Calculus of Variations
Lecture Notes

Riccardo Cristoferi

May 9 2016

Contents

Introduction

These lecture notes are based on the undergraduate course 21-470 Selected Topics in Analysis I gave at Carnegie Mellon University in Spring 2016.

The aim of the course was to present the basic notions and results of the so called *classical* Calculus of Variations, *i.e.*, the theory of necessary and sufficient conditions for minimizers of variational problems of the form

$$\mathcal{F}(u) := \int_\Omega f\big(x, u(x), Du(x)\big)\, \mathrm{d}x\,,$$

where $u : \Omega \subset \mathbb{R}^N \to \mathbb{R}^M$ and $f : \Omega \times \mathbb{R}^N \times \mathbb{R}^M \to \mathbb{R}$ (the *lagrangian*).
The notes are (almost) self-contained and can be read with just basic notions in analysis and the fine ability of not seeing errors[1] in the proofs.

They are divided in four parts:

(i) classical variational problems,
(ii) necessary conditions,
(iii) sufficient conditions,
(iv) *to infinity*[2]*...and beyond!*

In Chapter 1 we present four classical problems in the Calculus of Variations that we are eventually going to solve: brachistochrone, minimal surfaced of revolution, hanging cable and isoperimetric problem in the plane.

Chapter 2 is a brief review of the study minimum problems in finite dimension.

First order necessary conditions minimizers have to satisfy are presented in Chapter 3 (for one dimensional scalar problems) and in Chapter 4 (for the general case). In particular we treat the case of broken extremals, natural boundary conditions, isoperimetric problems and holomic constraints.

Chapter 5 is then devoted to the second order necessary conditions: coercivity of the second variation, Legendre-Hadamard condition and Weierstrass's condition. The last two are presented only for the case of one dimensional variational problems.

Null-Lagrangians are presented in Chapter 6, while the relation between Lagrange multipliers and eigenvalues of the Laplace operator is in investigated in Chapter 7.

The sufficient conditions for (weak and strong) local minimality of one dimensional scalar variational problems will be treated in Chapter 8.

Due to the peculiarity and the importance of the problem, Chapter 9 is entirely devoted to the presentation of three different proofs of the isoperimetric inequality in the plane, each one of a different flavor.

Finally, we give a very brief overview to the modern approach to the Calculus of Variations.

Most of the material of the classical part is based on [5].

[1] Of course, every error you may find was clearly made on purpose, just to add a bit of fun to the static (and stable) perfection of math!
[2] of the modern theory

CHAPTER 1

Some classical problems

We would like to start by introducing some classical problems in the Calculus of Variations. Besides their own interest, these problems will give us a reason to introduce and to study the theoretical problems we will treat.

1.1. The brachistochrone problem

This problem is consider the *birth* of the Calculus of Variations, and was posed by Jackob Bernoulli in 1696:

Let us consider two points P and Q in a vertical plane. Find the smooth curve joining the two points such that a particle subject to gravity starting from rest at the higher point will slide without friction along the curve to reach the lower point in the shortest possible time

The solution of such a problem (if it exists!) is called *brachistochrone* (from the greek: brachistos=shortest, chronos=time).

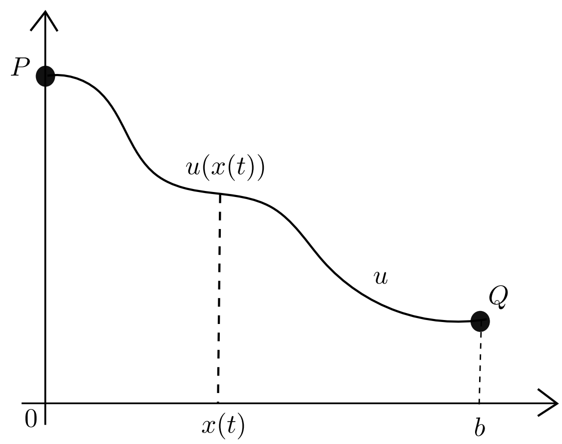

FIGURE 1. The mathematical setting of the brachistochrone problem.

We now want to find a mathematical formulation of the problem. For, let us assume that the points P and Q are given by

$$P = (0, y_1), \qquad Q = (b, y_2).$$

Moreover, assume that the solution is given by a graph of a smooth function

$$u : [0, b] \to \mathbb{R}.$$

At time t, the particle will be at the position $(x(t), u(x(t)))$ (see Figure 1). The length of the path traveled by the particle is

$$l(t) := \int_0^{x(t)} \sqrt{1 + \big(u'(x(t))\big)^2}\, \mathrm{d}x\,.$$

Thus, the velocity of the particle can be written as

$$v(t) := \frac{\mathrm{d}l}{\mathrm{d}s}_{|s=t} = x'(t)\sqrt{1 + \big(u'(x(t))\big)^2}\,,$$

from which we deduce

$$x'(t) = \frac{v(t)}{\sqrt{1 + \big(u'(x(t))\big)^2}}\,. \tag{1.1}$$

On the other hand, it is also possible to compute the velocity $v(t)$ by using the conservation of the energy. So, let $m > 0$ be the mass of the particle. Then

$$\frac{m}{2}(v(0))^2 + mgu(x(0)) = \frac{m}{2}(v(t))^2 + mgu(x(t))\,.$$

Recalling that, by hypothesis, $v(0) = 0$, we have that

$$(v(t))^2 = 2g(y_1 - u(x(t)))\,.$$

This forces the right-hand side to be non negative. The physical interpretation is that the particle will no go above the initial vertical position. By inserting this expression in (1.1) we get

$$x'(t) = \frac{\sqrt{2g(y_1 - u(x(t)))}}{\sqrt{1 + \big(u'(x(t))\big)^2}}\,.$$

We now make another reasonable assumption: the particle won't go back. This is reflected in the mathematical assumption that the function

$$t \mapsto x(t)$$

is invertible. Thus, we have that

$$t'(x) = \frac{1}{x'(t)} = \frac{\sqrt{1 + \big(u'(x)\big)^2}}{\sqrt{2g(y_1 - u(x))}}\,.$$

Hence, the total time the particle needs to go from P to Q along the path described by the function u is

$$\mathcal{T}(u) := \frac{1}{2g}\int_0^b \frac{\sqrt{1 + \big(u'(x)\big)^2}}{\sqrt{2g(y_1 - u(x))}}\, \mathrm{d}x\,.$$

Thus, the mathematical formulation of the brachistochrone problem is the following:

$$\min_{u \in \mathcal{A}} \mathcal{T}(u)\,,$$

where the admissible class of functions is

$$\mathcal{A} := \{u \in C^1([a, b]) \;:\; u(0) = y_1,\, u(b) = y_2\}\,.$$

A variant of the brachistochrone problem. A variant of the above problem was stated by Jakob Bernoulli one year later. The problem is the same, but the point Q is just required to lies on the vertical line $\{x = b\}$. Clearly, a solution of this problem (if it exists!) must be a solution of the preceding problem for the particular value of y_2 assumed at the minimum point.

1.2. The hanging cable problem

Suppose we want to find the shape of an inextensible cable hanging under its own weight, when its ends are pinned at two given points P and Q.

In order to derive the mathematical formulation, we recall that the quantity to minimize is the potential energy of the cable. Let us suppose the cable has a uniform cross-section and has a uniform density ρ. Moreover we will make the (reasonable) assumption that the shape assumed by the cable is the graph of a regular function $u \in C^1([a,b])$ (see Figure 2).

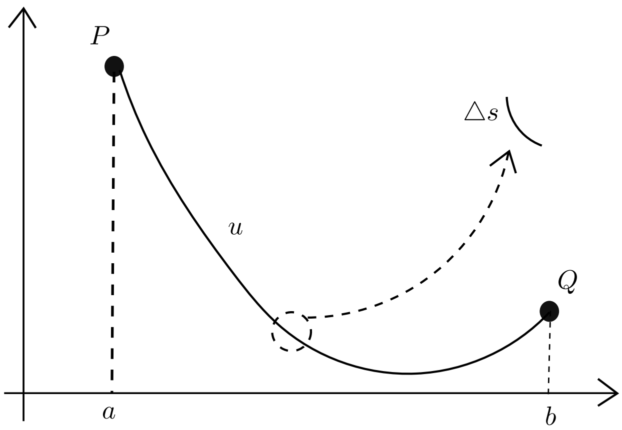

FIGURE 2. The mathematical setting of the hanging cable problem.

First of all, let us denote by l the total length of the cable. Then every admissible function will have to satisfy

$$\int_a^b \sqrt{1+(u'(x))^2}\,\mathrm{d}x = l\,.$$

To write down the total potential energy of the cable, let us use a physical argument: consider a small piece of the cable $\triangle s$ at height y. Its mass will be $\rho\triangle s$. Thus, the potential energy of this piece will be

$$gy\rho\triangle s\,.$$

Summing up all these contributions we get

$$\mathcal{F}(u) := \int_{\mathrm{graph}(u)} g\rho u\,\mathrm{d}\sigma(s)\,,$$

where $\sigma(s)$ is the measure on the graph of u. This integral can be rewritten as

$$\mathcal{F}(u) = \int_a^b g\rho u(x)\sqrt{1+(u'(x))^2}\,\mathrm{d}x\,.$$

The problem is thus

$$\min_{u\in\mathcal{A}} F(u)\,,$$

where the admissible class of functions is

$$\mathcal{A} := \{u \in C^1([a,b]) \;:\; u(0)=y_1,\, u(b)=y_2,\, \int_a^b \sqrt{1+(u'(x))^2}\,\mathrm{d}x = l\}\,.$$

1.3. Minimal surfaces of revolution

Let us consider a multidimensional problem that can be reduced to a one dimensional one. Let us consider two circumferences C_1 and C_2 of radius r_1 and r_2 respectively. We want to find, among the surfaces in $\mathbb{R}^3$ having those circumferences as a boundary, the one having minimal surface area.

In order to make this multidimensional problem a one dimensional one, we restrict ourselves to the case of the so called *surfaces of revolution*, *i.e.*, surfaces that are obtained by rotating the graph of a function $u : [a, b] \to \mathbb{R}$ with respect to the x-axes (see Figure 3).

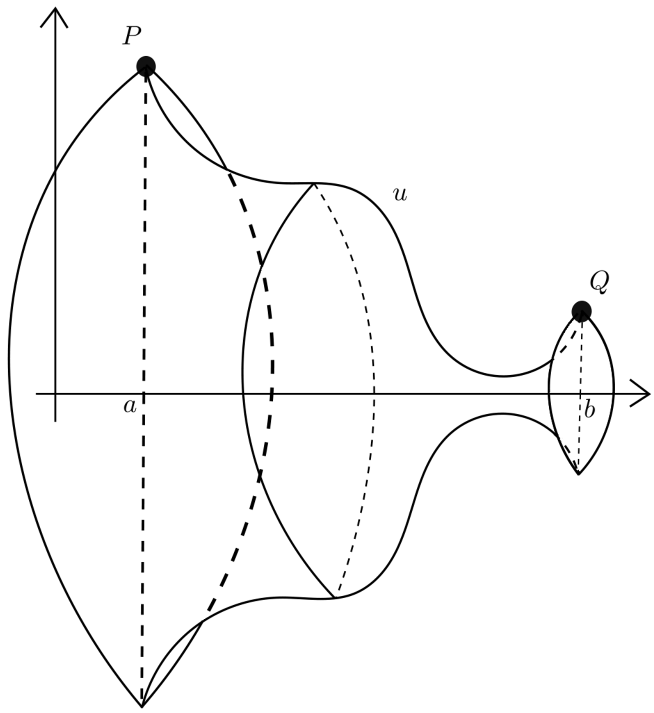

FIGURE 3. A surface of revolution.

Under these assumptions, the surface area of the surface generated by the function u (assuming u to be of class C^1 such that $u > 0$, that means no auto intersections of the surface) can be written as

$$F(u) := \int_a^b 2\pi u(x)\sqrt{1 + (u'(x))^2}\, \mathrm{d}x\,.$$

Thus we are leading to the following minimization problem:

$$\min_{u \in \mathcal{A}} \mathcal{F}(u)\,,$$

where

$$\mathcal{A} := \{u \in C^1([a, b]),\, u(a) = r_1, u(b) = r_2, u > 0\}\,.$$

The novelty of this problem is that we ask $u > 0$. But this problem has another important peculiarity: it does not always have a solution. Indeed, we will see that a solution will exists only if $b - a$ is sufficiently small with respect to r_1 and r_2.

1.4. The isoperimetric problem

This is referred as the oldest[1] problem in the Calculus of Variations. The myth says the queen Dido landed on the coast of North Africa and asked to the locals for a small piece of land as a temporary place where to stay. She was told that she could have as many land as she could enclose by an oxhide. no such a good answer! But queen Dido was smart: she cut the oxhide into very thin stips that she used to encircle an entire hill. This is, according to the legend, how Carthage has been founded.

The above problem can be stated as follows:

Find, among all planar simple[2] *closed curves of* fixed length*, the one that encloses the* maximum *area.*

1.4.1. An equivalent problem. Before translating this problem into mathematical symbols, we would like to formulate another problem:

Find, among all planar simple closed curves enclosing a fixed area*, the one with* minimal *length.*

These two problems are equivalent, that is: a curve solves the first one if and only if it solves the second one. Let us be more precise: let γ be a simple closed curve in the plane and suppose it encloses maximum area among all the curves with fixed length $l > 0$. Let $m > 0$ be this area. Then γ is the simple closed curve in the plane having minimal length among all the curves enclosing area m.

To prove the above statement we introduce some objects: for a simple closed planar curve γ, let E_γ be the region enclosed by γ, and let $A(E_\gamma)$ its area. Finally, let $L(\gamma)$ denotes the length of the curve. We the following quantity:

$$\frac{A(E_\gamma)}{L^2(\gamma)},$$

called the *isoperimetric ratio*. The reason to introduce it is that it is invariant under dilation of the space, *i.e.*, for any curve γ and any $\lambda > 0$ let us denote by $\lambda\gamma$ the image of the curve γ under the dilatation $(x, y) \mapsto (\lambda x, \lambda y)$. Then, we have that

$$\frac{A(E_{\lambda\gamma})}{L^2(\lambda\gamma)} = \frac{\lambda^2 A(E_\gamma)}{\lambda^2 L^2(\gamma)} = \frac{A(E_\gamma)}{L^2(\gamma)}. \tag{1.2}$$

Moreover, the two statements above can be phrased as the problem of maximizing the isoperimetric ratio under the respective constraints. Indeed the first problem is

$$\max_{A(E_\gamma)=m} \frac{A(E_\gamma)}{L^2(\gamma)},$$

while the second is

$$\max_{L(\gamma)=l} \frac{A(E_\gamma)}{L^2(\gamma)}.$$

[1]The brachistochrone one was the first mathematically stated and solved!

[2]See Definition 11.2

In order to prove the equivalence of the two problems, we prove that:

$$\max_{A(E_\gamma)=m} \frac{A(E_\gamma)}{L^2(\gamma)} = \max_{L(\gamma)=l} \frac{A(E_\gamma)}{L^2(\gamma)} .$$

We first prove $\leq$: let $\bar\gamma$ be a solution[3] of the left-hand side. Then consider the curve

$$\widetilde\gamma := \frac{l}{L(\bar\gamma)}\bar\gamma .$$

Then we have that $L(\widetilde\gamma) = l$, and thus $\widetilde\gamma$ is an admissible competitor for the problem on the right-hand side. Thus

$$\frac{A(E_{\bar\gamma})}{L^2(\bar\gamma)} = \frac{A(E_{\widetilde\gamma})}{L^2(\widetilde\gamma)} \leq \max_{L(\gamma)=l} \frac{A(E_\gamma)}{L^2(\gamma)} ,$$

where in the first equality we have used (1.2). This proves $\leq$.

In order to prove $\geq$ we reason in a similar way.

1.4.2. The mathematical formulation. In order to formulate the problem in mathematical terms, let us consider a simple closed curve $\gamma : [0,1] \to \mathbb{R}^2$ and let us denotes its components by γ_1 and γ_2 respectively. Then we have that

$$L(\gamma) = \int_0^1 \sqrt{(\gamma_1'(x))^2 + (\gamma_2'(x))^2}\, \mathrm{d}x ,$$

and

$$\begin{aligned} A(E_\gamma) &= \int_0^1 \gamma_1(x)\gamma_2'(x)\, \mathrm{d}x = -\int_0^1 \gamma_2(x)\gamma_1'(x)\, \mathrm{d}x \\ &= \frac{1}{2}\int_0^1 (\gamma_1(x)\gamma_2'(x) - \gamma_2(x)\gamma_1'(x))\, \mathrm{d}x , \end{aligned}$$

where these last equalities follow from the Gauss-Green Theorem (see Appendix, section 11.1). Thus, the mathematical formulation of the isoperimetric problem is the following:

$$\min\{L(\gamma) \ : \ \gamma : [0,1] \to \mathbb{R}^2 \text{ is a simple closed curve with } A(E_\gamma) = 1\} .$$

Equivalently, we can consider the following problem

$$\max\{A(E) \ : \ E\text{is a region in the plane enclosed by a simple closed curve}\gamma\text{with } L(\gamma) = 1\} .$$

The novelty of this problem is that it is a multidimensional parametric problem, that is, a problem where out unknown is not a function, but a more complex object that can be described by using a (set of) parameter(s). As we will see in Chapter 9, the solution of such a problem requires *ad hoc* techniques.

[3] Since we are in a classical setting, we give for grant that a solution actually exists!

