Chapter 0

Introduction

0.1 Brief historical comments

The calculus of variations is one of the classical branches of mathematics. It was Euler who, looking at the work of Lagrange, gave the present name, not really self explanatory, to this field of mathematics.

In fact the subject is much older. It starts with one of the oldest problems in mathematics: the isoperimetric inequality. A variant of this inequality is known as the Dido problem (Dido was a semi historical Phoenician princess and later a Carthaginian queen). Several more or less rigorous proofs were known since the times of Zenodorus around 200 BC, who proved the inequality for polygons. There are also significant contributions by Archimedes and Pappus. Important attempts for proving the inequality are due to Euler, Galileo, Legendre, L'Huilier, Riccati, Simpson or Steiner. The first proof that agrees with modern standards is due to Weierstrass and it has been extended or proved with different tools by Blaschke, Bonnesen, Carathéodory, Edler, Frobenius, Hurwitz, Lebesgue, Liebmann, Minkowski, H.A. Schwarz, Sturm, and Tonelli among others. We refer to Porter [90] for an interesting article on the history of the inequality.

Other important problems of the calculus of variations were considered in the seventeenth century in Europe, such as the work of Fermat on geometrical optics (1662), the problem of Newton (1685) for the study of bodies moving in fluids (see also Huygens in 1691 on the same problem) or the problem of the brachistochrone formulated by Galileo in 1638. This last problem had a very strong influence on the development of the calculus of variations. It was resolved by Johann Bernoulli in 1696 and almost immediately after also by Jakob, his brother, Leibniz and Newton. A decisive step was achieved with the work of

Euler and Lagrange who found a systematic way of dealing with problems in this field by introducing what is now known as the Euler-Lagrange equation. This work was then extended in many ways by Bliss, Bolza, Carathéodory, Clebsch, Hahn, Hamilton, Hilbert, Kneser, Jacobi, Legendre, Mayer, Weierstrass, just to quote a few. For an interesting historical book on the one dimensional problems of the calculus of variations, see Goldstine [55].

In the nineteenth century and in parallel to some of the work that were mentioned above, probably, the most celebrated problem of the calculus of variations emerged, namely the study of the Dirichlet integral; a problem of multiple integrals. The importance of this problem was motivated by its relationship with the Laplace equation. Many important contributions were made by Dirichlet, Gauss, Thompson and Riemann among others. It was Hilbert who, at the turn of the twentieth century, solved the problem and was immediately after imitated by Lebesgue and then Tonelli. Their methods for solving the problem were, essentially, what are now known as the direct methods of the calculus of variations. We should also emphasize that the problem has been very important in the development of analysis in general and more notably functional analysis, measure theory, distribution theory, Sobolev spaces or partial differential equations. This influence is studied in the book by Monna [77].

The problem of minimal surfaces has also had, almost at the same time as the previous one, a strong influence on the calculus of variations. The problem was formulated by Lagrange in 1762. Many attempts to solve the problem were made by Ampère, Beltrami, Bernstein, Bonnet, Catalan, Darboux, Enneper, Haar, Korn, Legendre, Lie, Meusnier, Monge, Müntz, Riemann, H.A. Schwarz, Serret, Weierstrass, Weingarten and others. Douglas and Rado in 1930 gave, simultaneously and independently, the first complete proof. One of the first two Fields medals was awarded to Douglas in 1936 for having solved the problem. Immediately after the results of Douglas and Rado, many generalizations and improvements were made by Courant, Leray, MacShane, Morrey, Morse, Tonelli and many others since then. We refer for historical notes to Dierkes-Hildebrandt-Küster-Wohlrab [39] and Nitsche [82].

In 1900 at the International Congress of Mathematicians in Paris, Hilbert formulated 23 problems that he considered to be important for the development of mathematics in the twentieth century. Three of them (the 19th, 20th and 23rd) were devoted to the calculus of variations. These “predictions” of Hilbert have been amply justified all along the twentieth century and the field is at the turn of the twenty first one as active as in the previous century.

