lambda g) - \frac{d}{dx}\left(\frac{\partial}{\partial y'}(f - \lambda g)\right)\right) v dx = 0$$

and by the fundamental lemma we get that

$$\frac{\partial}{\partial y}(f - \lambda g) - \frac{d}{dx}\left(\frac{\partial}{\partial y'}(f - \lambda g)\right) = 0$$

This is called the Euler Lagrange equation for this case. We have thus proved:

Theorem 6.2 *An extremum of $I(y)$ subject to $J(y) = J_0$ satisfies the Euler Lagrange equation*

$$\boxed{\frac{\partial}{\partial y}(f - \lambda g) - \frac{d}{dx}\left(\frac{\partial}{\partial y'}(f - \lambda g)\right) = 0} \tag{11}$$

for $I - \lambda J$ for some λ called the Lagrange Multiplier.

Remark This proof can be adapted to more derivatives or more independent variables.

We now solve the examples given at the start of this subsection.

1. **Catenary** We have $I(y) := \rho g \int_{x_1}^{x_2} y\sqrt{1 + (y')^2} dx$ and $J(y) := \rho \int_{x_1}^{x_2} \sqrt{1 + (y')^2} dx =: M$. y satisfies the E-L equation for $I - \lambda J$ for some λ. This functional is

$$\rho \int_{x_1}^{x_2} (gy - \lambda)\sqrt{1 + (y')^2} dx$$

and we use the optical analogy to solve it. This corresponds to light moving with speed $c = \frac{1}{gy - \lambda}$ and has solution of

$$y = \lambda + \frac{c_1}{\rho g} \cosh \frac{\rho g x}{c_1} + c_2$$

and we have three conditions and three unknowns and so we can solve to find c_1, c_2, λ

2. **Isoperimetric Problem** We want to maximise $A(x, y)$ while keeping $L(x, y) = l$ fixed. $(x(t), y(t))$ is a parameterisation of a closed curve. The extremising curve will satisfy the E-L equations for $A - \lambda L$ and so we get

$$(A - \lambda L)(x, y) = \int_{t_1}^{t_2} \left(\frac{1}{2}(x\dot{y} - y\dot{x}) - \lambda\sqrt{\dot{x}^2 + \dot{y}^2}\right) dt$$

and we have two E-L equations and so we have

$$\begin{aligned} \tfrac{1}{2}\dot{y} - \tfrac{d}{dt}\left(-\tfrac{1}{2}y - \lambda\tfrac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}}\right) &= 0 \\ -\tfrac{1}{2}\dot{x} - \tfrac{d}{dt}\left(\tfrac{1}{2}x - \lambda\tfrac{\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}}\right) &= 0 \end{aligned}$$

Note that both equations are time derivatives and so we get

$$\frac{d}{dt}\left(y - \lambda\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}}\right) = 0$$
$$\frac{d}{dt}\left(-x + \lambda\frac{\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}}\right) = 0$$

and integrating once gives

$$y - \lambda\frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2}} = B$$
$$-x + \lambda\frac{\dot{y}}{\sqrt{\dot{x}^2+\dot{y}^2}} = -C$$

and hence we get that $(x-C)^2 + (y-B)^2 = \lambda^2$ and this is a circle centre (C, B) and radius λ. Therefore $2\pi\lambda = l$ and so $\lambda = \frac{l}{2\pi}$.

To handle example three we need a new method.

6.1.5 One functional constrained by a function

As far as I can gather, we cannot in general constrain a functional with respect to a given function, but we can do if the functional is a function of curves.

Suppose we have curves $\underline{x}(t) = (x_1(t), ..., x_n(t))$ joining two points $\underline{x}_{(1)}$ and $\underline{x}_{(2)}$ at times t_1 and t_2 and the curves satisfy $g(t, \underline{x}(t), \dot{\underline{x}}(t)) = 0$. We have some functional $I(\underline{x}) = \int_{t_1}^{t_2} f(t, \underline{x}(t), \dot{\underline{x}}(t))dt$ and we aim to extremise amongst these curves. An example of this is finding geodesics in a surface.