1.4.3. The non parametric case. We now want to consider a special variant of the above problem, that is when we ask our curve to be made by two parts: a segment and a graph over this segments in such a way that the resulting curve a simple and closed one (see Figure 4).

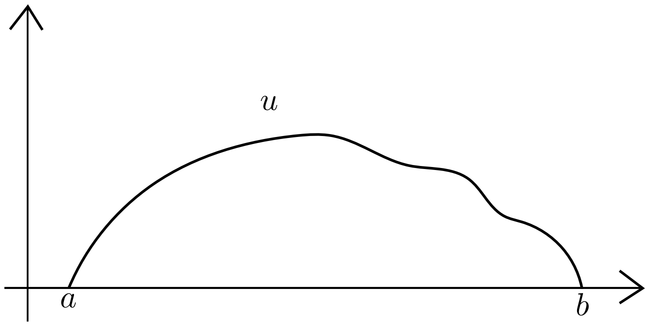

FIGURE 4. Ad admissible curve for the non parametric isoperimetric problem.

In other words, we consider functions

$$u : [a, b] \to \mathbb{R}\,, \qquad \text{s.t. } u(a) = u(b) = 0,\; u > 0 \text{ on } (a, b)\,.$$

In such a situation, since the 'segment' part of the curve is fixed, the constrain we have to impose for the length of the graph itself, *i.e.*, we have to impose that

$$\int_a^b \sqrt{1 + (u'(x))^2}\, \mathrm{d}x = l\,.$$

Moreover, we have a simpler formula for the area of the enclosed region:

$$\int_1^b u(x)\, \mathrm{d}x\,.$$

Thus, in the non parametric case, the isoperimetric problem becomes

$$\min\left\{ \int_1^b u(x)\, \mathrm{d}x \;:\; u \in C^1([a, b]),\, u(a) = u(b) = 0,\, \int_a^b \sqrt{1 + (u'(x))^2}\, \mathrm{d}x = l \right\}.$$

CHAPTER 2

Minimization on $\mathbb{R}^N$

In this section we want to recall some well known facts about sufficient and necessary conditions for local minimizers for functions $f:\mathbb{R}^N \to \mathbb{R}$. The reason to do so is that these simple concepts will be at the bootom of the idea needed for developing the results we will present about necessary and sufficient conditions in the classical Calculus of Variations.

2.1. The one dimensional case

Let us consider a function $f:(a,b)\to\mathbb{R}$.

Definition 2.1. We say that a point $\bar{x}\in(a,b)$ is a *point of local minimum for* f, if there exists $\delta>0$ such that

$$f(\bar{x}) \le f(x)\,,$$

for all $x\in(\bar{x}-\delta,\bar{x}+\delta)$. Moreover, we say that the local minimum is *isolated* whether equality in the above condition holds true only if $x=\bar{x}$.

The first necessary condition for minimality we present was discovered by Fermat in 1635. Let $\bar{x}\in(a,b)$ be a point of local minimum for f. Then it is easy to see that the difference quotient (for $|h|\ll 1$, *i.e.*, h is very small)

$$\frac{f(\bar{x}+h)-f(\bar{x})}{h}\,,$$

has to change sign accordling to the sign of h. In particular, if f is differentiable at $\bar{x}$, then we must have

$$f'(\bar{x}) = 0\,.$$

Definition 2.2. A point $x\in(a,b)$ for which $f'(x)=0$ is called a *critical point* of f.

Thus, being a critical point is a necessary condition for local minimality, but not a sufficient one. Indeed, if we consider the function $f(x):=x^3$, we have that $x=0$ is a critical point, but not a point of local minimum for f.

Another necessary condition for local minimizers can be deduce by assuming f to be twice differentiable in $(\bar{x}-\varepsilon,\bar{x}+\varepsilon)$, for some $\varepsilon>0$. Under there hypothesis, if $\bar{x}$ is a point of local minimum for f, by Taylor's formula we have that

$$f(x) = f(\bar{x}) + f'(\bar{x}) + \frac{1}{2}f''(\xi)(\xi-\bar{x})^2 = f(\bar{x}) + \frac{1}{2}f''(\xi)(\xi-\bar{x})^2\,,$$

for some ξ between x and $\bar{x}$, where we have used the fact that a local minimum is a critical point. Thus, by taking the limit as $x\to\bar{x}$, we get

$$f''(\bar{x}) \ge 0\,.$$

REMARK 2.3. Notice that we cannot conclude that $f''(x) \geq 0$ holds for all $x \in (\bar{x}-\varepsilon, \bar{x}+\varepsilon)$, form some $\varepsilon > 0$. Take for example the C^2 function

$$f(x) := x^6 + \sin\Big(\frac{1}{x}\Big)(x^8 - x^6)\,.$$

Then $x = 0$ is a point of global minimum for f, but f'' oscillates in any neighborhood of $x = 0$.

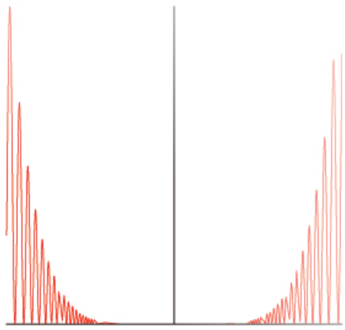

FIGURE 1. The graph of the function f close to the origin.

DEFINITION 2.4. We say that a critical point $\bar{x} \in \mathbb{R}$ is *stable* if $f''(\bar{x}) \geq 0$. We say that it is *strictly stable* if $f''(\bar{x}) > 0$.

As before, by considering at the function $f(x) := x^3$, we see that being stable is not sufficient for beeing a local minimum. But it turns out that a strictly stable critical point is a local minimum for f. To see it, just apply Taylor's formula.

Thus, it holds that:

Necessary conditions:

$$\bar{x} \text{ is a point of local minimum} \;\Rightarrow\; \begin{cases} \bar{x} \text{ is critical} \\ \text{and stable} \end{cases}$$

Sufficient conditions:

$$\begin{cases} \bar{x} \text{ is critical} \\ \text{and strictly stable} \end{cases} \;\Rightarrow\; \bar{x} \text{ is an isolated point of local minimum}$$

REMARK 2.5. We notice that there is a gap between necessary and sufficient conditions for local minimality. So, we expect the same phenomenon to hold also for problems in the Calculus of Variations.

Question: what about conditions for detecting points that are local minima, but not isolated ones?

2.2. The multidimensional case

We now consider functions $f : \Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^N$ is an open set. By using the same arguments as above, adapted to the multidimensional setting, we can prove similar necessary and sufficient conditions. We first need a couple of definitions.

DEFINITION 2.6. A point $\bar{x} \in \Omega$ for which $\nabla f(\bar{x}) = 0$ is called a *critical point*.

DEFINITION 2.7. A critical point $\bar{x} \in \Omega$ for which $D^2 f(\bar{x}) \geq 0$ is called *stable*. It is called *strictly stable* whenever $D^2 f(\bar{x}) > 0$.

DEFINITION 2.8. A critical point $\bar{x} \in \Omega$ for which $D^2 f(\bar{x}) \geq 0$ does not have 0 as an eigenvalue is called *non degenerate*.

DEFINITION 2.9. An $N \times N$ matrix A is said to be *semi-positive definite*, and we write $A \geq 0$, if

$$Av \cdot v \geq 0\,,$$

for each $v \in \mathbb{R}^N$, where $\cdot$ denotes the standard scalar product on $\mathbb{R}^N$.

Moreover, we say that matrix A is *positive definite*, and we write $A > 0$, if

$$Av \cdot v > 0\,,$$

for each $v \in \mathbb{R}^N \setminus \{0\}$.

In our case, the matrix we are dealing with is $D^2 f(\bar{x})$ that, thanks to Schwarz's theorem, happens to be symmetric. Thus, all its eigenvalues $\lambda_1, \cdots, \lambda_N$ are *real*. This allows to rephrase the conditions of being (semi-)positive definite with respect to the sign of the eigenvalues. Indeed, it holds that

$$D^2 f(\bar{x}) \geq 0 \Leftrightarrow \lambda_i \geq 0 \quad \forall i = 1, \ldots, N\,,$$

and that

$$D^2 f(\bar{x}) > 0 \Leftrightarrow \lambda_i > 0 \quad \forall i = 1, \ldots, N\,,$$

Usually these conditions are the one used to checked whether $D^2 f(\bar{x})$ is (semi)-positive or not.

Finally, we recall that similar necessary and sufficient conditions hold hold true for the multi dimensional case, as well as for the gap between the two.

Necessary conditions:

$$\bar{x} \text{ is a point of local minimum } \Rightarrow \begin{cases} \bar{x} \text{ is critical} \\ \text{and stable} \end{cases}$$

Sufficient conditions:

$$\begin{cases} \bar{x} \text{ is critical} \\ \text{and strictly stable} \end{cases} \Rightarrow \quad \bar{x} \text{ is an isolated point of local minimum}$$

2.3. Local representation close to non degenerate points

In the previous section we have seen that a strictly stable critical point is a local minimum. When we think to an isolated local minimum, the picture we have in mind is the one of the function $x \mapsto |x|^2$ (see Figure 2). Next theorem tells us that our intuition was correct.

LEMMA 2.10 (Morse Lemma). *Let $f : \mathbb{R}^N \to \mathbb{R}$ be a function of class C^3, and let $\bar{x} \in \mathbb{R}^N$ be a non degenerate critical point of f. Then there exists a diffeomorphism of class C^1*

$$\Phi : U \to B_\delta(\bar{x}) \,,$$

where $U \subset \mathbb{R}^N$ is an open set and $B_\delta(\bar{x})$ denotes the ball of radius δ centered at $\bar{x}$, such that

$$f\big(\Phi(y)\big) = f(\bar{x}) - \sum_{i=1}^{q} |y_i|^2 + \sum_{i=q+1}^{N} |y_i|^2 \,,$$

where $q \in \{1, \ldots, N\}$ is the dimension of the space generated by all the eigenvectors with negative eigenvalues, and it is call the index.

Thus, the behavior of a function close to non degenerate critical points is completely understood: up to a change of coordinates, there are directions where it behaves like t^2 and others where it is like $-t^2$. In particular, when the point $\bar{x}$ happens to be a non degenerate local minimum, the local behavior of f is the same as those of $x \mapsto (x - \bar{x})^2$.

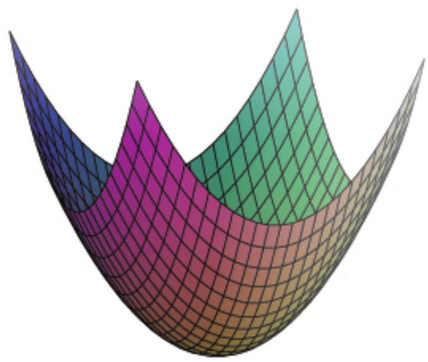

FIGURE 2. The typical behavior of a function near a non degenerate local minimum point.

On the other hand, the behavior of a function close to a degenerate critical point can be very wild, as the following three examples will show us.

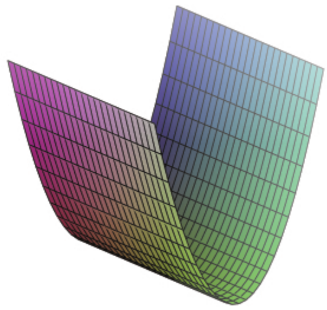

FIGURE 3. The graph of the function $f(x,y) := x^2$.

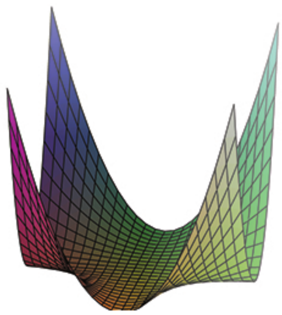

FIGURE 4. The graph of the function $f(x,y) := x^2y^2$.

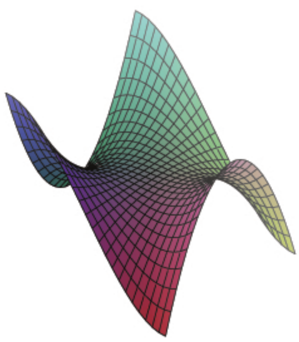

FIGURE 5. The graph of the function $f(x,y) := x(x^2 - 3y^2)$ (the so called *monkey saddle*).

2.4. Convex functions

In this section we would like to recall some facts about convex functions.

DEFINITION 2.11. A function $f : C \to \mathbb{R}$, where $C \subset \mathbb{R}^N$ is a convex set, is said to be *convex* if

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)\,,$$

for all $x_1, x_2 \in C$ and all $\lambda \in [0,1]$. Moreover, the function is called *strictly convex* whenever equality holds only if $x_1 = x_2$, or $\lambda \in \{0,1\}$.

REMARK 2.12. Anytime we will speak about convex functions, we will assume that they are defined on a convex set.

We now present some important properties of convex functions. A convex function is

- continuous,
- differentiable almost everywhere[1],
- twice differential for almost-every point. This is Alexandrov's Theorem.

It is possible to characterize convexity with properties of the gradient or the hessian of the function.

PROPOSITION 2.13. *Let $f : C \to \mathbb{R}$, where $C \subset \mathbb{R}^N$ is a convex set. It hold:*

(1) if f is of class C^1, then

$$f \text{ is convex } \Leftrightarrow f(y) \geq f(x) + \nabla f(x)\cdot(y-x)\,.$$

(2) If f is of class C^2, then

$$f \text{ is convex } \Leftrightarrow D^2 f \geq 0\,.$$

From the first characterization of the above proposition[2] we can deduce the following result, saying that critical points happen to be minimizers.

COROLLARY 2.14. *Let $f : C \to \mathbb{R}$ be a convex function. Then any critical point is a minimizer of f. In particular, if f is strictly convex, then there can be only one minimizer.*

REMARK 2.15. Notice that in the above result, we are not claiming that a convex function must have a minimizer. Just think to the convex function $f(x) := e^x$ on $\mathbb{R}$.

Finally, we present a result concerning the relation between convex functions and integrals. In order to justify it, just recall that

$$f\Big(\sum_{i=1}^n a_i x_i\Big) \leq \sum_{i=1}^n a_i f(x_i)\,,$$

for every $x_1, \ldots, x_n \in \mathbb{R}$ and every $a_1, \ldots, a_n \geq 0$ such that $\sum_{i=1}^n a_i = 1$. Formally sending $n \to \infty$, we get the following result.

THEOREM 2.16 (Jensen's inequality). *Let $f : \Omega \to \mathbb{R}$ be a convex function, where $\Omega \subset \mathbb{R}^N$ is an open set, and let us denote by $|\Omega|$ its (Lebesgue) measure. Then*

$$f\Big(\int_\Omega u(x)\,\mathrm{d}x\Big) \leq \frac{1}{|\Omega|}\int_\Omega f\big(u(x)\big)\,\mathrm{d}x\,,$$

for each function $u \in L^1(\Omega) := \{v : \Omega \to \mathbb{R} \;:\; \exists \int_\Omega |u(x)|\,\mathrm{d}x < \infty\}$.

---

[1] With respect to the Lebesgue measure.

[2] For a proof of such a characterization, see Appendix Section 11.2.

2.5. On the notion of differentiability

We now want to recall two important notions of differentiability.

DEFINITION 2.17. A function $f : \mathbb{R}^N \to \mathbb{R}$ is called *Fréchet differentiable* at the point $\bar{x} \in \mathbb{R}^N$ if there exists a linear map $L : \mathbb{R}^N \to \mathbb{R}$ such that

$$f(\bar{x} + v) = f(\bar{x}) + L(v) + o(|v|)\,,$$

as $|v| \to 0$, for all $v \in \mathbb{R}^N$.

In this case, the function L is called the *Fréchet differential* of f at $\bar{x}$.

REMARK 2.18. The above condition means that there exists $\delta > 0$ such that if $|v| < \delta$, we can write

$$f(\bar{x} + v) = f(\bar{x}) + L(v) + g(v)\,,$$

where $g : \mathbb{R}^N \to \mathbb{R}$ is a function such that $\frac{g(v)}{|v|} \to 0$ as $|v| \to 0$. This is the usual notion of differentiability.

DEFINITION 2.19. A function $f : \mathbb{R}^N \to \mathbb{R}$ is called *Gateaux differentiable* at the point $\bar{x} \in \mathbb{R}^N$ if there exists a map $L : \mathbb{R}^N \to \mathbb{R}$ such that

$$f(\bar{x} + \varepsilon v) = f(\bar{x}) + \varepsilon L(v) + o(\varepsilon)\,,$$

as $|\varepsilon| \to 0$, for all $v \in \mathbb{R}^N$.

In this case, the function L is called the *Gateaux differential* of f at $\bar{x}$.

REMARK 2.20. The above condition means that there exists $\delta > 0$ such that if $\varepsilon < \delta$, we can write

$$f(\bar{x} + \varepsilon v) = f(\bar{x}) + \varepsilon L(v) + g(\varepsilon)\,,$$

where $g : \mathbb{R} \to \mathbb{R}$ is a function such that $\frac{g(\varepsilon)}{\varepsilon} \to 0$ as $\varepsilon \to 0$.

Notice that here we are *not* requiring L to be linear. The map L can be also characterized as follows:

$$L(v) = \lim_{\varepsilon \to 0} \frac{f(\bar{x} + \varepsilon v) - f(\bar{x})}{\varepsilon}\,,$$

for all $v \in \mathbb{R}^N$.