Finally we should mention that we will not speak of many important topics of the calculus of variations such as Morse or Liusternik-Schnirelman theories. The interested reader is referred to Ekeland [41], Mawhin-Willem [76], Struwe [97] or Zeidler [104].

0.2 Model problem and some examples

We now describe in more details the problems that we consider. The model case takes the following form

$$(P) \quad \inf \left\{ I(u) = \int_{\Omega} f(x, u(x), \nabla u(x))\, dx : u \in X \right\} = m.$$

This means that we want to minimize the integral, $I(u)$, among all functions $u \in X$ (and we call m the minimal value that can take such an integral), where

- $\Omega \subset \mathbb{R}^n$, $n \geq 1$, is a bounded open set, a point in Ω is denoted by $x = (x_1, \cdots, x_n)$;

- $u : \Omega \to \mathbb{R}^N$, $N \geq 1$, $u = \left(u^1, \cdots, u^N\right)$, and hence

$$\nabla u = \left(\frac{\partial u^j}{\partial x_i} \right)_{1 \leq i \leq n}^{1 \leq j \leq N} \in \mathbb{R}^{N \times n};$$

- $f : \overline{\Omega} \times \mathbb{R}^N \times \mathbb{R}^{N \times n} \longrightarrow \mathbb{R}$, $f = f(x, u, \xi)$, is continuous;

- X is the space of admissible functions (for example, $u \in C^1\left(\overline{\Omega}\right)$ with $u = u_0$ on $\partial\Omega$).

We are concerned with finding a *minimizer* $\overline{u} \in X$ of (P), meaning that

$$I(\overline{u}) \leq I(u), \ \forall u \in X.$$

Many problems coming from analysis, geometry or applied mathematics (in physics, economics or biology) can be formulated as above. Many other problems, even though not entering in this framework, can be solved by the very same techniques.

We now give several classical examples.

Example: Fermat principle. We want to find the trajectory that should follow a light ray in a medium with non-constant refraction index. We can formulate the problem in the above formalism. We have $n = N = 1$,

$$f(x, u, \xi) = g(x, u)\sqrt{1 + \xi^2}$$

and

$$(P) \quad \inf \left\{ I(u) = \int_a^b f(x, u(x), u'(x))\, dx : u(a) = \alpha,\ u(b) = \beta \right\} = m.$$

Example: Newton problem. We seek for a surface of revolution moving in a fluid with least resistance. The problem can be mathematically formulated as follows. Let $n = N = 1$,

$$f(x,u,\xi) = f(u,\xi) = 2\pi u \frac{\xi^3}{1+\xi^2}$$

and

$$(P) \quad \inf \left\{ I(u) = \int_a^b f(u(x), u'(x))\, dx : u(a) = \alpha,\ u(b) = \beta \right\} = m.$$

We will not treat this problem in the present book and we refer to Buttazzo-Kawohl [18] for a review article on this subject.

Example: brachistochrone. The aim is to find the shortest path between two points that follows a point mass moving under the influence of gravity. We place the initial point at the origin and the end one at $(b, -\beta)$, with $b, \beta > 0$. We let the gravity act downwards along the y-axis and we represent any point along the path by $(x, -u(x))$, $0 \le x \le b$.

In terms of our notations, we have that $n = N = 1$ and the function, under consideration, is

$$f(x,u,\xi) = f(u,\xi) = \frac{\sqrt{1+\xi^2}}{\sqrt{2gu}}$$

and

$$(P) \quad \inf \left\{ I(u) = \int_0^b f(u(x), u'(x))\, dx \ : u \in X \right\} = m$$

where

$$X = \left\{ u \in C^1([0,b]) : u(0) = 0,\ u(b) = \beta \text{ and } u(x) > 0,\ \forall x \in (0,b] \right\}.$$

The shortest path turns out to be a *cycloid*.

