Chapter 3

Calculus of variations

Let $[a, b] \subset \mathbb{R}$ and let $F \in C^1([a, b] \times \mathbb{R} \times \mathbb{R})$. Throughout this chapter I will be denoting the arguments of F by (x, y, z): $(x, y, z) \mapsto F(x, y, z)$.

For given $\alpha, \beta \in \mathbb{R}$ put $\mathcal{C} := \{ u \in C^1[a, b] \big| \ u(a) = \alpha,\ u(b) = \beta \}$.

Define the map (integral functional) $J : \mathcal{C} \to \mathbb{R}$ by $J(u) := \int_a^b F(x, u(x), u'(x))\, dx$.

Main problem of the calculus of variations: find a $u \in \mathcal{C}$ such that

$$J(u) = \min_{v \in \mathcal{C}} J(v) \tag{3.1}$$

or

$$J(u) = \max_{v \in \mathcal{C}} J(v). \tag{3.2}$$

Of course, problem (3.2) is reduced to (3.1) by changing the sign of F, so further on we will deal with (3.1).

Why do we have only one variable x rather than m? Just to make the exposition simpler. Most sensible problems of mathematical physics admit a variational formulation, with m usually taking values 1, 2 or 3. In particular, all problems considered in the first two chapters of the course admit a variable formulation.

Two examples providing motivation for the above abstract setting are given in Handout 4 prepared by the previous lecturer, Hartmut Logemann. These are examples of a geodesic on a surface and the brachistochrone problem.

For me personally the main motivation is the Dirichlet problem for the Laplace operator considered in Chapter 1. The fact that this problem can be set in variational terms was addressed in Question 3 of Exercise Sheet 3. This matter will be addressed again in Question 1 of Exercise Sheet 7.

Put $\mathcal{C}_0 := \{ u \in C^1[a, b] \big| \ u(a) = u(b) = 0 \}$. We will need the following

Lemma 3.1 *(du Bois-Reymond) Let $f \in C[a, b]$. If*

$$\int_a^b f(x)\, g'(x)\, dx \ = \ 0\,, \qquad \forall g \in \mathcal{C}_0\,,$$

then $f \equiv \text{const}$.

Proof Set

$$c := \frac{1}{b - a} \int_a^b f(x)\, dx$$

and

$$h(x) := \int_a^x (f(t) - c)\, dt\,.$$

Then $h \in C^1[a,b]$ and $h(a) = h(b) = 0$, so $h \in \mathcal{C}_0$. Also, by the hypothesis of the lemma

$$\int_a^b f(x)\, h'(x)\, dx \;=\; 0\,.$$

Hence,

$$\int_a^b (f(x) - c)\, h'(x)\, dx \;=\; \int_a^b f(x)\, h'(x)\, dx \;-\; c(h(b) - h(a)) \;=\; 0\,.$$

But $h'(x) = f(x) - c$, so

$$\int_a^b (f(x) - c)^2\, dx \;=\; 0\,.$$

As the function $f(x) - c$ is continuous on $[a,b]$ this implies

$$f(x) - c = 0, \qquad \forall x \in [a,b].$$

□

We are now in a position to derive a necessary condition for $u \in \mathcal{C}$ to satisfy (3.1).

Theorem 3.1 *Under the above assumptions on F, if $u \in \mathcal{C}$ satisfies (3.1) then*

$$\frac{d}{dx}\left(\frac{\partial F}{\partial z}(x, u(x), u'(x))\right) - \frac{\partial F}{\partial y}(x, u(x), u'(x)) = 0\,, \tag{3.3}$$

$\forall x \in [a,b]$.

Equation (3.3) is called the *Euler–Lagrange equation.*

Note that it is not a priori obvious that the function $\frac{\partial F}{\partial z}(x, u(x), u'(x))$ is differentiable on $[a,b]$, so the fact that it turns out to be differentiable is nontrivial.

