§2.1 Calculus of variations

Let us start by reviewing how to find the maxima and minima of a function $f(s)\colon \mathbb{R} \to \mathbb{R}$. As you will recall, this can be done by solving the equation

$$\frac{df}{ds} = 0$$

as a function of s. As an example, if our function is $f(s) = \frac{1}{2}s^2 - s$ we have

$$\frac{df}{ds} = s - 1$$

so the function has an extremum (a minimum, in this case) at $s = 1$. An alternative way of formulating the same condition makes use of the definition of derivative as encoding the change in the function under small changes in s. For a small $\delta s \in \mathbb{R}$, we have

$$f(s + \delta s) = f(s) + \frac{df(s)}{ds}\delta s + R(s, \delta s)$$

where $R(s, \delta s)$ is an error term. It is convenient to introduce the notation

$$\delta f \coloneqq f(s + \delta s) - f(s)$$

so the statement above becomes

$$\delta f = \frac{df(s)}{ds}\delta s + R(s, \delta s)\,.$$

We note that the usual definition of the derivative implies

$$\lim_{\delta s \to 0} \frac{R(s, \delta s)}{\delta s} = 0\,.$$

In these cases we say that "δf vanishes to first order in δs". The functions that we will study will almost always admit a well-behaved Taylor expansion, so this result implies that $R(s, \delta s)$ is at least of quadratic order in δs. Henceforth we will encode this vanishing to first order by writing $\mathcal{O}((\delta s)^2)$ instead of $R(s, \delta s)$.

So, finally, we can say that the extrema of $f(s)$ are located at the points where

$$\delta f = \mathcal{O}((\delta s)^2)\,.$$

The same reasoning can be applied in the case of functions of multiple variables. Consider a function $f(s_1, \ldots, s_N)\colon \mathbb{R}^N \to \mathbb{R}$, and introduce a small displacement $s_i \to s_i + \delta s_i$. In this case the partial derivatives $\partial f/\partial s_i$ are defined by

$$\delta f = \sum_{i=1}^{N} \frac{\partial f}{\partial s_i}\delta s_i + \mathcal{O}(\delta s^2)$$

where the error term includes terms vanishing faster than δs_i (so terms of the form δs_1^2, $\delta s_1 \delta s_2, \ldots$). Stationary points[1] of f are located wherever δf vanishes to first order in δs_i.

In fact, we need to go one step further, and work with *functionals*: these are maps from *functions* to $\mathbb{R}$. One (heuristic, but sometimes useful) way of thinking of them is as the limit of the previous multi-variate case when the number of variables N goes to infinity. From instance, we could have a functional $S[y(t)]$ defined by

$$S[y(t)] = \int_a^b y(t)^2 \, dt$$

for some fixed choice of (a, b). I emphasize that one should think of S as the analogue of f above, and the different functions $y(t)$ as the "points" in the domain of this functional.

We want to define a meaning for a function $y(t)$ to give an extremal value for the functional S. In analogy with what happened in the finite dimensional case above, we can study the variation of S as we displace $y(t)$ slightly. We need to be a bit careful when specifying which class of functions $y(t)$ we are going to include in our extremization problem. In the case of interest to us, we will extremize over the set of smooth[2] functions $y(t)$ with fixed values at the endpoints a and b. That is, we fix $y(a) = y_a$ and $y(b) = y_b$, for some fixed values of y_a and y_b.

Definition 2.1.1. We say that a function $y(t)$ is *stationary* (for the functional S) if

$$\left. \frac{dS[y(t) + \epsilon z(t)]}{d\epsilon} \right|_{\epsilon=0} = 0$$

for *all* smooth $z(t)$ such that $z(a) = z(b) = 0$.

Remark 2.1.2. This definition encodes the idea that the path $y(t)$ extremises the action S in the space of paths starting at y_a and ending at y_b. To see this, think of the function $z(t)$ as an arbitrary choice of direction in the space of deformations. The condition is then saying that $S[y(t)]$ is extremised along this direction. Since $z(t)$ is arbitrary, we have that the condition is satisfied in every direction in function space. It might help to compare with how you could impose that an ordinary function $f(\mathbf{x})$ depending on a vector $\mathbf{x}$ has an extremum at $\mathbf{x}_0$. You could impose

$$\left. \frac{df(\mathbf{x}_0 + \epsilon \mathbf{z})}{d\epsilon} \right|_{\epsilon=0} = \left[\frac{d}{d\epsilon} \left(f(\mathbf{x}_0) + \epsilon \sum_i z_i \left. \frac{\partial f(\mathbf{x})}{\partial x_i} \right|_{\mathbf{x}=\mathbf{x}_0} + \mathcal{O}(\epsilon^2) \right) \right]_{\epsilon=0} = \sum_{i=1}^n z_i \left. \frac{\partial f(\mathbf{x})}{\partial x_i} \right|_{\mathbf{x}=\mathbf{x}_0} = 0$$

for every choice of vector $\mathbf{z}$. This is equivalent to imposing the familiar condition

$$\left. \frac{\partial f(\mathbf{x})}{\partial x_i} \right|_{\mathbf{x}=\mathbf{x}_0} = 0$$

for all i, which ensures the existence of an extremum (or saddle point) at $\mathbf{x} = \mathbf{x}_0$.