Example: minimal surface of revolution. We have to determine among all surfaces of revolution of the form

$$v(x,y) = (x, u(x)\cos y, u(x)\sin y)$$

with fixed end points $u(a) = \alpha,\ u(b) = \beta$ one with minimal area. We still have $n = N = 1$,

$$f(x,u,\xi) = f(u,\xi) = 2\pi u \sqrt{1+\xi^2}$$

and

$$(P) \quad \inf \left\{ I(u) = \int_a^b f(u(x), u'(x))\, dx : u(a) = \alpha,\ u(b) = \beta,\ u > 0 \right\} = m.$$

Solutions of this problem, when they exist, are *catenoids*. More precisely the minimizer is given, $\lambda > 0$ and μ denoting some constants, by

$$u(x) = \lambda \cosh \frac{x+\mu}{\lambda}.$$

Example: mechanical system. Consider a mechanical system with M particles whose respective masses are m_i and positions at time t are

$$u^i(t) = \left(x^i(t), y^i(t), z^i(t)\right) \in \mathbb{R}^3,\ 1 \le i \le M.$$

Let

$$T(u') = \frac{1}{2}\sum_{i=1}^{M} m_i \left|\left(u^i\right)'\right|^2 = \frac{1}{2}\sum_{i=1}^{M} m_i \left(\left(\left(x^i\right)'\right)^2 + \left(\left(y^i\right)'\right)^2 + \left(\left(z^i\right)'\right)^2\right)$$

be the kinetic energy and denote the potential energy with $U = U(t,u)$. Finally let

$$f(t,u,\xi) = T(\xi) - U(t,u)$$

be the Lagrangian. In our formalism we have $n = 1$ and $N = 3M$.

Example: Dirichlet integral. This is the most celebrated problem of the calculus of variations. Here we have $n > 1$, $N = 1$ and

$$(P) \quad \inf\left\{ I(u) = \frac{1}{2}\int_\Omega |\nabla u(x)|^2\, dx : u = u_0 \text{ on } \partial\Omega \right\}.$$

As for every variational problem we associate a differential equation which is nothing other than *Laplace equation*, namely $\Delta u = 0$.

Example: minimal surfaces. This problem is almost as famous as the preceding one. The question is to find among all surfaces $\Sigma \subset \mathbb{R}^3$ (or more generally in $\mathbb{R}^{n+1}$, $n \ge 2$) with prescribed boundary, $\partial\Sigma = \Gamma$, where Γ is a simple closed curve, one that is of minimal area. A variant of this problem is known as *Plateau problem*. One can realize experimentally such surfaces by dipping a wire into a soapy water; the surface obtained when pulling the wire out from the water is then a minimal surface.

The precise formulation of the problem depends on the kind of surfaces that we are considering. We have seen above how to write the problem for minimal surfaces of revolution. We now formulate the problem for more general surfaces.

Case 1: nonparametric surfaces. We consider (hyper) surfaces of the form

$$\Sigma = \left\{ v(x) = (x, u(x)) \in \mathbb{R}^{n+1} : x \in \overline{\Omega} \right\}$$

with $u : \overline{\Omega} \to \mathbb{R}$ and where $\Omega \subset \mathbb{R}^n$ is a bounded domain. These surfaces are therefore graphs of functions. The fact that $\partial\Sigma$ is a preassigned curve Γ, reads now as $u = u_0$ on $\partial\Omega$, where u_0 is a given function. The area of such a surface is given by

$$\text{Area}\,(\Sigma) = I\,(u) = \int_\Omega f\,(\nabla u\,(x))\,dx$$

where, for $\xi \in \mathbb{R}^n$, we have set

$$f\,(\xi) = \sqrt{1 + |\xi|^2}\,.$$

The problem is then written in the usual form

$$(P) \quad \inf\left\{ I\,(u) = \int_\Omega f\,(\nabla u\,(x))\,dx : u = u_0 \text{ on } \partial\Omega \right\}.$$

Associated with (P) we have the so-called *minimal surface equation*

$$(E) \quad Mu \equiv \left(1 + |\nabla u|^2\right)\Delta u - \sum_{i,j=1}^{n} u_{x_i} u_{x_j} u_{x_i x_j} = 0$$

which is the equation that any minimizer u of (P) should satisfy. In geometrical terms, this equation just expresses the fact that the corresponding surface Σ has everywhere vanishing *mean curvature*.