The two notions seem very related, but they are in fact very different (when $N > 1$!). Indeed Fréchet differentiability is the requirement that there exists a tangent plane to the graph of f at the point $(\bar{x}, f(\bar{x}))$, while Gateaux differentiability concerns with the existence of directional derivatives. Clearly, a function f that is Fréchet differentiable is also Gateaux differentiable, but the opposite is not true, as can be seen in the following example:

EXAMPLE 2.21. Let us consider the following function (introduced by Peano):

$$f(x) := \begin{cases} \left(\frac{xy^2}{x^2+y^4}\right)^2 & (x,y) \neq (0,0)\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

This function is Gateaux differentiable at the point $(0,0)$, but not Fréchet differentiable. Indeed, for any $(a,b) \in \mathbb{R}^2$ we have that

$$L\big((a,b)\big) = \lim_{\varepsilon \to 0} \Big(\frac{\varepsilon}{\varepsilon^2 + \varepsilon^4}\Big)^2 \frac{ab^b}{a^b + b^4} = 0\,,$$

and thus the Gateaux derivative of f at $(0,0)$ is $L \equiv 0$.

On the other hand, let us consider the family of vectors

$$v_\alpha := (\alpha b^2, b)\,.$$

Then

$$\lim_{|v_\alpha|\to 0} \frac{f(v_\alpha)}{|v_\alpha|} = \frac{\alpha^2}{(\alpha^2+1)^2} \frac{1}{|b|\sqrt{1+\alpha^2 b^2}} \to \infty\,,$$

and thus f is not Fréchet differentiable at $(0,0)$.

Moreover, it is also possible to have existence of the Gateaux derivative, but the Gateaux derivative may fail to be linear, as the following example shows.

EXAMPLE 2.22. Let us consider the following function:

$$f(x) := \begin{cases} \frac{x^3}{x^2+y^2} & (x,y) \neq (0,0)\,, \\ 0 & \text{otherwise}\ . \end{cases}$$

Then the Gateaux derivative L of f at $(0,0)$ is given by

$$L\big((a,b)\big) := \begin{cases} \frac{a^3}{a^2+b^2} & (x,y) \neq (0,0)\,, \\ 0 & \text{otherwise}\ , \end{cases}$$

that is a non linear function.

CHAPTER 3

First order necessary conditions for one dimensional scalar functions

3.1. The first variation - C^2 theory

We now want to apply the previous machinery to the case of *variational integrals*, *i.e.*, functionals $f : C^1([a,b]) \to \mathbb{R}$ of the form

$$\mathcal{F}(u) := \int_a^b f\big(x, u(x), u'(x)\big)\,\mathrm{d}x\,,$$

in order to derive some (first order) necessary conditions minimizers have to satisfy.

3.1.1. The Euler-Lagrange equation - strong form. We start by fixing some notation.

Definition 3.1. The function $f : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is called *lagrangian*.

So, fix a function $u \in C^1([a,b])$ and take a direction $\varphi \in C^1([a,b])$. The idea of Euler was to consider the directional derivative of $\mathcal{F}$ at u_0 in the direction φ, that is, to consider the function $\Phi : (-\varepsilon, \varepsilon) \to \mathbb{R}$ given by

$$\Phi(\varepsilon) := \mathcal{F}(u + \varepsilon\varphi)\,, \tag{3.1}$$

and to derive it.

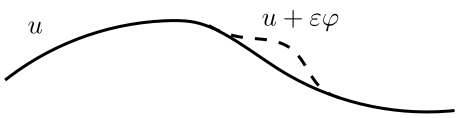

Figure 1. The variations of Euler

For, we need the following technical result.

Lemma 3.2. *Let $g : [a,b] \times [c,d] \to \mathbb{R}$ be a continuous function such that there exists the partial derivative with respect to the second variable and it is continuous. Let us define the function $G : [c,d] \to \mathbb{R}$ as*

$$G(t) := \int_a^b g(x,t)\,\mathrm{d}x\,.$$

Then G is a function of class C^1 and

$$G'(t) = \int_a^b \frac{\partial g}{\partial t}(x,t)\,\mathrm{d}x\,.$$

PROOF. Fix $t \in [c,d]$ and, for h sufficiently small, let us consider the incremental ratio

$$\frac{G(t+h)-G(t)}{h} = \int_a^b \frac{g(x,t+h)-g(x,t)}{h}\,\mathrm{d}x = \int_a^b \frac{\partial g}{\partial t} g(x,t+\theta h)\,\mathrm{d}x\,,$$

where in the last step we have used the mean value theorem and $\theta \in (0,1)$ depends on x,t,h. Since $\frac{\partial g}{\partial t}$ is continuous, it is uniformly continuous on $[c,d]$. Thus, fixed $\varepsilon > 0$, we can find $\delta > 0$ such that, if $|h| < \delta$, then

$$\Big|\frac{\partial g}{\partial t} g(x,t+h) - \frac{\partial g}{\partial t} g(x,t)\Big| < \varepsilon\,,$$

for all $x \in [a,b]$. Then

$$\begin{aligned}
\Big|\frac{G(t+h)-G(t)}{h} - \frac{\partial g}{\partial t} g(x,t)\,\mathrm{d}x\Big| &= \Big|\int_a^b \Big(\frac{\partial g}{\partial t} g(x,t+\theta h) - \frac{\partial g}{\partial t} g(x,t)\,\mathrm{d}x\Big)\Big| \\
&\le \int_a^b \Big|\frac{\partial g}{\partial t} g(x,t+\theta h) - \frac{\partial g}{\partial t} g(x,t)\Big|\,\mathrm{d}x \\
&\le \varepsilon(b-a)\,.
\end{aligned}$$

Since ε is arbitrary, we conclude. □

So, let us suppose our lagrangian f to be of class C^1. Actually, we only need f to have continuous partial derivatives with respect to the variables p and ξ. The above result tells us that the function Φ is differentiable at $\varepsilon = 0$, and

$$\Phi'(0) = \int_a^b \big[f_p\big(x,u(x),u'(x)\big)\varphi(x) + f_\xi\big(x,u(x),u'(x)\big)\varphi'(x)\big]\,\mathrm{d}x\,.$$

DEFINITION 3.3. We introduce the operator $\delta\mathcal{F} : C^1([a,b]) \times C^1([a,b]) \to \mathbb{R}$ as follows:

$$\delta\mathcal{F}(u,\varphi) := \Phi'(0)\,,$$

where Φ is defined as in (3.1), provided the right-hand side exists. The quantity $\delta\mathcal{F}(u,\varphi)$ will be called *first variation of $\mathcal{F}$ at u in the direction φ*.

REMARK 3.4. Without supposing f of class C^1 we do not know whether the derivative of Φ at $\varepsilon = 0$ exists or not, and, in the affirmative case, how we can write it.

We now want to focus our attention on functions $u \in C^1([a,b])$ that are minima of $\mathcal{F}$ in some admissible class of functions $\mathcal{A} \subset C^1([a,b])$. We will see that the conclusions that we will derive really depend on the properties of the admissible class we are working in.

In the following we will suppose that the admissible class is

$$\mathcal{A} := \{w \in C^1([a,b]) \;:\; w(a) = \alpha, w(b) = \beta\}\,,$$

for some fixed values $\alpha, \beta \in \mathbb{R}$. In this case, the variations φ we can consider are the ones that keep fixed the boundary values. For this reason we will consider only functions

$$\varphi \in C_0^1([a,b]) := \{w \in C^1([a,b]) \;:\; w(a) = w(b) = 0\}\,.$$

We know that, if the operator δ is well defined, then we must have

$$\delta\mathcal{F}(u,\varphi) = 0\,,$$

for each $\varphi \in C_0^1([a,b])$. By Lemma 3.2 this can be rephrased as

$$\int_a^b \big[f_p\big(x,u(x),u'(x)\big)\varphi(x) + f_\xi\big(x,u(x),u'(x)\big)\varphi'(x)\big]\,\mathrm{d}x = 0\,, \tag{3.2}$$

for each $\varphi \in C_0^1([a,b])$. This equation is called the *weak Euler-Lagrange equation* of $\mathcal{F}$.

DEFINITION 3.5. A function $u \in C^1([a,b])$ satisfyng (3.2) for each $\varphi \in C_0^1([a,b])$ is called a *weak estremal* of $\mathcal{F}$.

The idea is to obtain a more nice equation that minimizers of $\mathcal{F}$ on $\mathcal{A}$ have to satisfy. For this reason we suppose that the lagrangian f is of class C^2 and that the minimum point u is of class C^2. With these hypothesis in force it is possible, by using integration by parts, to write the weak Euler-Lagrange equation as follows

$$\begin{aligned}
0 &= \int_a^b f_p\big(x,u(x),u'(x)\big)\varphi(x)\,\mathrm{d}x + \int_a^b f_\xi\big(x,u(x),u'(x)\big)\varphi'(x)\,\mathrm{d}x \\
&= \int_a^b f_p\big(x,u(x),u'(x)\big)\varphi(x)\,\mathrm{d}x + \big[f_\xi\big(x,u(x),u'(x)\big)\varphi(x)\big]_a^b \\
&\quad - \int_a^b \Big(\frac{\mathrm{d}}{\mathrm{d}x} f_\xi\big(x,u(x),u'(x)\big)\Big)\varphi(x)\,\mathrm{d}x \\
&= \int_a^b \Big(f_p\big(x,u(x),u'(x)\big) - \frac{\mathrm{d}}{\mathrm{d}x} f_\xi\big(x,u(x),u'(x)\big)\Big)\varphi(x)\,\mathrm{d}x\,,
\end{aligned}$$

where in the last step we have used the fact that $\varphi(a) = \varphi(b) = 0$.

We still need a further step, in oder to obtain a nice equation from the above condition. The following technical result will tell us how.

LEMMA 3.6 (Fundamental lemma of the Calculus of Varitaions). *Suppose we have a function $g \in C^0([a,b])$ such that*

$$\int_a^b g(x)\varphi(x)\,\mathrm{d}x = 0\,,$$

for all $\varphi \in C_0^1([a,b])$. Then $g \equiv 0$ on $[a,b]$.

PROOF. Let assume by the sake of contradiction that there exists a point $\bar{x} \in [a,b]$ such that $g(\bar{x}) \neq 0$. Without loss of generality, we can assume $g(\bar{x}) > 0$. Since g in continuous on $[a,b]$, there exists $\delta > 0$ such that

$$g(x) > \frac{g(\bar{x})}{2} > 0\,,$$

for every $x \in (\bar{x}-\delta, \bar{x}+\delta) \cap [a,b]$. By continuity of g it is also possible to suppose $\bar{x} \in (a,b)$. The idea is to construct a function $\varphi \in C_0^1([a,b])$ such that $\varphi > 0$ on $(\bar{x}-\delta, \bar{x}+\delta)$ and it is zero otherwise. Let us for a moment taking for grant the existence of such a function φ. Then we would have

$$0 = \int_a^b g(x)\varphi(x)\,\mathrm{d}x = \int_{\bar{x}-\delta}^{\bar{x}+\delta} g(x)\varphi(x)\,\mathrm{d}x > \frac{\bar{x}}{2}\int_{\bar{x}-\delta}^{\bar{x}+\delta} \varphi(x)\,\mathrm{d}x > 0\,.$$

Since this is impossible, we conclude that $g \equiv 0$ on $[a,b]$.

Let us now construct such a function φ. define $\varphi : [a,b] \to \mathbb{R}$ as follows:

$$\varphi(x) := \begin{cases} \big(x-(\bar{x}-\delta)\big)^2\big(\bar{x}+\delta-x\big)^2 & x \in (\bar{x}-\delta, \bar{x}+\delta) \\ 0 & \text{otherwise} \end{cases}$$

It is easy to see that the function φ satisfies all the required properties. □

Thus, we have obtained the following result

THEOREM 3.7 (Euler-Lagrange equation - strong form). *Let $f : [a,b]\times\mathbb{R}\times\mathbb{R}$ be a function of class C^2. Suppose that the functional $\mathcal{F} : C^1([a,b]) \to \mathbb{R}$ given by*

$$\mathcal{F}(u) := \int_a^b f\big(x, u(x), u'(x)\big)\,\mathrm{d}x\,,$$

admits a minimum $u \in C^2([a,b]) \cap \mathcal{A}$. Then the following equation holds

$$f_p\big(x, u(x), u'(x)\big) = \frac{\mathrm{d}}{\mathrm{d}x} f_\xi\big(x, u(x), u'(x)\big)\,, \tag{3.3}$$

for each $x \in [a,b]$.

DEFINITION 3.8. Equation (3.3) is called the *(strong) Euler-Lagrange equation* of $\mathcal{F}$. A function $u \in C^2([a,b])$ that satisfies that equation is called a *(strong) extremal* of $\mathcal{F}$.

REMARK 3.9. The above theorem does not assert any kind of existence result. It is just a necessary condition minimizers of $\mathcal{F}$ over $\mathcal{A}$ have to satisfy. Moreover, the fact that a minimum u is of class C^2 is something that we assume *a priori*, and it is in general not garanteed.

REMARK 3.10. We notice that, since f and u are supposed to be of class C^2, we can write the right-hand side of (3.3) as

$$f_{\xi x}\big(x, u(x), u'(x)\big) + f_{\xi p}\big(x, u(x), u'(x)\big)u'(x) + f_{\xi\xi}\big(x, u(x), u'(x)\big)u''(x)\,.$$

3.1.2. Interlude on test functions and local minimality. Before continuing we would like to spend a couple of words about two important questions: the choice of the space of admissible variations, and the concept of local minimizers.

So far, we have used the space $C^1_0([a,b])$ as the space of admissible variations, or *test functions* for minimizers over the class $\mathcal{A}$. This space has been choosen *ad hoc* for the particular situation we are dealing with. Suppose we want to derive a similar necessary condition for minimizers of functionals $\mathcal{F} : C^2([a,b]) \to \mathbb{R}$ of the type

$$\mathcal{F}(u) := \int_a^b f\big(x, u(x), u'(x), u''(x)\big)\,\mathrm{d}x\,,$$

over the class $\mathcal{B} := \{w \in C^2([a,b]) \,:\, w(a) = w\alpha,\, w(b = \beta)\}$. In this case the space of test functions will be $C^2_0([a,b])$. Similarly, if the lagrangian depends on the k-th derivative of u, then the space of test functions will be $C^k_0([a,b])$, and so on. Thus, it is custom to take as the *standard space of test functions* the space

$$C^\infty_C([a,b]) := \{w \in C^\infty([a,b]) \,:\, \operatorname{supp} w \subset\subset [a,b]\}\,,$$

that is, the space of C^∞ functions whose support (*i.e.*, the closure of the set where the function is not zero) is compactly contained in $[a,b]$. The choice of this space must be motivated. First of all we notice that $C^\infty_C([a,b]) \subset C^k_0([a,b])$ for every $k \in \mathbb{N}$. Thus, this space can be used for lagrangians depending on any order of derivatives. Moreover, it turns out that the Fundamental Lemma of the Calculus of Variations holds true even for test functions in this small space. That is, we have

LEMMA 3.11 (Fundamental lemma - second version). *Suppose we have a function $g \in C^0([a,b])$ such that*

$$\int_a^b g(x)\varphi(x)\,\mathrm{d}x = 0\,,$$

for all $\varphi \in C^\infty_C([a,b])$. Then $g \equiv 0$ on $[a,b]$.

REMARK 3.12. Actually, a more general statement holds true: we can take $g \in L^1\big((a,b)\big)$ and obtain the same conclusion!

The idea of the proof is the same of the one for the Lemma 3.6. We just have to construct the suitable test function $\varphi \in C_C^\infty([a,b])$ with the same properties of the one we constructed in Lemma 3.6. This means that, even if we only know that the directional derivative along directions in this smaller space is zero, this is enough to obtain the differential form of the necessary condition $\delta\mathcal{F}(u,\cdot) = 0$. We can think to an analogue in finite dimension: consider a function $f \in C^1(\mathbb{R}^N)$ and, insted of considering variation in all directions, we just consider $\frac{\partial f}{\partial v}(\bar{x}) = 0$ for all unit vectors $v \in \mathbb{R}^N$ having rational coordinates (the analogous of the set $C_C^\infty([a,b])$. Indeed, it turns out that this space is dense in $C_0^k([a,b])$ with respect to the norm $\|\cdot\|_{C^k}$, for all $k \in \mathbb{N}$).

We now discuss the concept of *local minimizers*. As you know, in infinite dimension, not all the norms are equaivalent[1]. This means that the concept of locality depends on the norm we choose in our space. Let now focus on the space $C^1([a,b])$. The natural norm associated to it is the so called C^1-norm $\|\|_{C^1}$ given by

$$\|u\|_{C^1} := \max_{[a,b]} |u| + \max_{[a,b]} |u'| =: \|u\|_{C^0} + \|u'\|_{C^0}\,.$$

But it is possible to consider also the C^0-norm on it (basically we do not care about the derivative!):

$$\|u\|_{C^0} := \max_{[a,b]} |u|\,.$$

Clearly $\|u\|_{C^0} \le \|u\|_{C^1}$, but the two norms are not equivalent, as we can see by considering the funcions $u_n(x) := \frac{1}{n}\sin(nx)$. Hence the norm $\|\cdot\|_{C^1}$ is *stronger* than the norm $\|\cdot\|_{C^0}$. Then we have two notions of local minimality:

DEFINITION 3.13. A function $u \in \mathcal{A}$ such that

$$\mathcal{F}(u) \le \mathcal{F}(v)\,,$$

for all $v \in \mathcal{A}$ with $\|u - v\|_{C^1} < \delta$, for some $\delta > 0$, is called a *weak local minimizer of* $\mathcal{F}$. If equality holds only when $v = u$, we say that is a *strict weak local minimizer of* $\mathcal{F}$.