Proof of Theorem 3.1 Suppose that $u \in \mathcal{C}$ to satisfies (3.1). Take arbitrary $w \in \mathcal{C}_0$. Then for any $\tau \in \mathbb{R}$ we have

$$u + \tau w \in \mathcal{C}\,.$$

The function τw is often called the *variation* of u and is denoted δu.

Define $g : \mathbb{R} \to \mathbb{R}$ by $g(\tau) := J(u + \tau w)$. Clearly, g has a minimum at $\tau = 0$. Under our assumptions on F the function g is differentiable, so

$$\frac{dg}{d\tau}(0) = 0.$$

Differentiating under the integral sign (which is justified under our assumptions on F) we get

$$\left[\int_a^b \frac{d}{d\tau} F(x, u + \tau w, u' + \tau w')\, dx\right]_{\tau=0} = 0.$$

(Here and in some subsequent formulae I suppress, for the sake of brevity, the dependence of u, u' and w, w' on x.) Hence,

$$\begin{aligned} 0 = &\left[\int_a^b \left(\frac{\partial F}{\partial y}(x, u + \tau w, u' + \tau w')\; w \;+\; \frac{\partial F}{\partial z}(x, u + \tau w, u' + \tau w')\; w'\right) dx\right]_{\tau=0} \\ &= \int_a^b \left(\frac{\partial F}{\partial y}(x, u, u')\; w \;+\; \frac{\partial F}{\partial z}(x, u, u')\; w'\right) dx\,. \end{aligned} \tag{3.4}$$

Put

$$A(x) := \int_a^x \frac{\partial F}{\partial y}(t, u(t), u'(t))\, dt\,, \qquad B(x) := \frac{\partial F}{\partial z}(x, u(x), u'(x))\,. \tag{3.5}$$

Then

$$\int_a^b \frac{\partial F}{\partial y}\, w\,dx = \int_a^b A'\,w\,dx = A\,w|_a^b - \int_a^b A\,w'\,dx = -\int_a^b A\,w'\,dx$$

since $w \in \mathcal{C}_0$. (Here and in some subsequent formulae I suppress, for the sake of brevity, the dependence of $\frac{\partial F}{\partial y}$ and $\frac{\partial F}{\partial z}$ on (x, u, u').)

Formula (3.4) can now be rewritten as

$$\int_a^b (Bw' - Aw')\,dx = 0\,.$$

This is true for any $w \in \mathcal{C}_0$, so by Lemma 3.1

$$B - A = \text{const}$$

which implies

$$\frac{d}{dx}(B - A) = 0. \tag{3.6}$$

Substituting (3.5) into (3.6) we arrive at (3.3) □

Remark 3.1 *If we work in the space of twice continuously differentiable functions then there is no need to use the du Bois-Reymond Lemma in the derivation of the Euler–Lagrange equation. This matter is addressed in Question 5 of Exercise Sheet 7. The reason we had to use the du Bois-Reymond Lemma in the proof of Theorem 3.1 is that we wanted to derive the Euler–Lagrange equation under minimal smoothness assumptions.*

Definition 3.1 *A function satisfying the Euler–Lagrange equation (3.3) is called an* extremal.

If $u \in \mathcal{C}$ is a solution of the variational problem (3.1) then u is an extremal, but not every extremal is a solution of the variational problem (3.1). Extremals are merely 'candidates' for solutions of the variational problem (3.1). In other words, Theorem 3.1 provides a necessary condition for solutions of the variational problem (3.1).

We will now examine the Euler–Lagrange equation (3.3) for several special cases and go through a few examples.

Special case 1: F does not depend on z. Abusing notation, I will write in this case $F(x, y, z) = F(x, y)$. Then the Euler–Lagrange equation (3.3) becomes

$$\frac{\partial F}{\partial y}(x, u(x)) = 0\,.$$

This is no longer a differential equation for u, but an implicit definition of the function $x \mapsto u(x)$.