Let $\underline{x}_h$ be a variation of an extremum $\underline{x}$, and $\underline{x}_h = \underline{x} + h\underline{u} + o(h^2)$ which satisfy the constraint for all h. Then $h = 0$ is a critical point for $I(\underline{x}_h)$ as a function of h and so

$$\left.\frac{d}{dh}I(\underline{x}_h)\right|_{h=0} = 0$$

and thus we get that

$$\int_{t_1}^{t_2}\left[\sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}u_i + \frac{\partial f}{\partial \dot{x}_i}\dot{u}_i\right)\right]dt = 0 \tag{12}$$

Differentiating the constraint $g(t, \underline{x}(t), \dot{\underline{x}}(t)) = 0$ gives

$$\sum_{i=1}^{n}\left(\frac{\partial g}{\partial x_i}u_i + \frac{\partial g}{\partial \dot{x}_i}\dot{u}_i\right) = 0 \text{ for all } t \tag{13}$$

Pick a function $\lambda(t)$, multiply (13) by λ and subtract from the integrand of (12). Therefore we get

$$\begin{aligned} 0 &= \int_{t_1}^{t_2}\sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}u_i + \frac{\partial f}{\partial \dot{x}_i}\dot{u}_i - \lambda(t)\left(\frac{\partial g}{\partial x_i}u_i + \frac{\partial g}{\partial \dot{x}_i}\dot{u}_i\right)\right)dt \\ &= \int_{t_1}^{t_2}\left[\sum_{i=1}^{n}\left(\frac{\partial}{\partial x_i}(f-\lambda g)u_i + \frac{\partial}{\partial \dot{x}_i}(f-\lambda g)\dot{u}_i\right)\right]dt \end{aligned}$$

Observe that $\underline{u}(t_1) = 0 = \underline{u}(t_2)$ and then integrating the $\dot{u}_i$ terms by parts gives

$$0 = \int_{t_1}^{t_2}\left[\sum_{i=1}^{n}\left(\frac{\partial}{\partial x_i}(f-\lambda g) - \frac{d}{dt}\left(\frac{\partial}{\partial \dot{x}_i}(f-\lambda g)\right)\right)u_i\right]dt$$

We can pick $\lambda(t)$ so that one of the coefficients, say of u_i, is zero. We can do this since setting $\frac{\partial}{\partial x_1}(f-\lambda g) - \frac{d}{dt}\left(\frac{\partial}{\partial \dot{x}_1}(f-\lambda g)\right) = 0$ gives a first order linear inhomogeneous ODE for λ which can be solved by the integrating factor method. The constraint (12) amongst the u_i then determines u_1 in terms of $u_2, ..., u_n$ and then the latter can be varied freely. Hence the condition for an extremum becomes

$$0 = \sum_{i=2}^{n} \left(\int_{t_1}^{t_2} \left[\left(\frac{\partial}{\partial x_i}(f-\lambda g) - \frac{d}{dt}\left(\frac{\partial}{\partial \dot{x}_i}(f-\lambda g) \right) \right) u_i \right] dt \right)$$

and since now $u_2, ..., u_n$ are arbitrary, vanishing at t_1 and t_2 we can apply the fundamental theorem to each u_i in turn taking the rest of $u_2, ..., u_n$ to be zero giving

$$\frac{\partial}{\partial x_i}(f-\lambda g) - \frac{d}{dt}\left(\frac{\partial}{\partial \dot{x}_i}(f-\lambda g) \right) = 0 \text{ for } i = 2, ..., n$$

For $i = 1$ this equation was the way we chose $\lambda(t)$ and hence the equation becomes

$$\boxed{\frac{\partial}{\partial x_i}(f-\lambda g) - \frac{d}{dt}\left(\frac{\partial}{\partial \dot{x}_i}(f-\lambda g) \right) = 0 \text{ for } i = 1, 2, ..., n} \tag{14}$$

We have thus proved

Theorem 6.3 *To extremise a functional I given by a Lagrangian f amongst curves $\underline{x}(t)$ with fixed endpoints subject to a constraint $g(t, \underline{x}(t), \dot{\underline{x}}(t)) = 0$ there is a function $\lambda(t)$ such that the Euler-Lagrange equations for the Lagrangian $f - \lambda g$ (namely (14)) are satisfied.*

Example 6.3 *Geodesics on a surface in $\mathbb{R}^3$ given by an equation $g(\underline{x}) = 0$. Geodesics are paths of shortest length and so minimise$\int_{t_1}^{t_2} \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} dt$. Then there is a function $\lambda(t)$ such that a geodesic $\underline{x}(t)$ satisfies the E-L equations of the Lagrangian $\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} - \lambda(t)g(\underline{x})$. We thus have, for the x equation,*

$$-\lambda(t)\frac{\partial g}{\partial x} - \frac{d}{dt}\left(\frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} \right) = 0$$