DEFINITION 3.14. A function $u \in \mathcal{A}$ such that

$$\mathcal{F}(u) \le \mathcal{F}(v)\,,$$

for all $v \in \mathcal{A}$ with $\|u - v\|_{C^0} < \delta$, for some $\delta > 0$, is called a *strong local minimizer of* $\mathcal{F}$. If equality holds only when $v = u$, we say that is a *strict strong local minimizer of* $\mathcal{F}$

Clearly a strong local minimizer is also a weak local minimzier. The opposite is **not** true, as we can see in the following example.

EXAMPLE 3.15 (Sheffer). Consider the functional

$$\mathcal{F}(u) := \int_0^1 \left[\big(u'(x)\big)^2 - \big(u'(x)\big)^4\right],$$

[1]Recall that two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a space X are called *equivalent* if there exist two constants $C_1, C_2 > 0$ such that

$$C_1\|x\|_1 \le \|x\|_2 \le C_2\|x\|_1\,,$$

for all $x \in X$.

defined over the set $C_0^1([0,1])$. Let us prove that $u \equiv 0$ is a strict weak local minimizer, but not a strong one. The idea is the following: the lagrangian $g(\xi) := \xi^2(1-\xi^2)$ as an isolated local minimum at $\xi = 0$ (see Figure 2).

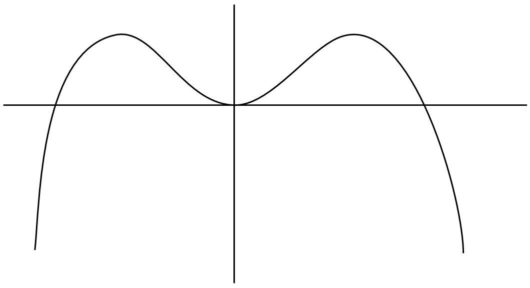

FIGURE 2. The function g.

This is why $u \equiv 0$ is a strict weak local minimizer of $\mathcal{F}$. On the other hand, it is possible to construct a sequence of functions $(u_n)_n$ such that $u_n \to 0$ uniformly in $[0,1]$ and $\mathcal{F}(u_n) \to -\infty$. For, the idea is to make the derivative of the u_n's to explode. So, we define

$$u_n(x) := \frac{1}{k}\sin(2\pi k^2 x)\,.$$

Then it is easy to see that $u_n \to 0$ with respect to the norm $\|\cdot\|_{C^0}$. Moreover $u_n'(x) = 2\pi k\cos(2\pi k^2 x)$. Thus u_n does not converge to 0 in C^1 and

$$\begin{aligned}\mathcal{F}(u_n) &= \int_0^1 \left[4\pi^2 k^2 \cos^2(2\pi k^2 x) - 16\pi^4 k^4 \cos^4(2\pi k^2 x)\right]\,\mathrm{d}x \\ &= 2\pi^2 k^2(1-3k^2\pi) \to -\infty\,.\end{aligned}$$

This proves that $u \equiv 0$ is not a strong local minimizer of $\mathcal{F}$.

Thus, each time we want to consider local minimziers, we have to specify which metric (or topology) we are considering.

REMARK 3.16. Clearly Theorem 3.7 holds true also for weak local minimziers of $\mathcal{F}$, and thus also for strong local minimizers of $\mathcal{F}$.

3.1.3. The Du Bois-Reymon equation - strong form. We now want to derive another first order necessary condition for local minimizers[2] of $\mathcal{F}$. For the moment we give the result without explaing the idea underlying the derivation of this equation, delaying it for when we'll develope the C^1 theory.

THEOREM 3.17 (The Du Bois-Reymon[3] equation - strong form). *Let us consider a function $f : [a,b]\times\mathbb{R}\times\mathbb{R}$ of class C^2. Suppose that the functional $\mathcal{F} : C^1([a,b]) \to \mathbb{R}$ given by*

$$\mathcal{F}(u) := \int_a^b f\big(x, u(x), u'(x)\big)\,\mathrm{d}x\,,$$

admits a weak local minimum $u \in C^2([a,b]) \cap \mathcal{A}$. Then the following equation holds

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[f\big(x,u(x),u'(x)\big) - u'(x) f_\xi\big(x,u(x),u'(x)\big)\right] = f_x\big(x,u(x),u'(x)\big)\,, \tag{3.4}$$

[2]Here we do not specify if we are considering weak or local minimizers, because we want this condition to hold for both.

[3]It is just one person, Paul Du Bois-Reymond!

for each $x \in [a,b]$.

PROOF. By a direct computation we have that

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}x}\big[f\big(x,u(x),u'(x)\big) - u'(x) f_\xi\big(x,u(x),u'(x)\big)\big] \\
&\quad = f_x\big(x,u(x),u'(x)\big) + f_p\big(x,u(x),u'(x)\big)u'(x) + f_\xi\big(x,u(x),u'(x)\big)u''(x) \\
&\qquad - u''(x) f_\xi\big(x,u(x),u'(x)\big) - u'(x)\frac{\mathrm{d}}{\mathrm{d}x} f_\xi\big(x,u(x),u'(x)\big) \\
&\quad = f_x\big(x,u(x),u'(x)\big) + u'(x)\big[f_p\big(x,u(x),u'(x)\big) - \frac{\mathrm{d}}{\mathrm{d}x} f_\xi\big(x,u(x),u'(x)\big)\big]\,.
\end{aligned}$$

□

REMARK 3.18. Equation (3.4) is also called *second form of the Euler-Lagrange equation.*

3.1.4. On the existence of minimizers. In all the previous necessary conditions, we had given for grant the existence of a minimizer. Here we want to show that, even in a very simple case, existence of minimziers mail fail to be true.

The example we are going to show is the so called *Euler's paradox.* Let us consider the function

$$f(\xi) := (1-\xi^2)^2\,.$$

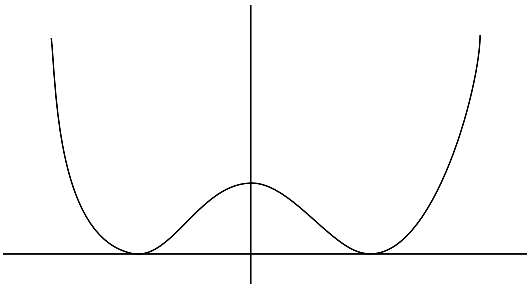

FIGURE 3. The so called double well potential.

Define the functional

$$\mathcal{F}(u) := \int_0^1 f\big(u'(x)\big)\,\mathrm{d}x\,,$$

for all $u \in C^1_0([0,1])$. Then we have $\mathcal{F}(u) \geq 0$, but there is no function such that $\mathcal{F}(u) = 0$. Indeed, such a function u could have only $u' \in +1, -1$, and has to satisfy $u(0) = u(1) = 0$. But this is not compatible with the requirement $u \in C^1$.

If we enlarge our class of admissible functions to the one of piecewise-C^1 functions, that is

$$\begin{aligned}
C^1_{0,pw}([0,1]) := \{u \in C_0([0,1]) \ : \ \exists 0 = x_1 < \cdots < x_N = 1\,,\, u \in C^1((x_i, x_{i+1}))\,,\\
\text{for all } i = 1,\dots,N-1\}\,,
\end{aligned}$$

we can easily seen that each zig-zag function $v \in C^1_{0,pw}([0,1])$ such that $v' \in \{-1,1\}$, where v' exists, is a minimizer of $\mathcal{F}$.

Thus, by considering the function $u(x) := 1 + \left|x - \frac{1}{2}\right|$ and smoothing out the edge, it is possible to prove that:

$$\inf_{C_0^1([0,1])} \mathcal{F} = 0\,.$$

This example shows us that the existence of a minimizer is something that we can not taking from grant!

Finally, let us consider the Euler-Lagrange equation of the above functional. It is

$$4u'(x)[u'(x)^2 - 1]^2 = 0\,.$$

Thus, the only admissible C^1 solution is $u \equiv 0$, that turns out to be a local maximzer of $\mathcal{F}$.

3.2. The first variation - C^1 theory

In the previous chapter we derived two important first order necessary conditions (the Euler-Lagrange equation and the Du Bois-Reymon one) by assuming our minimizer to be of class C^2. The following example shows that, in general, this is an assumption that we cannot make *a priori*.

Example 3.19. Let us consider the functional

$$\mathcal{F}(u) := \int_{-1}^{1} u(x)^2 (2x - u'(x))^2 \,\mathrm{d}x\,,$$

and consider the problem of minimizing it among functions $v \in C^1([-1,1])$ such that $v(-1) = 0$, $v(1) = 1$. It is easy to see that the functional is uniquely minimized by the function

$$u(x) := \begin{cases} 0 & x \in [-1,0] \\ x^2 & x \in [0,1]\,. \end{cases}$$

This function is $C^1([-1,1]) \setminus C^2([-1,1])$.

So, it would be useful to have first order necessary conditions that hold true for (local) minimizers that are only of class $C^1([a,b])$.

3.2.1. The Euler-Lagrange equation.

By looking at what we did in order to derive the Euler-Lagrange equation, we notice that the fundamental step where the additional regularity of the minimizer u (and of the lagrangian f) really matters is when we integrate by parts. Indeed, by just supposing f to be of class C^1 and the minimizer u to be of class C^1, we know that:

$$\int_a^b [\, f_p((x, u(x), u'(x)))\varphi(x) + f_\xi\big(x, u(x), u'(x)\big)\varphi'(x) \,] \,\mathrm{d}x = 0\,, \tag{3.5}$$

for all $\varphi \in C_c^\infty((a,b))$. From this condition we would like to obtain a differential equation. So, let's do it step by step: let us suppose to have two *continuous* functions $g, h : [a,b] \to \mathbb{R}$ such that the following is true

$$\int_a^b [\, g(x)\varphi(x) + h(x)\varphi'(x) \,] \,\mathrm{d}x = 0\,, \tag{3.6}$$

for every $\varphi \in C_c^\infty((a,b))$. We would like to derive some relation between g and h. Since we cannot integrate by part, we have to handle the first term. The technical result that will help us is the following:

LEMMA 3.20 (Du Bois-Reymond lemma). *Let $g : [a,b] \to \mathbb{R}$ be a continuous function such that*

$$\int_a^b g(x)\varphi'(x)\,\mathrm{d}x = 0\,,$$

for all $\varphi \in C_c^\infty((a,b))$. Then there exists $c \in \mathbb{R}$ such that $g(x) = c$ on $[a,b]$.

REMARK 3.21. Notice that if $\varphi \in C_c^\infty((a,b))$, then $\varphi' \in C_c^\infty((a,b))$. Thus, the above result tells us that if $\int_a^b g(x)\varphi(x)\,\mathrm{d}x = 0$ only for functions φ that are derivatives, then I can conclude that g is constant, but I cannot conclude that the constant is 0, as we were able to do in the fundamental lemma of calculus of variations.

PROOF OF LEMMA 3.20. First of all we want to understand (characterize) the subsets of $C_c^\infty((a,b))$ of functions that are derivatives of functions in $C_c^\infty((a,b))$. Let $v \in C_c^\infty((a,b))$; then we have that $v' \in C_c^\infty((a,b))$ and

$$\int_a^b v'(x)\,\mathrm{d}x = v(b) - v(a) = 0\,.$$

we claim that this property characterize derivatives. More clearly, if we have a function $v \in C_c^\infty((a,b))$ such that

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

then there exists $\varphi \in C_c^\infty((a,b))$ such that $v = \varphi'$. Indeed, by defining

$$\varphi(x) := \int_a^x v(t)\,\mathrm{d}t\,,$$

we have that $\varphi \in C_c^\infty((a,b))$ and $\varphi' = v$.

The idea now is to use the fundamental lemma of calculus of variation to prove this result. So, let us take $\varphi \in C_c^\infty((a,b))$. In order to make this function admissible as a test function for our problem, we have to *transform it* into a derivative, *i.e.*, we have to transform it into a function whose integral over $[a,b]$ is zero. The simplest way to do it is by considering the function

$$\widetilde{\varphi}(x) := \varphi(x) - \frac{1}{b-a}\int_a^b \varphi(t)\,\mathrm{d}t\,.$$

This function satisfies all the requisites except one: in general it doesn't vanishes at the ending points! For this reason, we have to modify it a little bit: let us take a function $\omega \in C_0^\infty([a,b])$ such that $\int_a^b \omega(t)\,\mathrm{d}t = 1$. Thus, define the function

$$\widetilde{\varphi}(x) := \varphi(x) - \omega(x)\int_a^b \varphi(t)\,\mathrm{d}t\,.$$

Now, let us check that we did well: $\widetilde{\varphi} \in C_c^\infty((a,b))$, $\int_a^b \widetilde{\varphi}(x)\,\mathrm{d}x = 0$. So, we can use this function as a test function for our problem. Then

$$\begin{aligned} 0 &= \int_a^b g(x)\widetilde{\varphi}(x)\varphi(x)\,\mathrm{d}x \\ &= \int_a^b g(x)\varphi(x)\,\mathrm{d}x - \int_a^b \omega(x)g(x)\left(\int_a^b \varphi(t)\,\mathrm{d}t\right)\mathrm{d}x \\ &= \int_a^b g(x)\varphi(x)\,\mathrm{d}x - \int_a^b \omega(x)g(x)\left(\int_a^b \varphi(t)\,\mathrm{d}x\right)\mathrm{d}t \\ &= \int_a^b \left[g(x) - \int_a^b \omega(t)g(t)\,\mathrm{d}t\right]\varphi(x)\,\mathrm{d}x\,. \end{aligned}$$

Since this holds true for every $\varphi \in C_c^\infty((a,b))$, by using the fundamental lemma of calculus of variation we conclude that

$$g(x) \equiv \int_a^b \omega(t)g(t)\,\mathrm{d}t\,.$$

Notice that this equation makes sense even if g is on both sides! Indeed what it tells us is that g is a constant. $\square$

The previous result will help us in dealing with the expression (3.6).

Corollary 3.22. *Let $g, h : [a,b] \to \mathbb{R}$ be continuous functions. Suppose that*

$$\int_a^b [\,g(x)\varphi(x) + h(x)\varphi'(x)\,]\,\mathrm{d}x = 0\,,$$

holds for each $\varphi \in C_c^\infty((a,b))$. Then the function $h \in C^1([a,b])$ and

$$h'(x) = g(x)\,.$$

Proof. The idea to prove the corollary is to integrate by parts the second term and then use the previous result.

$$0 = \int_a^b [\,g(x)\varphi(x) + h(x)\varphi'(x)\,]\,\mathrm{d}x = \int_a^b \left[h(x) - \left(\int_a^x g(t)\,\mathrm{d}t\right)\right]\varphi'(x)\,\mathrm{d}x\,.$$

By the previous lemma we obtain that there exists a constant $c \in \mathbb{R}$ such that

$$h(x) = c + \int_a^x g(t)\,\mathrm{d}t\,.$$

This proves the desired result. $\square$

By applying the above result to our case we obtain the following necessary condition under the natural assumptions on f and u.

Theorem 3.23. *Let $f : [a,b]\times\mathbb{R}\times\mathbb{R}$ be a function of class C^1. Suppose that the functional $\mathcal{F} : C^1([a,b]) \to \mathbb{R}$ given by*

$$\mathcal{F}(u) := \int_a^b f\big(x, u(x), u'(x)\big)\,\mathrm{d}x\,,$$

admits a minimum $u \in C^1([a,b]) \cap \mathcal{A}$. Then the function

$$x \mapsto f_\xi\big(x, u(x), u'(x)\big)$$

is of class C^1 and there exists a constant $c \in \mathbb{R}$ such that the following equation holds true

$$f_\xi\big(x, u(x), u'(x)\big) = c + \int_a^x f_p\big(t, u(t), u'(t)\big)\, \mathrm{d}t\,,$$

for all $x \in [a, b]$.

Remark 3.24. Usually the above equation is written in the following form:

$$\frac{\mathrm{d}}{\mathrm{d}x} f_\xi\big(x, u(x), u'(x)\big) = f_p\big(x, u(x), u'(x)\big)\,.$$

Nevertheless, we prefer to write it in the integral form in order to remind ourselves that the left-hand side cannot (in general) be expanded by using the chain rule, since we are only assuming f and u to be of class C^1.

3.2.2. The Du Bois-Reymond equation. Since we were so lucky to be able to recover the Euler-Lagrange equation (it a weaker form!) just assuming the natural hypothesis on f and u, we now want to understand if we can be so lucky to recover also the Du Bois-Reymond equation under the same weak assumptions.

The idea of the proof of the theorem is the following: so far we have consider variations of a function u that can be considered *outer* variations. But since we are dealing with functions, we can also take advantage of the fact that we can vary also the independent variable.