Special case 2: F does not depend on y. Abusing notation, I will write in this case $F(x, y, z) = F(x, z)$. Then the Euler–Lagrange equation (3.3) becomes

$$\frac{d}{dx}\left(\frac{\partial F}{\partial z}(x, u'(x))\right) = 0$$

which is equivalent to the statement

$$\frac{\partial F}{\partial z}(x, u'(x)) = c_1$$

where c_1 is some constant. This gives us an implicit definition of $u'(x)$ in terms of x and c_1, $u'(x) = f(x, c_1)$. Integrating, we get

$$u(x) = \int_a^x f(t, c_1)\,dt \,+\, c_2\,.$$

Of course, the constants c_1 and c_2 have to be chosen in such a way that u satisfies the boundary conditions

$$u(a) = \alpha, \qquad u(b) = \beta.$$

The third special case is the most important in examples and applications.

Special case 3: F does not depend on x. Abusing notation, I will write in this case $F(x, y, z) = F(y, z)$. It turns out that in this case one can effectively lower the order of the Euler–Lagrange equation (3.3).

Lemma 3.2 *If $F(x, y, z) = F(y, z)$ and $u \in C^2[a, b]$ is an extremal, then*

$$E(x) := F(u(x), u'(x)) - u'(x)\frac{\partial F}{\partial z}(u(x), u'(x)) = c_1 \tag{3.7}$$

where c_1 is some constant.

Proof of Lemma 3.2 We have

$$E' = \frac{dE}{dx} = \frac{\partial F}{\partial y}u' + \frac{\partial F}{\partial z}u'' - u''\frac{\partial F}{\partial z} - u'\frac{d}{dx}\frac{\partial F}{\partial z} = -u'\left(\frac{d}{dx}\frac{\partial F}{\partial z} - \frac{\partial F}{\partial y}\right) \overset{\text{by (3.3)}}{=} 0.$$

□

An inspection of the above proof shows that if $u \in C^2[a, b]$ and $u'(x) \neq 0$ for all $x \in [a, b]$ and $E(x) = c_1$, then u satisfies the Euler–Lagrange equation (3.3). Thus, in a sense, Lemma 3.2 works both ways, and essentially it allows us to lower the order of the Euler–Lagrange equation (3.3).

One usually looks for u satisfying $E(x) = c_1$ in order to find solutions to the Euler–Lagrange equation. The equation $E(x) = c_1$ gives us an implicit definition of $u'(x)$ in terms of $u(x)$ and c_1, $u'(x) = f(u(x), c_1)$. Separating variables, we get

$$dx = \frac{du}{f(u, c_1)} \qquad \Longrightarrow \qquad x = g(u, c_1) + c_2.$$

Solving the latter equation for u, we get u as a function of x and the integration constants c_1 and c_2. Of course, the constants c_1 and c_2 have to be chosen in such a way that u satisfies the boundary conditions

$$u(a) = \alpha, \qquad u(b) = \beta.$$

There is an important class of problems within the Special Case 3, namely, problems with

$$F(y, z) = h(y)\sqrt{1 + z^2}\,. \tag{3.8}$$

For such problems identity (3.7) becomes

$$h(u)\sqrt{1 + (u')^2} - \frac{h(u)\,(u')^2}{\sqrt{1 + (u')^2}} = c_1,$$

$$\Updownarrow$$

$$\frac{h(u)}{\sqrt{1 + (u')^2}} = c_1.$$

Resolving with respect to u', we get

$$u' = \frac{du}{dx} = \sqrt{\left(\frac{h(u)}{c_1}\right)^2 - 1}\,.$$

We solve this equation by separating variables:

$$dx = \frac{du}{\sqrt{\left(\frac{h(u)}{c_1}\right)^2 - 1}}\,,$$

$$x = \int \frac{du}{\sqrt{\left(\frac{h(u)}{c_1}\right)^2 - 1}} + c_2\,. \tag{3.9}$$

Evaluating the integral and solving (3.9) for u, we get u as a function of x and the integration constants c_1 and c_2. Of course, the constants c_1 and c_2 have to be chosen in such a way that u satisfies the boundary conditions

$$u(a) = \alpha, \qquad u(b) = \beta.$$

I must admit that I carried out the above arguments sloppily because a) I divided by c_1 without checking whether c_1 can be zero, and b) I really should have written $\pm$ when resolving my equation with respect to u'. I will have to allow this sloppiness because careful analysis of problems of this type takes too much time.