[1] You might want to remind yourself of section 1.9 of the Calculus I Epiphany notes.

[2] My conventions are that smooth functions are those which have continuous derivatives to all orders.

Note 2.1.3

Consider the Taylor expansion in ϵ (which is a constant) of $S[y(t) + \epsilon z(t)]$:

$$S[y(t) + \epsilon z(t)] = S[y(t)] + \epsilon \left.\frac{dS[y(t) + \epsilon z(t)]}{d\epsilon}\right|_{\epsilon=0} + \frac{1}{2}\epsilon^2 \left.\frac{d^2 S[y(t) + \epsilon z(t)]}{d\epsilon^2}\right|_{\epsilon=0} + \dots$$

The condition for $y(t)$ to be stationary is that the term proportional to ϵ vanishes:

$$\delta S \coloneqq S[y(t) + \epsilon z(t)] - S[y(t)] = \mathcal{O}(\epsilon^2)\,.$$

It is useful to think of the combination $\epsilon z(t)$ as a small variation of $y(t)$, which we denote $\delta y(t) \coloneqq \epsilon z(t)$. We define $\mathcal{O}((\delta y(t))^n)$ to mean simply $\mathcal{O}(\epsilon^n)$. In particular, we can rewrite the stationary condition as

$$\delta S = \mathcal{O}((\delta y(t))^2)\,.$$

If you are ever confused about the expansions in $\delta y(t)$ below, you can replace $\delta y(t)$ with $\epsilon z(t)$, and expand in the constant ϵ. For instance, consider the integral

$$\int g(t)(\delta y(t))^n dt$$

For any positive integer n and any function $g(t)$. I claim that this is $\mathcal{O}((\delta y(t))^n)$. The proof is as follows: replacing $\delta y(t)$ with $\epsilon z(t)$ we have

$$\int g(t)\epsilon^n z(t)^n dt = \epsilon^n \int g(t) z(t)^n dt\,.$$

(Here we have used that ϵ is a constant.) In our $\mathcal{O}((\delta y(t)^n))$ notation, this means that

$$\int \mathcal{O}((\delta y(t))^n) dt = \mathcal{O}((\delta y(t))^n)\,.$$

That is, integration does not change the order in $\delta y(t)$ (which, once more, when talking about action functionals I define to be really just the order in ϵ).

Note 2.1.4

Because our interest in dynamical problems, we will often refer to the functions $y(t)$ as *paths*, so that the conditions above define what an *"stationary path"* is.

We are now in a position to introduce the action principle. Assume that we have an *action functional* (or simply *action*) $S\colon \{\text{functions}\} \to \mathbb{R}$, which takes functions, and

generates a real number. In the Lagrangian formalism all of the physical content of the theory can be summarized in the choice of action functional.

For now we also assume that we have a particle moving in one dimension, and we want to determine its motion. Its trajectory is given by a function $x(t)$, with t the time coordinate. For many physical problems equations of motion are second order in $x(t)$, so we need data to fix two integration constants. In the Lagrangian formalism these are given by fixing the initial and final positions. That is, we will assume that we know the initial position $x(t_0)$ of the particle at time t_0, and its final position $x(t_1)$ at time t_1.

The *action principle*[3] then states that for **arbitrary** smooth small deformations $\delta x(t)$ around the "true" path $x(t)$ (that is, the path that the particle will actually follow) we have

$$\delta S \coloneqq S[x + \delta x] - S[x] = \mathcal{O}((\delta x)^2)\,. \tag{2.1.1}$$

Or in other words:

Action principle:
It is possible to choose an action functional $S[x(t)]$ such that the paths described by physical particles are stationary paths of S.

In a moment we will need an important result known as the *fundamental lemma of the calculus of variations*. It goes as follows:

Lemma 2.1.5 (Fundamental lemma of the calculus of variations)**.** *Consider a function $f(x)$ continuous in the interval $[a, b]$ (we assume $a < b$) such that*

$$\int_a^b f(x)g(x)\,dx = 0$$

for all smooth functions $g(x)$ in $[a, b]$ such that $g(a) = g(b) = 0$. Then $f(x) = 0$ for all $x \in (a, b)$.

Proof. I will prove the result by contradiction. Assume that there is a point $p \in (a, b)$ such that $f(p) > 0$ (the case $f(p) < 0$ can be proven analogously). By continuity of $f(x)$, there will be some non-vanishing interval $[p_0, p_1]$ where $f(x) > 0$. Construct

$$g(x) = \begin{cases} \nu(x - p_0)\nu(p_1 - x) & \text{if } x \in (p_0, p_1) \\ 0 & \text{otherwise.} \end{cases}$$

with $\nu(x) = \exp(-1/x)$. It is an interesting exercise in first year calculus to prove that this function is smooth everywhere, including at p_0 and p_1 (it is an example of "bump

[3]This goes under various names in the literature. Common ones are *action principle*, *least action principle*, *extremal action principle* and (less precisely) *variational principle*. I will mostly use "action principle", which has the advantage of being concise.

functions", useful in many domains). Clearly $f(x)g(x) > 0$ for $x \in (p_0, p_1)$, and vanishes otherwise. This implies that

$$\int_a^b f(x)g(x)\,dx = \int_{p_0}^{p_1} f(x)g(x)\,dx > 0$$

which is a contradiction. □