First of all, let us take a diffeomorphism $\Psi \in C^1([a, b])$ such that $\Psi(a) = a$ and $\Psi(b) = b$. Then define the function

$$v(x) := u\big(\Psi(x)\big)\,.$$

It holds that

$$\begin{aligned} \mathcal{F}(v) &= \int_a^b f\big(x, v(x), v'(x)\big)\, \mathrm{d}x \\ &= \int_a^b f\big(x, u\big(\Psi(x)\big), u'\big(\Psi(x)\big)\Psi'(x)\big)\, \mathrm{d}x \\ &= \int_a^b f\left(\Phi(y), u(y), \frac{u'(y)}{\Phi'(y)}\right)\Phi'(y)\, \mathrm{d}y\,, \end{aligned}$$

where $\Phi := \Psi^{-1}$. Then, we define the function

$$g(y, q, \eta) := f\left(q, u(y), \frac{u'(y)}{\eta}\right)\eta\,,$$

and the energy:

$$\mathcal{G}(\Psi) := \int_a^b g\big(y, \Psi(y), \Psi'(y)\big)\, \mathrm{d}y\,,$$

for diffeomorphisms $\Psi \in C^1([a, b])$ that does not move the boundary points. Notice that

$$\mathcal{G}(\mathrm{Id}) = \mathcal{F}(u)\,.$$

Let us now suppose that $u \in C^1([a, b])$ is a weak local minimizer for $\mathcal{F}$, and that f is of class C^1. Then Id turns out to be a weak local minimizer for $\mathcal{G}$ and g will be of class C^1. We now want to compute the Euler-Lagrange equation for $\mathcal{G}$ at the identity. To do so, fix a function $\varphi \in C_c^\infty((a, b))$ and let us consider the family of functions:

$$\Psi_\varepsilon(x) := x + \varepsilon\varphi(x)\,.$$

Then there exists $\varepsilon_0 > 0$ such that if $|\varepsilon| < \varepsilon_0$ the function Ψ_e turns out to be a diffeomorphism of $[a,b]$ of class C^1 that does not move the boundary points. Thus, we consider the function

$$\widetilde{\Psi}(\varepsilon) := \mathcal{G}(\Psi_e)\,.$$

We know that this function has a local minimum for $\varepsilon = 0$. Thus the condition $\widetilde{\Psi}'(0) = 0$ leads to the (weak) Euler-Lagrange equation

$$\frac{\mathrm{d}}{\mathrm{d}y} g_\eta\big(y, y, 1\big) = g_q\big(y, y, 1\big)\,.$$

That is

$$\frac{\mathrm{d}}{\mathrm{d}x}\big[f_\xi\big(x, u(x), u'(x)\big)u'(x) - f\big(x, u(x), u'(x)\big)\big] = f_x\big(x, u(x), u'(x)\big)\,.$$

We have just obtained the following result

THEOREM 3.25. *Let $f : [a,b]\times\mathbb{R}\times\mathbb{R}$ be a function of class C^1. Suppose that the functional $\mathcal{F} : C^1([a,b]) \to \mathbb{R}$ given by*

$$\mathcal{F}(u) := \int_a^b f\big(x, u(x), u'(x)\big)\,\mathrm{d}x\,,$$

admits a minimum $u \in C^1([a,b]) \cap \mathcal{A}$.
Then there exists a constant $c \in \mathbb{R}$ such that the following equation holds true

$$f_\xi\big(x, u(x), u'(x)\big)u'(x) = c + f\big(x, u(x), u'(x)\big) + \int_a^x f_x\big(t, u(t), u'(t)\big)\,\mathrm{d}t\,, \tag{3.7}$$

for all $x \in [a,b]$.

REMARK 3.26. Notice that the Euler-Lagrange equation and the Du Bois-Reymond one are, in general, different equations, since there are different derivatives of f present in each one.

3.2.3. A regularity result. Now we would like to ask ourselves the following question: let us suppose that we have a (local) minimizer $u \in C^{(}[a,b])$ of $\mathcal{F}$. Is is possible to deduce from the weak Euler-Lagrange equation, without explicitly solving it, that u actually has more regularity? The following result gives us an answer.

THEOREM 3.27. *Let $f : [a,b]\times\mathbb{R}\times\mathbb{R}$ be a lagrangian of class C^2. Let $u \in C^1([a,b])$ be a solution of the weak Euler-Lagrange equation (3.5). Suppose that*

$$f_{\xi\xi}\big(x, u(x), u'(x)\big) \neq 0\,,$$

for all $x \in [a,b]$. Then $u \in C^2([a,b])$.

PROOF. We know that there exists $c \in \mathbb{R}$ such that the following equation holds true

$$f_\xi\big(x, u(x), u'(x)\big) = g(x) := c + \int_a^x f_p\big(t, u(t), u'(t)\big)\,\mathrm{d}t\,,$$

for all $x \in [a,b]$. The function g turns out to be of class C^1. Define the function

$$G(x,\xi) := f_\xi\big(x, u(x), \xi\big) - g(x)\,,$$

which is of class C^1. We know that

$$G(x, u'(x)) = 0\,, \qquad \text{for all } x \in [a,b]\,.$$

Moreover

$$G_\xi(x, u'(x)) \neq 0\,, \qquad \text{for all } x \in [a,b]\,.$$

By applying the implicit function theorem we obtain that

$$u'(x) = h(x)\,,$$

for some function h of class C^1. Thus, $u \in C^2([a,b])$. $\square$

REMARK 3.28. It is possible to extend the above theorem as follows: let us suppose that the lagrangian f is of class C^k and still satisfies the non degeneracy condition. Then, any C^1 solution of the weak Euler-Lagrange equation turns out to be of class C^k.

The following example shows that the non degeneracy condition is really needed in order to obtain such a regularity result.

EXAMPLE 3.29. Let $f \in C^2(\mathbb{R})$ be a convex function such that $f(x) = x$ in $[1,2]$ and with f' injective in $\mathbb{R} \setminus [1,2]$. Then the minimizers of the functional

$$\int_0^1 f\big(u'(x)\big)\,\mathrm{d}x\,,$$

with $u(0) = 0$, $u(1) = 1$ can present singularities.

Moreover, it is also fundamental that we have non degeneracy at *all* the points. Indeed, as the following example will show, loosing it in just only point, can of singularities.

EXAMPLE 3.30. Let us define the function

$$g(x,\xi) := \big(\sqrt{2}\xi - \sqrt{x^2+\xi^2}\big)^3\,,$$

and the lagrangian

$$f(x,\xi) := \int_0^\xi g(x,\eta)\,\mathrm{d}\eta\,.$$

Then

$$f_{\xi\xi}(x,\xi) = g_\xi(x,\xi) = 3\big(\sqrt{2}\xi - \sqrt{x^2+\xi^2}\big)^2\left(\sqrt{2} - \frac{\xi}{\sqrt{x^2+\xi^2}}\right).$$

So $f_{\xi\xi}(x,\xi) = 0 \Leftrightarrow \xi = |x|$. Take the function $u(x) := \frac{1}{2}x|x|$. Then

- u is a solution of the weak Euler-Lagrange equation,
- $f_{\xi\xi}\big(x,u'(x)\big) \equiv 0$ on $[a,b]$,
- $u \in C^1([-1,1]) \setminus C^2([-1,1])$.

REMARK 3.31. The condition we required in the preceding theorem is related to the convexity of the function

$$\xi \mapsto f\big(x,u(x),\xi\big)$$

at $\xi = u'(x)$.

3.3. Lagrangians of some special form

We now want to compute explicitly the Euler-Lagrange equation and the Du Bois-Reymond's one for some interesting examples. In all the following examples we will suppose to be in the regular case, *i.e.*, a lagrangian of class C^2 and a (local) minimizer of class C^2.

3.3.1. The case $f = f(x,p)$. In this case the (strong) Euler-Lagrange equation reduces to

$$f_p\big(x, u(x)\big) = 0\,, \qquad \text{on } [a,b]\,.$$

Thus, we are led to consider, for each $x \in [a,b]$, critical points of the function

$$p \mapsto f(x,p)\,.$$

If there exists a point $x \in [a,b]$ for which this function does not admits critical points, then we do not have existence of extremals of class C^2.

Moreover, in order to ensure the existence of a minimizer of $\mathcal{F}$ of class C^2, we need to have, for each $x \in [a,b]$, existence of a minimum point of the function

$$p \to f(x,p)\,.$$

Otherwise, it is easy to see that a minimizer does not exist.

Notice that bad behaviors can occur: suppose that, for each $x \in [a,b]$, there exists a unique minimizer p_x of the function $p \to f(x,p)$. So, we are tempted to define $u(x) := p_x$. The problem is that this function may lack to have the desired regularity! An example of this situation is given by the following function:

$$f(x,p) := \big((p+1)^2 - 1\big)g(-x)g(p) + \big((p-1)^2 - 1\big)g(x)g(-p)\,,$$

where

$$g(y) := \begin{cases} e^{-\frac{1}{y}} & y > 0\,, \\ 0 & y \geq 0\,. \end{cases}$$

Finally, notice that in this case the Du Bois-Reymond equation is the same as the Euler-Lagrange one.

3.3.2. The case $f = f(\xi)$. In this case the Euler-Lagrange equation is simply

$$\frac{\mathrm{d}}{\mathrm{d}x} f'\big(u'(x)\big) = 0\,, \qquad \text{on } [a,b]\,.$$

A particular C^1 solution of the above equation that satisfies the boundary conditions is given by

$$u(x) := \frac{\beta - \alpha}{b - a} x + \alpha\,.$$

The convex case. If f' is an injective function, then the above one is also the unique solution satisfying the boundary conditions, and thus it is the unique strong extremal of $\mathcal{F}$. In particular, if f is convex, then the above function is a minimizer of $\mathcal{F}$, and if we have strict convexity, it is also unique.

An example of such a situation is given by the curves of minimal length. In this case the lagrangian is given by

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

Thus, we have that the line segment is the unique solution of the problem.

The non convex case. If f is not convex, then solution to the Euler-Lagrange equation may fail to be a minimizer. For example, let us consider the lagrangian

$$f(\xi) := e^{-\xi^2}\,,$$

and suppose $\alpha = \beta = 0$ (boundary conditions). In this case the function $u \equiv 0$ is an extremal, but it is not a minimizer. Indeed, it is a maximizer. Moreover one can see that the minimum problem does not admit a solution.

3.3.3. The case $f = f(x,\xi)$. In this case the Euler-Lagrange equation is

$$\frac{\mathrm{d}}{\mathrm{d}x} f_\xi\big(x, u'(x)\big) = 0\,, \qquad \text{on } [a,b].$$

In this case there is no a simple solution as there was in the previous one. But let us suppose we are able to solve the above equation with respect to u', that is, we can find a function $g : [a,b] \times \mathbb{R} \to \mathbb{R}$ such that

$$u'(x) := g(x,c)\,,$$

is a solution of

$$f_\xi\big(x, u'(x)\big) = c\,.$$

Thus, the solutions of the Euler-Lagrange equation are given by

$$u(x) := \alpha + \int_a^x g(t,c)\ \mathrm{d}t\,.$$

REMARK 3.32. In this case the minimum problem

$$\min\Big\{\mathcal{F}(u) := \int_a^b f\big(x, u'(x)\big)\ \mathrm{d}x \ :\ u(a) = \alpha,\ u(b) = \beta\Big\}\,,$$

can be translated into the following:

$$\min\Big\{\mathcal{G}(v) := \int_a^b f\big(x, v(x)\big)\ \mathrm{d}x \ :\ \int_a^b v(x)\ \mathrm{d}x = \beta - \alpha\Big\}\,.$$

We will see how this new formulation will give us a geometric interpretation of the Euler-Lagrange equation for lagrangians of the type $f = f(x,\xi)$.

3.3.4. The case $f = f(p,\xi)$. In this case the Euler-Lagrange equation is

$$\frac{\mathrm{d}}{\mathrm{d}x} f_\xi\big(u(x), u'(x)\big) = f_p\big(u(x), u'(x)\big)\,, \qquad \text{on } [a,b]\,.$$

This is much more difficult to solve. However, from (3.4) we have that the quantity

$$f\big(u(x), u'(x)\big) - u'(x) f_\xi\big(u(x), u'(x)\big)\,,$$

is conserved along each solution of (3.3).

We want to give a physical interpretation of the above equations. For, let just for the moment, call t the independent variable (instead of x), and consider the lagrangian

$$f(p,\xi) := \frac{m}{2}|\xi|^2 - V(p)\,,$$

where $m > 0$ and $V \in C^1(\mathbb{R})$. The functional

$$\mathcal{F}(u) := \int_{t_1}^{t_2} \frac{m}{2}|u'(t)|^2 - V\big(u(t)\big)\ \mathrm{d}t\,,$$

is called the *action integral* of the motion given by $t \mapsto u(t)$. The quantity $-V'$ is called *potential energy*, and $\frac{m}{2}|u'(t)|^2$ the *kinetic energy*. The Euler-Lagrange equation becomes

$$mu''(x) = -V'\big(u(x)\big)\,,$$

and is also known as *Newton equation*. Moreover the quantity

$$E(p,\xi) := \frac{m}{2}|\xi|^2 + V(p)\,,$$

called the *total energy*, is conserved along any solution of the Euler-Lagrange equation. This fact is also knows as *conservation of the mechanical energy*, and it is due to the fact that the lagrangian does not depend explicitly on the (time) variable t.

REMARK 3.33. As we will see, the Euler-Lagrange equation and the Du Bois-Reymond one can be generalized to the case of functions $u : [a, b] \to \mathbb{R}^N$. Thus, using the above argument, we can recover Newton's equation and the conservation of the total mechanical energy for particles moving in $\mathbb{R}^N$.

Moreover we will show that, using the Calculus of Variations, it is possible to find a correspondence between symmetry of the physical system (invariances) and conserved quantities.

REMARK 3.34. In deriving the Du Bois-Reymond equation (3.4) we have used the Euler-Lagrange one. So, it is natural to ask whether the two equations are equivalent or not. In general the answer is no, that is, the are solutions of the Du Bois-Reymond equations that are not solutions of the Euler-Lagrange one. But there is one interesting case where the are actually equivalent. Suppose the lagrangian f is of the form $f = f(p, \xi)$ and suppose we are looking for non constant solutions u. Then

$$u \text{ solves } (3.3) \ \Leftrightarrow u \text{ solves } (3.4)\ .$$

Indeed, by using the special form of f (independence with respect to x), we have that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}&\left[f\big(x, u(x), u'(x)\big) - u'(x) f_\xi\big(x, u(x), u'(x)\big)\right] \\ &= u'(x)\Big[f_p\big(x, u(x), u'(x)\big) - \frac{\mathrm{d}}{\mathrm{d}x} f_\xi\big(x, u(x), u'(x)\big)\Big]\,.\end{aligned}$$

3.4. Solution of the brachistochrone problem

We now want to find explicitly the extremals of the brachistochrone functional. To do so, let us recall that we are dealing with the followgin lagrangian:

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

The Euler-Lagrange equation is

$$u''(x)u(x) = -\frac{1}{2}\big(1 + \big(u'(x)\big)^2\big)\,,$$

from which we deduce $u'' < 0$ and $u > 0$. Let us now compute the Du Bois-Reymond equation. We have

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

for some $c \in \mathbb{R}$. Thus[4]:

$$u(x)\big(1 + \big(u'(x)\big)^2\big) = \frac{1}{c^2}\,. \tag{3.8}$$

The above equation is of the form $u = g(u')$, with $g(\xi) := \frac{d}{(1+\xi^2)}$, where $d := \frac{1}{c^2}$. These kind of differential equations can be treated in a standard way. The idea is to write the solution u not as a function of x, but with respect to the variable u'. Let us be more explicit: first of all we notice that $x \mapsto u'(x)$ is invertible, since $u'' < 0$. So, if we let $[c, d] := u'([a, b])$ to be the range of the function u' we know that there exists a function[5] $x : [c, d] \to [a, b]$ such that

$$u'\big(x(\xi)\big) = \xi\,,$$

[4] Notice that the Euler-Lagrange equation is of second order, while the Du Bois-Reymond's one is of first order.

[5] The really bad notation will be useful in a moment!

that is, the function x is the inverse of u'. So, it is possible to represent the graph of a solution u of (3.8) of the form

$$\operatorname{graph}(u) = \{(x(\xi), y(\xi)) \,:\, y(\xi) = u\big(x(\xi)\big) = g(\xi),\, \xi \in [c,d]\}\,.$$

In order to find an explicit form of the above representation, we proceed as follows: since

$$u\big(x(\xi)\big) = g(\xi) \;\Rightarrow\; g'(\xi) = u'\big(x(\xi)\big)x'(\xi) = \xi x'(\xi)\,,$$

and in our case $g(\xi) := \frac{d}{1+\xi^2}$, we arrive at

$$\frac{-2\xi d}{(1+\xi^2)} = \xi x'(\xi)\,,$$

from which we obtain

$$x(\xi) = -2d \int \frac{\mathrm{d}\xi}{(1+\xi^2)^2}\,.$$

To solve explicitly the above integral, we make use of the change of variable $\xi = \tan\tau$. Since $\mathrm{d}\xi = (1+\tan^2\tau)\,\mathrm{d}\tau$, we get

$$\begin{aligned} x(p) &= -2d \int \frac{\mathrm{d}\tau}{1+\tan^2\tau} \\ &= -2d \int \cos^2\tau \,\mathrm{d}\tau \quad \big[\text{change of variable } \tau = \frac{s}{2}\big] \\ &= -\frac{d}{2}\int (1+\cos s)\,\mathrm{d}s = -\frac{d}{2}(s+\sin s) + k =: x(s)\,. \end{aligned}$$

We can easily forget about the constant k because it is just a translation. Thus, recalling that $\xi = \tan\frac{s}{2}$, we get

$$y(\xi) = g(\xi) = d\cos^2\frac{s}{2} = \frac{d}{2}(1+\cos s)\,.$$

We thus arrive to the parametric representation

$$\begin{cases} x(s) = -\frac{d}{2}(s+\sin s)\,, \\ y(s) = \frac{d}{2}(1+\cos s)\,. \end{cases}$$

Finally, by using the change of variable $t = s+\pi$ and by letting $R := -\frac{d}{2}$, we get (up to a translation in x)

$$\begin{cases} x(t) = R(t-\sin t)\,, \\ y(t) = -R(1-\cos t)\,. \end{cases}$$

The object we obtain is called *catenary* (see Figure 4).