Case 2: parametric surfaces. Nonparametric surfaces are clearly too restrictive from the geometrical point of view and one is led to consider *parametric surfaces*. These are sets $\Sigma \subset \mathbb{R}^{n+1}$ so that there exist a domain $\Omega \subset \mathbb{R}^n$ and a map $v : \overline{\Omega} \to \mathbb{R}^{n+1}$ such that

$$\Sigma = v\left(\overline{\Omega}\right) = \left\{ v\,(x) : x \in \overline{\Omega} \right\}.$$

For example, when $n = 2$ and $v = v\,(x_1, x_2) \in \mathbb{R}^3$, if we denote by $v_{x_1} \times v_{x_2}$ the normal to the surface (where $a \times b$ stands for the vectorial product of $a, b \in \mathbb{R}^3$ and $v_{x_1} = \partial v / \partial x_1$, $v_{x_2} = \partial v / \partial x_2$) we find that the area is given by

$$\text{Area}\,(\Sigma) = J\,(v) = \iint_\Omega |v_{x_1} \times v_{x_2}|\,dx_1 dx_2\,.$$

In terms of the notations introduced at the beginning of the present section we have $n = 2$ and $N = 3$.

Example: isoperimetric inequality. Let $A \subset \mathbb{R}^2$ be a bounded open set whose boundary, ∂A, is a sufficiently regular simple closed curve. Denote by

$L(\partial A)$ the length of the boundary and by $M(A)$ the measure (the area) of A. The isoperimetric inequality states that

$$[L(\partial A)]^2 - 4\pi M(A) \geq 0.$$

Furthermore, equality holds if and only if A is a disk (i.e. ∂A is a circle).

We can rewrite it into our formalism (here $n = 1$ and $N = 2$) by parametrizing the curve

$$\partial A = \left\{ u(x) = \left(u^1(x), u^2(x) \right) : x \in [a, b] \right\}$$

and setting

$$L(\partial A) = L(u) = \int_a^b \sqrt{\left((u^1)'\right)^2 + \left((u^2)'\right)^2}\, dx$$

$$M(A) = M(u) = \frac{1}{2} \int_a^b \left(u^1 (u^2)' - u^2 (u^1)' \right) dx = \int_a^b u^1 (u^2)'\, dx.$$

The problem is then to show that

$$(P) \quad \inf \{ L(u) : M(u) = 1;\ u(a) = u(b) \} = 2\sqrt{\pi}.$$

The problem can then be generalized to open sets $A \subset \mathbb{R}^n$ with sufficiently regular boundary, ∂A, and it reads as

$$[L(\partial A)]^n - n^n \omega_n [M(A)]^{n-1} \geq 0$$

where ω_n is the measure of the unit ball of $\mathbb{R}^n$, $M(A)$ stands for the measure of A and $L(\partial A)$ for the $(n-1)$ measure of ∂A. Moreover, if A is sufficiently regular (for example, convex), there is equality if and only if A is a ball.

0.3 Presentation of the content of the monograph

To deal with problems of the type considered in the previous section, there are, roughly speaking, two ways of proceeding: the classical and the direct methods. Before describing a little more precisely these two methods, it might be enlightening to first discuss minimization problems in $\mathbb{R}^N$.

Let $X \subset \mathbb{R}^N$, $F : X \to \mathbb{R}$ and

$$(P) \quad \inf \{ F(x) : x \in X \}.$$

The first method consists, if F is continuously differentiable, of finding solutions $\overline{x} \in X$ of

$$F'(x) = 0,\ x \in X.$$

Then, by analyzing the behavior of the higher derivatives of F, we determine if $\overline{x}$ is a minimum (global or local), a maximum (global or local) or just a stationary point.

The second method consists of considering a minimizing sequence $\{x_\nu\} \subset X$ so that

$$F(x_\nu) \to \inf\{F(x) : x \in X\}.$$

We then, with appropriate hypotheses on F, prove that the sequence is compact in X, meaning that

$$x_\nu \to \overline{x} \in X, \text{ as } \nu \to \infty.$$

Finally if F is lower semicontinuous, meaning that

$$\liminf_{\nu\to\infty} F(x_\nu) \geq F(\overline{x})$$

we have indeed shown that $\overline{x}$ is a minimizer of (P).