We aim to simplify this equation, and so we introduce the arclength parameter s. This is given by $\frac{ds}{dt} = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}$ and we change independent variables from t to s. Then $\frac{dx}{ds} = \frac{\dot{x}}{\frac{ds}{dt}} = \frac{\dot{x}}{\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}}$. Then dividing the equation by $\frac{ds}{dt}$ and putting $\mu = \frac{\lambda}{\frac{ds}{dt}}$ we get that

$$-\mu\frac{\partial g}{\partial x} - \frac{d^2x}{ds^2} = 0$$

and the y equation becomes

$$-\mu\frac{\partial g}{\partial y} - \frac{d^2y}{ds^2} = 0$$

and similarly the z equation becomes

$$-\mu\frac{\partial g}{\partial z} - \frac{d^2z}{ds^2} = 0$$

and so

$$-\mu = \frac{\frac{d^2x}{ds^2}}{\frac{\partial g}{\partial x}} = \frac{\frac{d^2y}{ds^2}}{\frac{\partial g}{\partial y}} = \frac{\frac{d^2z}{ds^2}}{\frac{\partial g}{\partial z}}$$

There is no general method to solve these equations. We now consider a special case of a **sphere** *in* $\mathbb{R}^3$*, and so we have that* $g(x,y,z) = x^2+y^2+z^2-R^2$*. This then gives us that*

$$\frac{1}{2x}\frac{d^2x}{dx^2} = \frac{1}{2y}\frac{d^2y}{dx^2} = \frac{1}{2z}\frac{d^2z}{dx^2}$$

and notice that

$$\frac{d}{ds}\left(z\frac{dy}{ds} - y\frac{dz}{ds}\right) = \frac{dz}{ds}\frac{dy}{ds} + z\frac{d^2y}{ds^2} - \frac{dy}{ds}\frac{dz}{ds} - y\frac{d^2z}{ds^2} = yz\left(\frac{1}{y}\frac{d^2y}{ds^2} - \frac{1}{z}\frac{d^2z}{ds^2}\right) = 0$$

and so $z\frac{dy}{ds} - y\frac{dz}{ds} = A$ *is constant. Similarly* $x\frac{dz}{ds} - z\frac{dx}{ds} = B$ *is constant and* $y\frac{dx}{ds} - x\frac{dy}{ds} = C$ *is constant. Therefore multiplying by* x*,* y *and* z *respectively on these equations gives* $0 = Ax+By+Cz$*. This is a plane through the origin perpendicular to* (A,B,C)*. Hence the path must lie in the intersection of the sphere and a plane through the origin. These are called* **Great Circles***. We have two solutions to the E-L equations satisfying the endpoint conditions so long as the endpoints are not antipodal. If two endpoints are poles then we get a continuum of great circles all of which are solutions to the problem.*

7 Constrained Motion

For particles moving with coordinates related by a constraint, say $g = 0$, then Hamilton's principle extremises $\int_{t_1}^{t_2} L dt$ where $L = T - V$ and now we are subjected to a constraint. We use the Lagrange Multiplier method, and so we have a function $\lambda(t)$ such that the motion satisfies the E-L equation for $T - V - \lambda g$

Example 7.1 *Consider free motion on a surface in* $\mathbb{R}^3$*. We then have by definition* $V = 0$ *and also* $T = \frac{1}{2}m(\dot{x}^2+\dot{y}^2+\dot{z}^2)$ *and* $g(x,y,z) = 0$*. Thus we want the E-L equations for* $\frac{1}{2}m(\dot{x}^2+\dot{y}^2+\dot{z}^2) - \lambda(t)g(x,y,z)$ *and so we have that*

$$\left.\begin{array}{l} -\lambda\frac{\partial g}{\partial x} - \frac{d}{dt}(m\dot{x}) = 0 \\ -\lambda\frac{\partial g}{\partial y} - \frac{d}{dt}(m\dot{y}) = 0 \\ -\lambda\frac{\partial g}{\partial z} - \frac{d}{dt}(m\dot{z}) = 0 \end{array}\right\} \implies m\ddot{\underline{x}} = -\lambda\nabla g$$

∇g *is a vector perpendicular to the surface at each point. If we eliminate* $\frac{\lambda}{m}$ *then we get that*

$$\frac{\ddot{x}}{\frac{\partial g}{\partial x}} = \frac{\ddot{y}}{\frac{\partial g}{\partial y}} = \frac{\ddot{z}}{\frac{\partial g}{\partial z}}$$

Observe that

$$\frac{d}{dt}(\dot{x}^2+\dot{y}^2+\dot{z}^2) = 2\dot{x}\ddot{x} + 2\dot{y}\ddot{y} + 2\dot{z}\ddot{z} = -\frac{2\lambda}{m}\left(\dot{x}\frac{\partial g}{\partial x} + \dot{y}\frac{\partial g}{\partial y} + \dot{z}\frac{\partial g}{\partial z}\right) = 0$$

and so $\dot{x}^2+\dot{y}^2+\dot{z}^2$ *is constant and so* $\frac{ds}{dt}$ *is constant so changing from* t *to* s *gives the geodesic equation. Hence the motion of a free particle is along a geodesic and constant speed.*