We now have to care about the boundary conditions. We do not have any loss of generality if we consider the point P to be the origin, and the point Q to be of the form (b,β), with $b>0$ and $\beta<0$. In this case, in order to pass by the point P, we just have to take $t_0 = 0$. We thus wonder if it is possible to make the catenary to pass also for a generic point Q as above. What we aimed at proving is that there exists just one couple (R,t_1) such that

$$\begin{cases} x(t_1) = R(t_1 - \sin t_1) = b\,, \\ y(t_1) = -R(1-\cos t_1) = \beta\,. \end{cases}$$

To prove this fact, let us consider the function

$$h(t) := \frac{y(t)}{x(t)} = \frac{\cos t - 1}{t - \sin t}\,.$$

We notice that

$$\lim_{t\to 0} h(t) = -\infty\,, \qquad h \le 0\,, \qquad h(2\pi) = 0\,.$$

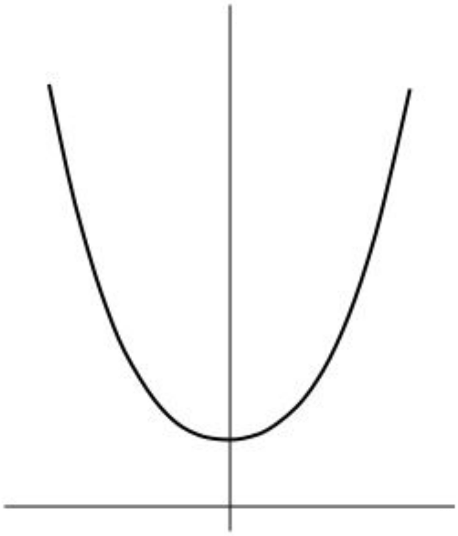

FIGURE 4. The graph of a catenary.

Since the function h is continuous, we conclude that for each $m < 0$ there exists $t_m > 0$ such that $h(t_m) = m$. We now want to prove that the function h is injective. For, we now consider

$$h'(t) = \frac{2(1-\cos t) - t\sin t}{(t-\sin t)^2}\,.$$

We claim that $h'(t) > 0$ for $t \in (0, 2\pi)$. To prove so, let us consider the function $n(t) := 2(1-\cos t) - t\sin t$. We have that

$$n'(t) = \sin t - t\cos t\,.$$

We notice that there exists just one[6] $\bar{t} \in (0, 2\pi)$ such that $n'(\bar{t}) = 0$. In particular, since $n(0) = n(2\pi) = 0$ and $n\left(\frac{\pi}{2}\right) > 0$, we can conclude that $n(t) > 0$ for $t \in (0, 2\pi)$. This tells us that, for a fixed R and for $t \in (0, 2\pi)$, the catenary meets each line passing through the origin in exactly one point. So, let $\bar{t} \in (0, 2\pi)$ be such that $\frac{y(\bar{t})}{x(\bar{t})} = -\frac{\beta}{b}$. Then we take the unique $R > 0$ such that $R(\bar{t} - \sin\bar{t}) = b$.

Thus, we have obtained that there exists exactly one critical point satifying the boundary condition. Bu so far we have no argument to conclude that the catenary turns out to be a (local) minimizer. Next section will provide us with a powerful method to conclude that the catenary is actually the unique global minimizer of the brachistochrone problem.

3.4.1. The method of coordinate transform. This method can be very useful to transform a difficult problem into a simper one. Let $u \in C^1([a,b])$ be a weak extremal for the lagrangian f such that $u' \neq 0$ on (a,b). Let us take a function $\Phi \in C^2(\mathbb{R})$ and suppose that there exists $\Phi^{-1} : u([a,b]) \to \mathbb{R}$ and it is of class C^1. Define implicitly a function

[6]Such a $\bar{t}$ will belong to $\in (\pi, \frac{3}{2}\pi)$. Indeed for $t \in (0, \frac{\pi}{2})$ we have that $\tan t > t$, while for $t \in (\frac{\pi}{2}) \cup (\frac{3}{2}\pi, 2\pi)$ we have $\tan t < 0$.

$v \in C^1([a,b])$ via
$$u(x) := \Phi\big(v(x)\big)\,.$$
Then
$$\mathcal{F}(u) + \int_a^b f\big(x, u(x), u'(x)\big)\,\mathrm{d}x = \int_a^b f\big(x, \Phi\big(v(x)\big), \Phi'\big(v(x)\big)v'(x)\big)\,\mathrm{d}x\,.$$
Introduce the lagrangian
$$g(x,p,\xi) := f\big(x, \Phi(p), \Phi'(p)\xi\big)\,.$$
Then g is of class C^1 and we can write its weak Euler-Lagrange equation as
$$g_\xi\big(x, v(x), v'(x)\big) = c + \int_a^x g_p\big(t, v(t), v'(t)\big)\,\mathrm{d}t\,,$$
that writes as
$$\begin{aligned}\Phi'\big(v(x)\big) f_\xi\big(x, u(x), u'(x)\big) = c + \int_a^b [\,&\Phi'\big(v(t)\big) f_p\big(t, u(t), u'(t)\big)\\ &+ f_\xi\big(t, u(t), u'(t)\big)\big(\Phi'\big(v(t)\big)\big)'\,]\,\mathrm{d}t\,.\end{aligned}$$
Now
$$\begin{aligned}&\int_a^x f_\xi\big(t, u(t), u'(t)\big)\big(\Phi'\big(v(t)\big)\big)'\,\mathrm{d}t =\\ &\qquad\Big[f_\xi\big(t, u(t), u'(t)\big)\Phi'\big(v(t)\big)\Big]_a^x - \int_a^x \Phi'\big(v(t)\big)\frac{\mathrm{d}}{\mathrm{d}t} f_\xi\big(t, u(t), u'(t)\big)\,\mathrm{d}t\,.\end{aligned}$$
Thus the weak Euler-Lagrange equation is
$$0 = \int_a^x \Phi'\big(v(t)\big)\Big[\frac{\mathrm{d}}{\mathrm{d}t} f_\xi\big(t, u(t), u'(t)\big) - f_p\big(t, u(t), u'(t)\big)\Big]\,\mathrm{d}t\,.$$
Thus, if u is a weak extremal for f, then v is also a weak extremal for g. The viceversa is true provided $\Phi' \neq 0$ on $[a,b]$.

3.4.2. Minimality of the cycloid.

Let us consider the function[7]
$$v := \frac{u^2}{2}\,.$$
Then
$$\mathcal{F}(u) = \int_0^b \frac{\sqrt{1 + \big(u'(x)\big)^2}}{\sqrt{u(x)}}\,\mathrm{d}x = \int_0^b \sqrt{\frac{1}{v^2(x)} + \big(v'(x)\big)^2}\,\mathrm{d}x\,.$$
The new lagrangian is
$$g(p,\xi) := \sqrt{\frac{1}{p^2} + \xi^2}\,,$$
that turns out to be strictly convex. The previous section ensue that that, since $u > 0$ everywhere, u is a critical point for the lagrangian f if and only if v is a critical point for the lagrangian g. The advantage of the latter one is that it is convex, and convexity is enough to conclude that a critical point is a minimizer, as it will be shown in the theorem below. In particular, this implies that the catenary is the unique solution of the brachistochrone problem.

[7]In order to be completely precise, let us take translate the catenary in the vertical direction, in order to have $u > 0$.

THEOREM 3.35. *Let $f : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a lagrangian of class C^1. Assume that for each $x \in [a,b]$ the function*

$$(p,\xi) \mapsto f(x,p,\xi)$$

is convex. Then any solution of the weak Euler-Lagrange equation is a minimizer. Moreover, if the above function is strictly convex, we also have uniqueness.

PROOF. Let $v \in \mathcal{A}$. By convexity[8] we have that

$$\begin{aligned} f(x,v(x),v'(x)) \geq f\big(x,u(x),u'(x)\big) &+ f_p\big(x,u(x),u'(x)\big)(v(x)-u(x)) \\ &+ f_\xi\big(x,u(x),u'(x)\big)(v'(x)-u'(x))\,, \end{aligned}$$

holds for every $x \in [a,b]$. By integrating we obtain

$$\mathcal{F}(v) \geq \mathcal{F}(u) + \int_a^b [f_p\big(x,u(x),u'(x)\big)(v(x)-u(x)) + f_\xi\big(x,u(x),u'(x)\big)(v'(x)-u'(x))]\,\mathrm{d}x\,.$$

The result follows since $v-u$ is an admissible test function[9], and thus the integral on the right-hand side vanishes.

To prove uniqueness in the case of strict convexity, let as assume that u_1 and u_2 are two (different) minimizers, and let us consider the function $v := \frac{u_1+u_2}{2}$. By strict convexity we have

$$f\big(x,v(x),v'(x)\big) < \frac{1}{2} f\big(x,u_1(x),u_1'(x)\big) + \frac{1}{2} f\big(x,u_2(x),u_2'(x)\big)\,.$$

Integrating we obtain that

$$\mathcal{F}(v) < \frac{1}{2}\mathcal{F}(u_1) + \frac{1}{2}\mathcal{F}(u_2) = \min_{\mathcal{A}} \mathcal{F}\,,$$

Clearly this is impossible. □

3.4.2.1. *Tautochrone property of the cycloid.* We would like to prove a formidable property of the cycloid, namely that the time needed to go from any point to the lower one is independent of the starting point.

First of all notice that the lower point of the cycloid is reached for $t = \pi$. If we start from a generic $\bar{t} \in [0,\pi)$, the time needed to reah the lower point is

$$\int_{\bar{t}}^{\pi} \frac{\sqrt{1-\cos t}}{\sqrt{\cos^2 \bar{t} - \cos^2 t}}\,\mathrm{d}t\,.$$

By using the fact that $\sqrt{1-\cos t} = \sqrt{2}\sin\frac{t}{2}$ and $\cos t = 2\cos^2\frac{t}{2} - 1$, we can rewrite the above integral as

$$\int_{\bar{t}}^{\pi} \frac{\sin\frac{t}{2}}{\cos^2\frac{\bar{t}}{2} - \cos^2\frac{t}{2}}\,\mathrm{d}t\,.$$

By using the change of variable $z = \frac{\cos\frac{t}{2}}{\cos\frac{\bar{t}}{2}}$, we get that it is equal to

$$\int_0^1 \frac{2}{\sqrt{1-z^2}}\,\mathrm{d}z = \pi\,.$$

Since this is independent of $\bar{t}$, we have proved the tautochrone property of the cycloid.

[8]See Lemma 11.9

[9]Actually it belongs to $C^1_0([a,b])$. But we know that we can use this space as a space of test functions!

3.5. Problems with free ending points

In this section we want to derive the analogous of the Euler-Lagrange equation and of the Du Bois-Reymond one when we have to minimize a functional

$$\mathcal{F}(u) := \int_a^b f\big(x, u(x), u'(x)\big)\,\mathrm{d}x\,,$$

over the class

$$\mathcal{A} := C^1([a,b])\,,$$

that is when we have no boundary condition. For, we reason as we did for the constrained case. Let $u \in C^1([a,b])$ be a (local) minimizer, and let us take a function $\varphi \in C^\infty([a,b])$. Notice that for each $\varepsilon \in \mathbb{R}$, the function $u + \varepsilon\varphi \in \mathcal{A}$. Hence we can consider the function

$$\Phi(\varepsilon) := \mathcal{F}(u + \varepsilon\varphi)\,,$$

and we know that $\Phi'(0) = 0$, that is

$$0 = \int_a^b [\, f_p\big(x, u(x), u'(x)\big)\varphi(x) + f_\xi\big(x, u(x), u'(x)\big)\varphi'(x)\,]\,\mathrm{d}x\,.$$

To understand what is going on, let us suppose for the moment that $f \in C^2$ and $u \in C^2$. So it is possible to integrate by parts to obtain

$$\begin{aligned}0 = \int_a^b \Big[\, f_p\big(x, u(x), u'(x)\big) - \frac{\mathrm{d}}{\mathrm{d}x} f_\xi\big(x, u(x), u'(x)\big)\,\Big]\varphi(x)\,\mathrm{d}x \\ + f_\xi\big(b, u(b), u'(b)\big)\varphi(b) - f_\xi\big(a, u(a), u'(a)\big)\varphi(a)\,.\end{aligned}$$

The above equality holds for all $\varphi \in C^\infty([a,b])$. Since $C_c^\infty((a,b)) \subset C^\infty([a,b])$, it particular it holds for every $\varphi \in C_c^\infty((a,b))$. Thus, by using what we already know, we conclude that

$$f_p\big(x, u(x), u'(x)\big) = \frac{\mathrm{d}}{\mathrm{d}x} f_\xi\big(x, u(x), u'(x)\big)\,, \qquad \text{on } [a,b]\,.$$

So we are left with

$$f_\xi\big(b, u(b), u'(b)\big)\varphi(b) - f_\xi\big(a, u(a), u'(a)\big)\varphi(a) = 0\,,$$

for all $\varphi \in C^\infty([a,b])$. It is easy to guess what this will implies.

Lemma 3.36. *Let $g : [a,b] \to \mathbb{R}$ be a continuous function such that*

$$g(b)\varphi(b) - g(a)\varphi(a) = 0\,,$$

for all $\varphi \in C^\infty([a,b])$. Then $g(b) = g(a) = 0$.

Proof. Assume by the sake of contradiction that $g(a) \neq 0$. Without loss of generality we can assume $g(a) > 0$. Let us consider the function

$$\varphi(x) := -\frac{g(a)}{b-a}x + g(a)\frac{b}{b-a}\,.$$

So $\varphi \in C^\infty([a,b])$. Thus we obtain

$$0 = \varphi(a)g(a) = g(a) > 0\,.$$

This is absurd. The same argument applies for the point b. □

So, we have the following result

THEOREM 3.37. *Let $f \in C^2 : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a lagrangian of class C^2. Let $u \in C^2([a,b])$ be a (local) minimizer for the functional $\mathcal{F}$ over the class $C^{(}[a,b])$. Then the following hold:*

$$f_p\big(x, u(x), u'(x)\big) = \frac{\mathrm{d}}{\mathrm{d}x} f_\xi\big(x, u(x), u'(x)\big)\,, \qquad \text{on } [a,b]\,,$$

and

$$f_\xi\big(a, u(a), u'(a)\big) = f_\xi\big(b, u(b), u'(b)\big) = 0\,. \tag{3.9}$$

DEFINITION 3.38. Conditions (3.9) are called *natural conditions*.

Finally, we would like to obtain the same result with the natural C^1 assumptions on f and u. First of all we notice that the Du Bois-Reymond lemma (see Lemma 3.20) applies also in this case, since $C_c^\infty((a,b)) \subset C^1([a,b])$. In particular we obtain that the function

$$x \mapsto f_\xi\big(x, u(x), u'(x)\big)$$

is of class $C^1([a,b])$. This allows us to integrate by parts as above and to obtain the same conclusion.

THEOREM 3.39. *Let $f \in C^1 : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a lagrangian of class C^2. Let $u \in C^1([a,b])$ be a (local) minimizer for the functional $\mathcal{F}$ over the class $C^{(}[a,b])$. Then the following hold:*

$$f_\xi\big(x, u(x), u'(x)\big) = c + \int_a^x f_p\big(t, u(t), u'(t)\big)\,\mathrm{d}t\,, \qquad \text{on } [a,b]\,,$$

for some $c \in \mathbb{R}$, and

$$f_\xi\big(a, u(a), u'(a)\big) = f_\xi\big(b, u(b), u'(b)\big) = 0\,.$$

REMARK 3.40. Clearly, if the class of admissible functions is

$$\mathcal{A} := \{u \in C^1([a,b]) \,:\, u(a) = \alpha\}\,,$$

the one obtains the natural boundary condition only at b.

3.6. Isoperimetric problems

We now want to investigate which kind of first order necessary conditions we can derive for local minimizers of the following problem: minimize

$$\mathcal{F}(u) := \int_a^b f\big(x, u(x), u'(x)\big)\,\mathrm{d}x\,,$$

over the class

$$\mathcal{A} := \{u \in C^1([a,b]) \,:\, u(a) = \alpha,\ u(b) = \beta,\ \mathcal{G}(u) = c\}\,,$$

where $c \in \mathbb{R}$ and

$$\mathcal{G}(u) := \int_a^b g\big(x, u(x), u'(x)\big)\,\mathrm{d}x\,.$$

This problem is a particular case of the following most general constrained minimization problem: minimize $\mathcal{F}$ over the class over the class

$$\mathcal{A} := \{u \in C^1([a,b]) \,:\, u(a) = \alpha,\ u(b) = \beta,\ G\big(x, u(x), u'(x)\big) = c \text{ on } [a,b]\}\,,$$

where $c \in \mathbb{R}$ and $G : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a C^1 function. This general case is most difficult to handle[10]

[10]The real difficulties arise when the function G depends on u'.

3.6.1. The Lagrange multiplier rule on $\mathbb{R}^N$. We now want to recall some results about constrained minimization problems on $\mathbb{R}^N$. Let us suppose we have two functions $F, G : \mathbb{R}^N \to \mathbb{R}$ of class C^1, and consider the problem

$$\min_{\mathcal{A}} F\,,$$

where

$$\mathcal{A} := \{x \in \mathbb{R}^N \,:\, G(x) = 0\} =: \{G = 0\}\,.$$

We would like to obtain some first order necessary conditions for local minimality. In general these conditions do not exist. Let us think to the case $F(x,y) := |(x,y)|^2$ and $G(x,y) := xy$. The minimum point is the origin, but no first order necessary conditions can be derived for that point. The problem relies on the fact that the minimum point is a point where ∇G vanishes, and thus we can not obtain a good description of $\{G = 0\}$ near that point.