We now consider two important examples of variational problems with integrand of the form (3.8).

Example 3.1 *(Geodesics on a cylinder) Here*

$$F(x, y, z) = F(z) = \sqrt{1 + z^2}\,.$$

In particular, F does not depend on y (Special case 2), so

$$\frac{\partial F}{\partial z}(u'(x)) = F'(u'(x)) = c_1\,.$$

Now,

$$F'(z) = \frac{z}{\sqrt{1 + z^2}}\,,$$

therefore

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

Note that F' is a strictly increasing function mapping $\mathbb{R}$ to $(-1, 1)$ so the constant c_1 has to be from $(-1, 1)$. Resolving with respect to $u'(x)$, we get

$$u'(x) = \frac{c_1}{\sqrt{1 - c_1^2}}\,.$$

Integrating, we get

$$u(x) = \frac{c_1}{\sqrt{1 - c_1^2}}\, x + c_2\,.$$

This is a circular helix *on the cylinder.*

Example 3.2 *(Brachistochrone) Here*

$$F(x, y, z) = F(y, z) = \sqrt{\frac{1 + z^2}{y}}\,,$$

so we are in the situation (3.8) with

$$h(y) = \frac{1}{\sqrt{y}}\,.$$

By (3.9),

$$x = \int \frac{du}{\sqrt{\frac{1}{c_1^2 u} - 1}} + c_2 \,.$$

Let us make the substitutions

$$u = \frac{\tau}{c_1^2}, \qquad \tau = \sin^2\frac{\theta}{2} \,.$$

Then

$$x - c_2 = \frac{1}{c_1^2}\int \sqrt{\frac{\tau}{1-\tau}}\, d\tau \;=\; \frac{1}{2c_1^2}\int (1-\cos\theta)\, d\theta \,,$$

therefore,

$$x - c_2 \;=\; \frac{1}{2c_1^2}(\theta - \sin\theta) \,.$$

This is a parametric representation of a cycloid. *(Recall the standard geometric definition of a cycloid: it is the locus of a point on the rim of a circle of radius a rolling along a straight line. The standard parametric representation of a cycloid is* $x = a(t - \sin t)$, $y = a(1 - \cos t)$.*)*

The final subject which we will consider in this chapter is that of *variational problems with integral constraints.* Such problems are also known as *isoperimetric problems.*

As usual, we denote $\mathcal{C} := \{u \in C^1[a,b] \,\big|\; u(a) = \alpha,\; u(b) = \beta\}$.

We now have a pair of functions $F, F^* \in C^1([a,b] \times \mathbb{R} \times \mathbb{R})$. Their arguments are denoted in the standard way: $F(x,y,z)$, $F^*(x,y,z)$. We define two integral functionals $J, J^* : \mathcal{C} \to \mathbb{R}$ by

$$J(u) := \int_a^b F(x, u(x), u'(x))\, dx \,, \qquad J^*(u) := \int_a^b F^*(x, u(x), u'(x))\, dx \,.$$

Let L be a given constant. Denote $\mathcal{C}^* := \{u \in \mathcal{C} \,|\; J^*(u) = L\}$.