We can proceed in a similar manner for problems of the calculus of variations. The first and second methods are then called, respectively, classical and direct methods. However, the problem is now considerably harder because we are working in infinite dimensional spaces.

Let us recall the problem under consideration

$$(P) \quad \inf\left\{I(u) = \int_\Omega f(x, u(x), \nabla u(x))\, dx : u \in X\right\} = m$$

where

- $\Omega \subset \mathbb{R}^n$, $n \geq 1$, is a bounded open set, points in Ω are denoted by $x = (x_1, \cdots, x_n)$;

- $u : \Omega \to \mathbb{R}^N$, $N \geq 1$, $u = (u^1, \cdots, u^N)$ and $\nabla u = \left(\frac{\partial u^j}{\partial x_i}\right)_{1\leq i\leq n}^{1\leq j\leq N} \in \mathbb{R}^{N\times n}$;

- $f : \overline{\Omega} \times \mathbb{R}^N \times \mathbb{R}^{N\times n} \longrightarrow \mathbb{R}$, $f = f(x, u, \xi)$, is continuous;

- X is a space of admissible functions which satisfy $u = u_0$ on $\partial\Omega$, where u_0 is a given function.

Here, contrary to the case of $\mathbb{R}^N$, we encounter a preliminary problem, namely: what is the best choice for the space X of admissible functions. A natural one seems to be $X = C^1(\overline{\Omega})$. There are several reasons, which will be clearer during the course of the book, that indicate that this is not the best choice. A better one is the *Sobolev space* $W^{1,p}(\Omega)$, $p \geq 1$. We say that $u \in W^{1,p}(\Omega)$, if u is (weakly) differentiable and if

$$\|u\|_{W^{1,p}} = \left[\int_\Omega (|u(x)|^p + |\nabla u(x)|^p)\, dx\right]^{1/p} < \infty$$

The most important properties of these spaces are recalled in Chapter 1.

In Chapter 2, we briefly discuss the *classical methods* introduced by Euler, Hamilton, Hilbert, Jacobi, Lagrange, Legendre, Weierstrass and others. The most important tool is the *Euler-Lagrange equation*, the equivalent of $F'(x) = 0$ in the finite dimensional case, that should satisfy any $\overline{u} \in C^2\left(\overline{\Omega}\right)$ minimizer of (P), namely (we write here the equation in the case $N = 1$)

$$(E) \quad \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left[f_{\xi_i}\left(x, \overline{u}, \nabla \overline{u}\right)\right] = f_u\left(x, \overline{u}, \nabla \overline{u}\right), \ \forall x \in \overline{\Omega}$$

where $f_{\xi_i} = \partial f / \partial \xi_i$ and $f_u = \partial f / \partial u$.

In the case of the Dirichlet integral

$$(P) \quad \inf\left\{I(u) = \frac{1}{2}\int_{\Omega} |\nabla u(x)|^2 \, dx : u = u_0 \text{ on } \partial\Omega\right\}$$

the Euler-Lagrange equation reduces to *Laplace equation*, namely $\Delta \overline{u} = 0$.

We immediately note that, in general, finding a C^2 solution of (E) is a difficult task, unless, perhaps, $n = 1$ or the equation (E) is linear. The next step is to know if a solution $\overline{u}$ of (E), sometimes called a stationary point of I, is, in fact, a minimizer of (P). If $(u, \xi) \to f(x, u, \xi)$ is convex for every $x \in \Omega$ then $\overline{u}$ is indeed a minimizer of (P); in the above examples this happens for the Dirichlet integral or the problem of minimal surfaces in nonparametric form. If, however, $(u, \xi) \to f(x, u, \xi)$ is not convex, several criteria, specially in the case $n = 1$, can be used to determine the nature of the stationary point. Such criteria are, for example, Jacobi, Legendre, Weierstrass, Weierstrass-Erdmann conditions or the fields theories.