So, let us suppose that $\bar{x} \in \mathcal{A}$ is a point of local minimum for F, and assume that $\nabla G(\bar{x}) \neq 0$. This means that, locally around $\bar{x}$, the set $\{G = 0\}$ is a submanifold of $\mathbb{R}^N$ of class C^1. Since $\bar{x}$ is a local minimum, if we consider a sequence $(x_n)_n \in \mathcal{A}$ converging to $\bar{x}$, we must have

$$F(\bar{x}) \leq F(x_n)\,,$$

for n large enough. Let us suppose that

$$\frac{x_n - \bar{x}}{|x_n - \bar{x}|} \to v\,,$$

Then, clearly, v belongs to the tangent space of the submanifold $\{G = 0\}$ at the point $\bar{x}$, that we will denote by $\mathrm{Tan}_{\{G=0\}}(\bar{x})$,. Moreover, each vector $v \in \mathrm{Tan}_{\{G=0\}}(\bar{x})$ can be obtained in that way (up to multiplying by its norm!). So, we have that

$$\nabla F(\bar{x}) \cdot v = 0\,,$$

for all $v \in \mathrm{Tan}_{\{G=0\}}(\bar{x})$. This means that $\nabla F(\bar{x})$ is orthogonal to $\mathrm{Tan}_{\{G=0\}}(\bar{x})$. By recalling[11] that ∇G is orthogonal to each level set of G, we infer that

THEOREM 3.41 (Lagrange multiplier Theorem). *Let $F, G : \mathbb{R}^N \to \mathbb{R}$ be functions of class C^1. Let $\bar{x} \in \mathbb{R}^N$ be a point of local minimum (or maximum) for F over the set $\{G = 0\}$. Assume that $\nabla G(\bar{x}) \neq 0$. Then there exists $\lambda \in \mathbb{R}$ such that*

$$\nabla F(\bar{x}) = \lambda \nabla G(\bar{x})\,.$$

Clearly, the above result generalizes to the case of multiple constraints.

3.6.2. The Lagrange multiplier rule for variational problems. The idea is to prove a similar result as in the finite dimensional case. The technical difficulty in proving this is that, basically, we cannot project the vector $v \in \mathbb{R}^N$ on the tangent space of $\mathcal{A}$ at the minimum point. But this turns out to be just a little modification in out argument.

[11] This is very easy to see: let us take a sequence $(x_n)_n$ with $G(x_n) = c$ for some $c \in \mathbb{R}$, and suppose $x_n \to \bar{x}$. By continuity $G(\bar{x}) = c$. Moreover (up to a subsequence), we can assume

$$\frac{x_n - \bar{x}}{|x_n - \bar{x}|} \to v \in \mathrm{Tan}_{\{G=c\}}(\bar{x})\,.$$

Thus

$$0 = \frac{G(x_n) - G(\bar{x})}{|x_n - \bar{x}|} \to \nabla G(\bar{x}) \cdot v\,.$$

Since each tangent vector can be obtained as limit of a sequence as above, we conclude.

THEOREM 3.42 (Lagrange multiplier rule). *Let $f, g : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be two functions of class C^1. Let $u \in \mathcal{A}$ be a point of local minimum for $\mathcal{F}$, where we define the admissible class as*

$$\mathcal{A} := \{u \in C^1([a,b]) \,:\, u(a) = \alpha,\, u(b) = \beta,\, \mathcal{G}(u) = c\}\,.$$

Assume there exists $\psi \in C_c^\infty((a,b))$ such that

$$\delta\mathcal{G}(u,\psi) \neq 0\,.$$

Then there exists $\lambda \in \mathbb{R}$ such that

$$\delta\mathcal{F}(u,\varphi) + \lambda\delta\mathcal{G}(u,\varphi) = 0\,,$$

for all $\varphi \in C_c^\infty((a,b))$. In particular, there exists a constant $c \in \mathbb{R}$ such that

$$\begin{aligned} &f_\xi\big(x,u(x),u'(x)\big) + \lambda g_\xi\big(x,u(x),u'(x)\big) \\ &\qquad = c + \int_a^x [\, f_p\big(x,u(x),u'(x)\big) + \lambda g_p\big(x,u(x),u'(x)\big)\,]\,\mathrm{d}x\,, \end{aligned}$$

for all $x \in [a,b]$.

PROOF. Take $\varphi \in C_c^\infty((a,b))$ and $\varepsilon > 0$. Consider the function $u + \varepsilon\varphi$. In general it is not true that it belongs to $\{\mathcal{G} = c\}$. The idea is to use the function ψ to modify it in order to satisfy the isoperimetric constraint. Heuristically, the idea is explained in Figure 5.

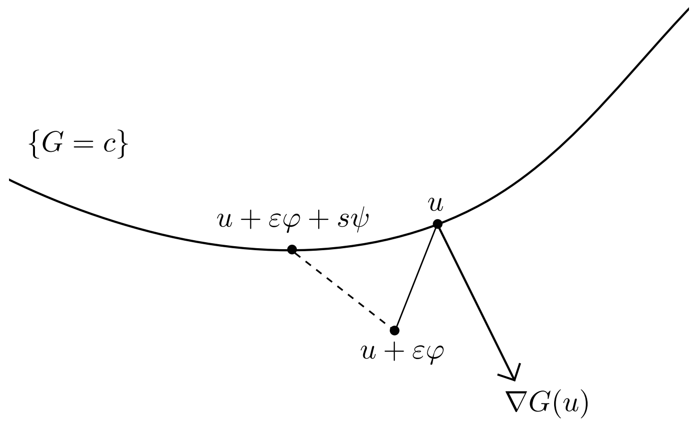

FIGURE 5. The main idea of the proof.

Let us consider the C^1 function

$$G(\varepsilon, s) := \mathcal{G}(u + \varepsilon\varphi + s\psi)\,.$$

We know that $G(0,0) = c$ and that $G_s(0,0) = \delta\mathcal{G}(u,\psi) \neq 0$. Thus it is possible to apply the implicit function theorem to obtain the existence of $\varepsilon_0 > 0$ and of a function $s : (-\varepsilon_0, \varepsilon_0) \to \mathbb{R}$ of class C^1 with $s(0) = 0$, such that

$$G(\varepsilon, s(\varepsilon)) = c\,,$$

for all $\varepsilon \in (\varepsilon_0, \varepsilon_0)$. This means that the function $u + \varepsilon\varphi + s(\varepsilon)\psi$ is admissible for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$. Moreover it holds:

$$s'(\varepsilon) = -\frac{G_\varepsilon(\varepsilon, s(\varepsilon))}{G_s(\varepsilon, s(\varepsilon))}\,.$$

In particular

$$s'(0) = -\frac{G_\varepsilon(0, s(0))}{G_s(0, s(0))} = -\frac{\delta\mathcal{G}(u, \varphi)}{\delta\mathcal{G}(u, \psi)}\,.$$

So, we can consider the function

$$\Phi(\varepsilon) := \mathcal{F}(u + \varepsilon\varphi + s(\varepsilon)\psi)\,,$$

and derive it at $\varepsilon = 0$ to obtain

$$0 = \Phi'(0) = \delta\mathcal{F}(u, \varphi) + \delta\mathcal{F}(u, \psi)s'(0) = \delta\mathcal{F}(u, \varphi) - \frac{\delta\mathcal{F}(u, \psi)}{\delta\mathcal{G}(u, \psi)}\delta\mathcal{G}(u, \varphi)\,.$$

The result follows by taking $\lambda := -\frac{\delta\mathcal{F}(u,\psi)}{\delta\mathcal{G}(u,\psi)}$ and by using the Du Bois-Reymond lemma. $\square$

Clearly, the extension of the above theorem to the case of multiple isoperimetric constraints is straightforward.

THEOREM 3.43. *Let f be a lagrangian of class C^1, and let $g_i : [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be functions of class C^1, for $i = 1, \ldots, m$. Consider the constrained minimization problem*

$$\min_{\mathcal{A}} \mathcal{F}\,,$$

where

$$\mathcal{A} := \{u \in C^1([a, b]) \,:\, u(a) = \alpha,\, u(b) = \beta,\, \mathcal{G}_i(u) = c_i\,,\ \textit{for all } i = i, \ldots, m\}\,,$$

where $c_i \in \mathbb{R}$ and

$$\mathcal{G}_i(u) := \int_a^b g_i\big(x, u(x), u'(x)\big)\,\mathrm{d}x\,.$$

Let $u \in \mathcal{A}$ be a point of local minimum for $\mathcal{F}$. Assume that there exists a function $\psi_i \in C_c^\infty((a, b))$ such that $\delta\mathcal{G}_i(u, \psi_i) \neq 0$. Then there exist $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ such that

$$\delta\mathcal{F}(u, \varphi) = \sum_{i=1}^m \lambda_i \delta\mathcal{G}_i(u, \varphi)\,,$$

for all $\varphi \in C_c^\infty((a, b))$.

REMARK 3.44. The above first order necessary conditions can be seen as the Euler-Lagrange equation for the functional

$$\widetilde{\mathcal{F}}_\lambda := \mathcal{F} + \lambda\mathcal{G}\,,$$

whose lagrangian is $\widetilde{f}_\lambda := f + \lambda g$. Let us suppose that, for each $x \in [a, b]$ the function

$$(p, \xi) \mapsto \widetilde{f}_\lambda(x, p, \xi)\,,$$

is convex. Then, we can conclude that any critical point of $\widetilde{\mathcal{F}}_\lambda$ is a minimizer for the constrained minimum problem. Moreover, if we have strict convexity, we gain uniqueness of the minimizer. Clearly, a minimizer for $\widetilde{\mathcal{F}}_\lambda$ is, in particular, a minimizer for the constrained problem for $\mathcal{F}$, up to choose the correct value of λ for which the minimizer of $\widetilde{\mathcal{F}}_\lambda$ turns out to be belong to the admissible class $\mathcal{A}$.

More precisely, we can solve the Euler-Lagrange equation for $\widetilde{\mathcal{F}}_\lambda$, getting a function u_λ and see if it is possible to find λ in such a way that $\widetilde{\mathcal{F}}_\lambda$ is convex and the functional $\mathcal{G}$ takes the desired value at the critical point u_λ.

REMARK 3.45. We can now make more precise the observation made in Remark 3.32. We noticed that the minimum problem

$$\min\Big\{\mathcal{F}(u) := \int_a^b f\big(x, u'(x)\big)\,\mathrm{d}x \;:\; u(a) = \alpha,\, u(b) = \beta\Big\},$$

can be translated into the following:

$$\min\Big\{\mathcal{G}(v) := \int_a^b f\big(x, v(x)\big)\,\mathrm{d}x \;:\; \int_a^b v(x)\,\mathrm{d}x = \beta - \alpha\Big\}.$$

This last problem fits into the just developed theory for isoperimetric problems. In this case the functional g is simply $g(p) := p$. Thus:

$$\delta\mathcal{G}(u, \varphi) = \int_a^b \varphi(x)\,\mathrm{d}x\,.$$

Thus, in order to satisfy the hypothesis of the above theorem, we just need to take a function $\psi \in C_c^\infty((a,b))$ such that $\int_a^b \psi(x)\,\mathrm{d}x \neq 0$ (that is[12], ψ is not a derivative!). Thus we obtain that, if $v \in C^1([a,b])$ with $\int_a^b v(x)\,\mathrm{d}x = \beta - \alpha$ is a local minimizer for $\mathcal{F}$, then there exists $\lambda \in \mathbb{R}$ such that

$$\frac{\mathrm{d}}{\mathrm{d}x} f_\xi\big(x, v(x)\big) = \lambda\,.$$

Since $v = u'$, we recover the well known Euler-Lagrange equation for lagrangians of the type $f = f(x, \xi)$.

REMARK 3.46. An interesting application of the Lagrange multipliers rule is the following: minimize

$$\int_a^b \frac{1}{2}|u'|^2\,\mathrm{d}x\,,$$

among all functions $u \in C_0^1([a,b])$ satisfying

$$\int_a^b u^2\,\mathrm{d}x = 1\,.$$

Then we obtain that a minimizer has to satisfy

$$u'' + \lambda u = 0\,, \qquad \text{on } [a,b]\,.$$

We will study this relation more in detail in Chapter 7.

3.7. Solution of the hanging cable problem

We now turn our attention to the hanging cable problem, that in mathematical terms asks to find the minimum of the functional

$$\mathcal{F}(u) := \int_0^b u(x)\sqrt{1 + \big(u'(x)\big)^2}\,\mathrm{d}x\,,$$

among all functions[13] $u \in C_0^1([a,b])$ satisfying

$$\mathcal{G}(u) := \int_0^b \sqrt{1 + \big(u'(x)\big)^2}\,\mathrm{d}x = l\,,$$

[12] Remember the Du Bois-Reymond lemma!

[13] For simplicity we consider the two ending points to be at the same height. Otherwise the problem of matching the boundary conditions will become a mess!

for some $l > b$. The problem is not written in the best possible way, that is, it turns out that is more convenient to consider the arc parametrization of the graph of the function u. So we introduce the arc parameter s given by

$$s(x) = \int_0^x \sqrt{1 + \big(u'(y)\big)^2}\, \mathrm{d}y\,.$$

Then it holds

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

and

$$\big(u(x(s))\big)' = u'\big(x(s)\big)x'(s)\,.$$

Thus

$$\big(u'(s)\big)^2 + \big(x'(s)\big)^2 = 1\,.$$

In particular we get $\big(u'(s)\big)^2 < 1$, where the strict inequality is due to the fact that the change of variable between x and s is a C^1 diffeomorphism. Then the problem translates as minimizing

$$\mathcal{F}(u) = g\rho \int_0^l u(s)\, \mathrm{d}s\,,$$

among all functions $u \in C^1_0([a,b])$ satisfying

$$\mathcal{G}(u) := \int_0^l \sqrt{1 - \big(u'(s)\big)^2}\, \mathrm{d}s = b\,.$$

Notice that in this case the two lagrangians are

$$f(p,\xi) = g\rho p\,, \qquad g(p,\xi) = \sqrt{1-\xi^2}\,.$$

Then

$$f_p(p_\xi) = g\rho\,, \qquad f_\xi(p,\xi) = 0\,,$$

and

$$g_p(p,\xi) = 0 \qquad g_\xi = -\frac{\xi}{\sqrt{1-\xi^2}}\,, \qquad g_{\xi\xi}(p,\xi) = \frac{2\xi^2 - 1}{(\xi^2-1)^2}\,.$$

Recalling that $\big(u'(s)\big)^2 < 1$, we have that g is strictly convex, and thus $f + \lambda g$ is strictly convex for every $\lambda > 0$. Thus the critical point we will find will be the only minimizer of the problem. The Lagrange multiplier's rule provides us the equation

$$\left(\frac{\lambda u'(s)}{\sqrt{1 - \big(u'(s)\big)^2}} \right)' = g\rho\,,$$

from which we deduce

$$u'(s) = \frac{c + g\rho s}{\sqrt{\lambda^2 + (c + g\rho s)^2}}\,,$$

for some $c \in \mathbb{R}$. By integrating we get

$$u(s) = \int_0^s \frac{c + g\rho t}{\sqrt{\lambda^2 + (c+g\rho t)^2}}\, \mathrm{d}t = \frac{1}{g\rho}\big[\, \sqrt{\lambda^2 + (c+g\rho s)^2} - \sqrt{\lambda^2 + c^2}\,\big]\,.$$

We now have to find a $\lambda > 0$ such that u satisfies $\mathcal{G}(u) = b$, and to find a corresponding $c \in \mathbb{R}$ to match the boundary conditions. By making the physical assumption that the shape of the cable is symmetric with respect to the middle point $\frac{l}{c}$, we deduce that we must have $u'(\frac{l}{2}) = 0$, and thus that $c = -g\rho\frac{l}{2}$. Notice that this choice of the constant c will make

u satisfying the boundary conditions $u(0) = u(b) = 0$. Let us now care about the integral constraint. We have

$$\begin{aligned}
x(s) &= \int_0^s \sqrt{1 - \big(u'(t)\big)^2}\,\mathrm{d}t = \\
&= \int_0^l \frac{1}{\sqrt{1 + \Big(\frac{c+g\rho t}{\lambda}\Big)^2}}\,\mathrm{d}t \qquad \Big[\text{substitution } z = \frac{c + g\rho t}{\lambda}\Big] \\
&= \frac{\lambda}{g\rho} \int_{-g\rho\frac{l}{2\lambda}}^{\frac{g\rho}{\lambda}(s-\frac{l}{2})} \frac{1}{\sqrt{1+z^2}}\,\mathrm{d}z \\
&= \frac{\lambda}{g\rho} \sinh^{-1}\Big(\frac{g\rho}{\lambda}\Big(s - \frac{l}{2}\Big)\Big) + \sinh^{-1}\Big(\frac{g\rho}{\lambda}\frac{l}{2}\Big).
\end{aligned}$$

In particular, by imposing $x\big(\frac{l}{2}\big) = \frac{b}{2}$ (recall the symmetry!), we have

$$\sinh^{-1}\Big(\frac{g\rho}{\lambda}\frac{l}{2}\Big) = \frac{b}{2}.$$

So

$$\frac{\lambda}{g\rho} \sinh^{-1}\Big(\frac{g\rho}{\lambda}\Big(s - \frac{l}{2}\Big)\Big) + \frac{b}{2},$$

from which we deduce

$$s(x) = \frac{l}{2} - \frac{\lambda}{g\rho}\sinh^{-1}\Big(\frac{g\rho}{\lambda}\Big(\frac{b}{2} - x\Big)\Big).$$

By inserting this expression in the formula for u we got above, we find

$$u(x) = \frac{\lambda}{g\rho}\cosh\Big(\frac{g\rho}{\lambda}\Big(\frac{b}{2} - x\Big)\Big) - \frac{1}{g\rho}\sqrt{\lambda^2 + \big(\frac{l}{2}\big)^2}.$$

Thus we discovered that the minimizer of the hanging cable problem is a *catenary*.