The problem is to find a $u \in \mathcal{C}^*$ such that

$$J(u) = \min_{v \in \mathcal{C}^*} J(v) \tag{3.10}$$

or

$$J(u) = \max_{v \in \mathcal{C}^*} J(v). \tag{3.11}$$

I give without proof the following

Theorem 3.2 *Suppose that* $u \in \mathcal{C}^*$ *is a solution of (3.10) and that it is not an extremal of the variational problem for* J^**. Then there exists a* $\lambda \in \mathbb{R}$ *such that* u *satisfies the Euler–Lagrange equation*

$$\frac{d}{dx}\left(\frac{\partial F_\lambda}{\partial z}(x, u(x), u'(x))\right) - \frac{\partial F_\lambda}{\partial y}(x, u(x), u'(x)) = 0 \,, \tag{3.12}$$

$\forall x \in [a,b]$*, where* $F_\lambda := F + \lambda F^*$.

Theorem 3.2 effectively reduces our *constrained* variational problem for the functional J to an *unconstrained* variational problem for the functional $J_\lambda := J + \lambda J^*$. Of course, now the Euler–Lagrange equation (3.12) has to be solved together with

$$J^*(u) = L. \tag{3.13}$$

The factor λ appearing in our construction is known as a *Lagrange multiplier.*

Theorem 3.2 is a development of a well known result from the analysis of constrained extrema of functions of several variables, see Section 3 in Handout 2 *Background Concepts from Analysis in Several Variables.*

Example 3.3 *(The catenary (or hanging cable) problem) Here*

$$F(x,y,z) = F(y,z) = g\rho y\sqrt{1+z^2}\,,$$

$$F^*(x,y,z) = F^*(z) = \sqrt{1+z^2}\,,$$

In order for the problem to make sense we assume (see Handout 5) that

$$L > \sqrt{(b-a)^2 + (\beta-\alpha)^2}\,. \tag{3.14}$$

Suppose that $u \in \mathcal{C}^*$ is a solution of (3.10). In order to apply Theorem 3.2 we need to check that u is not an extremal of the variational problem for J^*.

Seeking a contradiction, suppose that u is an extremal of the variational problem for J^*. Then, according to Example 3.1,

$$u(x) = c_1 x + c_2.$$

Since $u \in \mathcal{C}^*$, we have

$$c_1 a + c_2 = \alpha, \tag{3.15}$$

$$c_1 b + c_2 = \beta, \tag{3.16}$$

$$(b-a)\sqrt{1+c_1^2} = L. \tag{3.17}$$

Formulae (3.15), (3.16) imply

$$c_1 = \frac{\beta-\alpha}{b-a}\,.$$

Substituting this into (3.17) we get

$$\sqrt{(b-a)^2 + (\beta-\alpha)^2} = L$$

which contradicts (3.14).

Thus the application of Theorem 3.2 is justified.

We have

$$F_\lambda(x,y,z) = F_\lambda(y,z) = (g\rho y + \lambda)\sqrt{1+z^2}\,,$$

so we are in the situation (3.8) with

$$h(y) = g\rho y + \lambda.$$

By (3.9),

$$x = \int \frac{du}{\sqrt{\left(\frac{g\rho u+\lambda}{c_1}\right)^2 - 1}} + c_2\,.$$

Let us make the substitution

$$u = \frac{c_1 w - \lambda}{g\rho}\,.$$

Then

$$x - c_2 = \frac{c_1}{g\rho}\int \frac{dw}{\sqrt{w^2-1}} = \frac{c_1}{g\rho}\cosh^{-1} w\,,$$

therefore,

$$w(x) = \cosh\frac{g\rho}{c_1}(x-c_2)\,,$$

$$u(x) = \frac{c_1}{g\rho}\cosh\frac{g\rho}{c_1}(x-c_2) - \frac{\lambda}{g\rho}\,.$$

It is convenient to introduce new constants

$$C_1 = \frac{c_1}{g\rho}\,, \qquad \Lambda = -\frac{\lambda}{g\rho}\,.$$

Then the formula for $u(x)$ *becomes*

$$u(x) = C_1 \cosh \frac{x - c_2}{C_1} + \Lambda\,.$$

Of course, the constants C_1, c_2 *and* Λ *have to be chosen in such a way that* u *satisfies the conditions*

$$u(a) = \alpha, \qquad u(b) = \beta, \qquad J^*(u) = L.$$