In Chapters 3 and 4 we present the *direct methods* introduced by Hilbert, Lebesgue and Tonelli. The idea is to split the discussion of the problem into two parts: *existence* of minimizers in Sobolev spaces and then *regularity* of the solution. We start by establishing, in Chapter 3, the existence of minimizers of (P) in Sobolev spaces $W^{1,p}(\Omega)$. In Chapter 4 we see that, sometimes, minimizers of (P) are more regular than in a Sobolev space, for example they are in C^1 or even in C^∞, if the data Ω, f and u_0 are sufficiently regular.

We now briefly describe the ideas behind the proof of the existence of minimizers in Sobolev spaces. As for the finite dimensional case we start by considering a minimizing sequence $\{u_\nu\} \subset W^{1,p}(\Omega)$, which means that

$$I(u_\nu) \to \inf\left\{I(u) : u = u_0 \text{ on } \partial\Omega \text{ and } u \in W^{1,p}(\Omega)\right\} = m, \text{ as } \nu \to \infty.$$

The first step consists of showing that the sequence is compact, i.e., that the sequence converges to an element $\overline{u} \in W^{1,p}(\Omega)$. This, of course, depends on the topology that we have on $W^{1,p}$. The natural one is the one induced by the norm, that we call *strong convergence* and that we denote by

$$u_\nu \to \overline{u} \text{ in } W^{1,p}.$$

However, it is, in general, not an easy matter to show that the sequence converges in such a strong topology. It is often better to weaken the notion of convergence and to consider the so-called *weak convergence*, denoted by $\rightharpoonup$. To obtain that

$$u_\nu \rightharpoonup \overline{u} \text{ in } W^{1,p}, \text{ as } \nu \to \infty$$

is much easier and it is enough, for example if $p > 1$, to show (up to the extraction of a subsequence) that

$$\|u_\nu\|_{W^{1,p}} \leq \gamma$$

where γ is a constant independent of ν. Such an estimate follows, for instance, if we impose a *coercivity* assumption on the function f of the type

$$\lim_{|\xi|\to\infty} \frac{f(x,u,\xi)}{|\xi|} = +\infty, \ \forall (x,u) \in \overline{\Omega} \times \mathbb{R}.$$

We observe that the Dirichlet integral, with

$$f(x,u,\xi) = \frac{1}{2}|\xi|^2,$$

satisfies this hypothesis but not the minimal surface in nonparametric form, where

$$f(x,u,\xi) = \sqrt{1+|\xi|^2}.$$

The second step consists of showing that the functional I is lower semicontinuous with respect to weak convergence, namely

$$u_\nu \rightharpoonup \overline{u} \text{ in } W^{1,p} \ \Rightarrow \ \liminf_{\nu\to\infty} I(u_\nu) \geq I(\overline{u}).$$

We will see that this conclusion is true if

$$\xi \to f(x,u,\xi) \text{ is convex, } \forall (x,u) \in \overline{\Omega} \times \mathbb{R}.$$

Since $\{u_\nu\}$ was a minimizing sequence, we deduce that $\overline{u}$ is indeed a minimizer of (P).

In Chapter 5 we consider the problem of minimal surfaces. The methods of Chapter 3 cannot be directly applied. In fact the step of compactness of the minimizing sequences is much harder to obtain, for reasons that we explain in Chapter 5. There are, moreover, difficulties related to the geometrical nature of the problem; for instance, the type of surfaces that we consider, or the notion of area. We present a method due to Douglas and refined by Courant and Tonelli to deal with this problem. However the techniques are, in essence, direct methods similar to those of Chapter 3.

In Chapter 6 we discuss the isoperimetric inequality in $\mathbb{R}^n$. Depending on the dimension, the way of solving the problem is very different. When $n = 2$, we present a proof which is essentially the one of Hurwitz and is in the spirit of the techniques developed in Chapter 2. In higher dimensions the proof is more geometrical; it uses as a main tool the *Brunn-Minkowski theorem*.