3.8. Broken extremals

In this section we want to extend the class of admissible function in order to allow jumps of the derivatives. We first need to introduce exactly what kind of functions we want to consider.

DEFINITION 3.47. We say that a function $u \in C^0([a,b])$ is *piecewise* C^1 if there exist $a = x_0 < x_1 < \cdots < x_{k+1} = b$ such that $u_{|[t_i,t_{i+1}]} \in C^1([t_i,t_{i+1}])$ for all $k = 0,\ldots,k$. This space will be denoted by $C^1_{pw}([a,b])$.

REMARK 3.48. Notice that the limits

$$u_l'(x_i) := \lim_{x\to x_i^-} u'(x), \qquad u_r'(x_i) := \lim_{x\to x_i^+} u'(x)$$

exist and are finite.

Let us consider the admissible class

$$\mathcal{A} := \{u \in C^1_{pw}([a,b]) \,:\, u(a) = \alpha,\, u(b) = \beta\}.$$

3.8.1. The first Weierstrass-Erdmann condition. Suppose the lagrangian f is of class C^1 and let $c \in \mathcal{A}$ be a *weak* local minimizer. Consider, for $\varphi \in C_c^\infty((a,b))$ the usual variation of u and the corresponding function $\Phi(\varepsilon) := \mathcal{F}(u+\varepsilon\varphi)$. Then

$$\begin{aligned}
0 = \Phi'(0) &= \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big[\int_a^b f\big(x, u(x)+\varepsilon\varphi(x), u'(x)+\varepsilon\varphi'(x)\big)\,\mathrm{d}x\Big]_{|\varepsilon=0} \\
&= \sum_{i=1}^{k+1} \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big[\int_{t_{i-1}}^{t_i} f\big(x, u(x)+\varepsilon\varphi(x), u'(x)+\varepsilon\varphi'(x)\big)\,\mathrm{d}x\Big]_{|\varepsilon=0} \\
&= \sum_{i=1}^{k+1} \int_{t_{i-1}}^{t_i} \big[f_p\big(x,u(x),u'(x)\big)\varphi(x) + f_\xi\big(x,u(x),u'(x)\big)\varphi'(x)\big]\,\mathrm{d}x \\
&= \int_a^b \big[f_p\big(x,u(x),u'(x)\big)\varphi(x) + f_\xi\big(x,u(x),u'(x)\big)\varphi'(x)\big]\,\mathrm{d}x\,.
\end{aligned}$$

Let us now introduce the function

$$A(t) := \int_a^t f_p\big(x,u(x),u'(x)\big)\,\mathrm{d}x\,.$$

Then $A \in C([a,b])$. By looking at the proof of the Du Bois-Reymond lemma (see 3.20) it is easy to see that the same result holds for every point of continuity of the functions g and h. In particular we get that

$$f_\xi\big(x,u(x),u'(x)\big) = c + A(x)\,, \qquad \forall x \in [a,b]\setminus\{x_1,\dots,x_k\}\,.$$

Since the right-hand side is continuous, we get that

$$f_\xi\big(x_i,u(x_i),u_l'(x_i)\big) = f_\xi\big(x_i,u(x_i),u_r'(x_i)\big)\,, \qquad \forall i = 1,\dots,k\,. \tag{3.10}$$

The above condition is called the *first Weierstrass-Erdmann condition* and says that in a discontinuity point x_i for the derivative of a weak local minimizer, the jump of the derivative can be only between points ξ's for which the function

$$(x_i,u(x_i),\xi) \mapsto f\big(x_i,u(x_i),\xi\big)$$

have the same derivative.

3.8.2. The second Weierstrass-Erdmann condition. We now want to derive a condition for *strong* local minimizers of $\mathcal{F}$ similar to the Du Bois-Reymond one. We notice that, in contrast with what we've just done, here we need u to be a strong local minimizer and not only a weak one. The reason is that, since the derivative of u can jump, we cannot say that u' is uniformly continuous on $[a,b]$. With the strong local minimality assumption in force, we can reason in the same way as for the derivation of the Du Bois-Reymond equation and getting

$$f\big(x,u(x),u'(x)\big) - u'(x)f_\xi\big(x,u(x),u'(x)\big) = c + \int_a^x f_x\big(t,u(t),u'(t)\big)\,\mathrm{d}t\,,$$

for all $x \in [a,b]\setminus\{x_1,\dots,x_k\}$. Since the function on the right-hand side is continuous, we get

$$f\big(x_i,u(x_i),u_l'(x_i)\big) - u_l'(x_i)f_\xi\big(x_i,u(x_i),u_l'(x_i)\big) = f\big(x_i,u(x_i),u_r'(x_i)\big) - u_r'(x_i)f_\xi\big(x_i,u(x_i),u_r'(x_i)\big)\,, \tag{3.11}$$

for all $i = 1,\dots,k$. The above condition is called the *second Weierstrass-Erdmann condition.*

We want to give a geometrical interpretation of these two conditions, for a strong local minimizer. Consider, for all $i = 1,\dots,k$, the function

$$g(x_i,u(x_i),\xi) := f\big(x_i,u(x_i),\xi\big)\,.$$

Then conditions (3.10) and (3.11) say that in a point of discontinuity of the derivative of u, the derivative can jump only between points with *the same* tangent line for the function g. Indeed, in a generic point $\xi \in \mathbb{R}$ we have that the tangent line to the function g is given by

$$f\big(x_i, u(x_i), \xi\big) + f_\xi\big(x_i, u(x_i), \xi\big)(\xi - u'(x_i))\,.$$

Then the two Weierstrass-Erdmann conditions will give us that this line must be the same at the points $u_l'(x_i)$ and $u_r'(x_i)$.

3.8.3. An example. Let us consider the lagrangian of the Weierstrass's paradox

$$f(\xi) := (\xi^2 - 1)^2\,.$$

Then

$$f_\xi(\xi) = 4\xi(\xi^2 - 1)\,, \qquad f(\xi) - \xi f_\xi(\xi) = (\xi^2 - 1)(-3\xi^2 - 1)\,.$$

From (3.10) and (3.11) we get that, in a corner point x_i, we must have

$$4u_l'(x_i)\big(\big(u_l'(x_i)\big)^2 - 1\big)^2 = 4u_r'(x_i)\big(\big(u_r'(x_i)\big)^2 - 1\big)^2\,,$$

and

$$\big(\big(u_l'(x_i)\big)^2 - 1\big)\big(-3\big(u_l'(x_i)\big)^2 - 1\big) = \big(\big(u_r'(x_i)\big)^2 - 1\big)\big(-3\big(u_r'(x_i)\big)^2 - 1\big)\,.$$

Clearly this is possible only if $\big(u_l'(x_i)\big)^2 = \big(u_r'(x_i)\big)^2 = 1$.

CHAPTER 4

First order necessary conditions for general functions

4.1. The Euler-Lagrange equation

In this section we want to generalize the ideas leading to the first order necessary conditions to the case of functions

$$u : \Omega \to \mathbb{R}^M ,$$

where $\Omega \subset \mathbb{R}^N$ is an open set with boundary of class C^1 (see Appendix, Definition 11.10). For such a functions we have to consider lagrangians of the type

$$f : \Omega \times \mathbb{R}^M \times \mathbb{R}^{M\times N} \to \mathbb{R} .$$

We will still denote by $(x, p, \xi) = (x_1, \dots x_N, p^1, \dots, p^M, \xi_1^1, \dots, \xi_M^N)$ the variables to which we apply the function f. Moreover, in order to make the formulas more light, we will use the following notation

$$f_{x_\alpha} := \frac{\partial f}{\partial x_\alpha} , \qquad f_{p^i} := \frac{\partial f}{\partial p^i} , \qquad f_{\xi_\alpha^i} := \frac{\partial f}{\partial \xi_\alpha^i} .$$

The symbol D_α will denote the (total) derivative with respect to the variable α. Finally, we will consider the Einstein convention about repeated indexes, namely we will sum over repeated indexes (the Greek ones from 1 to N and the latin ones from 1 to M). For instance

$$a_\alpha^i b^i c_\alpha = \sum_{i=1}^{M} \sum_{\alpha=1}^{N} a_\alpha^i b^i c_\alpha .$$

The definitions of weak a strong local minimizers are the same as in the one dimensional scalar case.

Let us suppose[1] that the lagrangian f is of class C^2. Let $u \in C^2(\bar{\Omega}; \mathbb{R}^M)$ be a weak local minimizer for the functional

$$\mathcal{F}(u) := \int_\Omega f\big(x, u(x), Du(x)\big) ,$$

over the class

$$\mathcal{A} := \{v \in C^1(\bar{\Omega}; \mathbb{R}^M) \;\; v_{|\partial\Omega} = u_{|\partial\Omega}\} .$$

Take a test function $\varphi \in C_c^\infty(\Omega; \mathbb{R}^M)$ and consider the variation $u + \varepsilon\varphi$ and the function

$$\Phi(\varepsilon) := \mathcal{F}(u + \varepsilon\varphi) .$$

Then[2]

$$\begin{aligned} 0 = \Phi'(0) &= \int_\Omega \big[f_{p^i}\big(x, u(x), Du(x)\big)\varphi^i(x) + f_{\xi_\alpha^i}\big(x, u(x), Du(x)\big) D_\alpha \varphi^i\big] \,\mathrm{d}x \\ &= \int_\Omega \big[f_{p^i}\big(x, u(x), Du(x)\big) - D_\alpha f_{\xi_\alpha^i}\big(x, u(x), Du(x)\big)\big]\varphi^i(x) \,\mathrm{d}x , \end{aligned}$$

[1]Clearly, we know that we can relax the regularity assumption on f by just requiring the C^2 regularity with respect to p and to ξ.

[2]By reasoning in the same way we did for the scalar one dimensional case.

holds for every test function $\varphi \in C_c^\infty(\Omega;\mathbb{R}^M)$. To pass from the integral necessary condition to the differential one, we need the following generalization of the fundamental lemma of Calculus of Variations.

LEMMA 4.1. *Let $g \in C^0(\Omega)$ such that*

$$\int_\Omega g(x)\psi(x)\,\mathrm{d}x = 0\,, \qquad \forall \psi \in C_c^\infty(\Omega)\,.$$

Then $g \equiv 0$ on Ω.

PROOF. Let us suppose there exists $\bar{x} \in \Omega$ such that $g(\bar{x}) \neq 0$. As usual we can suppose $g(\bar{x}) > 0$ and that, by continuity, that $g(x) > 0$ in $\bar{B}_\delta(\bar{x})$, for some $\delta > 0$. Let us consider the function

$$\psi(x) := \begin{cases} e^{\frac{1}{|x-\bar{x}|^2-\delta^2}} & \text{if } |x-\bar{x}| < \delta\,, \\ 0 & \text{otherwise}\ . \end{cases}$$

Then $\psi \in C_c^\infty(\Omega)$ and $\psi > 0$ in $B_\delta(\bar{x})$. Thus we obtain

$$0 = \int_\Omega g(x)\psi(x)\,\mathrm{d}x = \int_{B_\delta(\bar{x})} g(x)\psi(x)\,\mathrm{d}x > 0\,.$$

□

By applying the previous lemma for each $i = 1, \dots, M$, we obtain the following system of differential equations: for all $i = 1, \dots, M$ it holds

$$f_{p^i}\big(x, u(x), Du(x)\big) - D_\alpha f_{\xi^i_\alpha}\big(x, u(x), Du(x)\big) = 0\,, \qquad \text{in } \Omega\,.$$

The above system can be written in a more compact form as follows: for a $M \times N$ (M rows and N columns) matrix $A = A(x)$ we denote by $\mathrm{div}A$ the vector of $\mathbb{R}^N$ whose i^{th} component is the divergence of the i^{th} row of A. Using this notation we can write the above system as

$$f_p\big(x, u(x), Du(x)\big) - \mathrm{div}\big(f_\xi\big(x, u(x), Du(x)\big)\big) = 0\,, \qquad \text{in } \Omega\,.$$

Thus we have obtained the following result

DEFINITION 4.2. The *Euler operator* L_f of the lagrangian f is defined as:

$$L_f(u) := f_p\big(x, u(x), Du(x)\big) - \mathrm{div}\big(f_\xi\big(x, u(x), Du(x)\big)\big)\,.$$

THEOREM 4.3. *Let $\Omega \subset \mathbb{R}^N$ be an open bounded set with boundary of class C^1. Suppose that the lagrangian $f : \Omega \times \mathbb{R}^M \times \mathbb{R}^{M\times N} \to \mathbb{R}$ is of class C^2. Let $u \in C^2(\bar{\Omega};\mathbb{R}^M)$ be a weak local minimizer of $\mathcal{F}$ among C^1 functions with the same boundary value. Then it holds*

$$L_f(u) = 0\,, \qquad \textit{in } \Omega\,. \tag{4.1}$$

4.1.0.1. *Some examples.*

(i) Let us consider the so called *Dirichlet functional*

$$\int_\Omega \frac{1}{2}|\nabla u|^2\,\mathrm{d}x\,.$$

The Euler-Lagrange equation for this functional is

$$\triangle u = 0\,, \qquad \text{in } \Omega\,.$$

Notice that, since the lagrangian is stictly convex, every solution of the above equation (that matches the desired boundary conditions) will be the uniqeu minimizer of the Dirichet functional.

(ii) Let us consider the functional
$$\int_\Omega \left(\frac{1}{2}|\nabla u|^2 + fu\right) \mathrm{d}x\,.$$
The Euler-Lagrange equation for this functional is
$$\triangle u = f\,, \qquad \text{in } \Omega\,,$$
the so called *Poisson equation*.

(iii) Let us consider the area functional for functions $u : \Omega \to \mathbb{R}$:
$$\mathcal{A}(u) := \int_\Omega \sqrt{1 + |\nabla u|^2}\, \mathrm{d}x\,.$$
Its Euler-Lagrange equation writes as
$$\operatorname{div}\left(\frac{\nabla u}{\sqrt{1+|\nabla u|^2}}\right) = 0\,, \qquad \Omega\,.$$
The above equation is known as the *minimal surface equation*.

(iv) Let us consider the the functional
$$\int_\Omega \big(\sqrt{1+|\nabla u|^2} + Hu\big)\, \mathrm{d}x\,.$$
Then we obtain the condition
$$\operatorname{div}\left(\frac{\nabla u}{\sqrt{1+|\nabla u|^2}}\right) = H\,, \qquad \Omega\,.$$
The above equation is called the *prescribed mean curvature equation*. The reason is that the quantity on the right-hand side is the mean curvature of $\operatorname{graph}(u)$.

4.2. Natural boundary conditions

We now investigate the case of weak local minimizers $u \in C^2(\Omega;\mathbb{R}^M)$ of the functional $\mathcal{F}$ in the class
$$C^1(\Omega;\mathbb{R}^M)\,.$$
Reasoning as above, for a test function $\varphi \in C^\infty(\bar\Omega;\mathbb{R}^M)$ we obtain
$$\int_\Omega L_f(u)\cdot\varphi\, \mathrm{d}x + \int_{\partial\Omega} \varphi^i f_{\xi^i_\alpha}\nu_\alpha\, \mathrm{d}\sigma = 0\,, \qquad \forall \varphi \in C^\infty(\bar\Omega;\mathbb{R}^M)\,,$$
where σ denotes the surface measure on $\partial\Omega$. Since $C^\infty_c(\Omega;\mathbb{R}^M) \subset C^\infty(\bar\Omega;\mathbb{R}^M)$, we still get the necessary condition $L_f(u) = 0$ in Ω. Thus we are left with the condition
$$\int_{\partial\Omega} \varphi^i f_{\xi^i_\alpha}\nu_\alpha\, \mathrm{d}\sigma = 0\,, \qquad \forall\varphi \in C^\infty(\bar\Omega;\mathbb{R}^M)\,.$$
We need a generalization of the fundamental lemma of the Calculus of Variations for boundaries of regular sets.

LEMMA 4.4. *Let $g \in C(\bar\Omega)$ be such that*
$$\int_{\partial\Omega} g\varphi\, \mathrm{d}\sigma = 0\,, \qquad \forall\varphi\in C^\infty(\bar\Omega)\,.$$
Then $g_{|\partial\Omega} = 0$.

PROOF. Let us suppose there exists $\bar{x} \in \partial\Omega$ such that $g(\bar{x}) \neq 0$. Without loss of generality, we can assume $g(\bar{x}) > 0$. Since the boundary of Ω is of class C^1, we can assume (up to a rotation and a translation) that there exist $r > 0$ and $\Psi : \mathbb{R}^{N-1} \to \mathbb{R}$ of class C^1 such that

$$\partial\Omega \cap B_r(\bar{x}) = \{(x', x_N) \in \mathbb{R}^{N-1} \times \mathbb{R} \ : \ |x'| < r, \, x_n = \psi(x')\} ,$$

and

$$\Omega \cap B_r(\bar{x}) = \{(x', x_N) \in \mathbb{R}^{N-1} \times \mathbb{R} \ : \ |x'| < r, \, x_n > \psi(x')\} .$$

By continuity of g there exists $\delta > 0$ such that

$$g(x) > 0 , \qquad \forall x \in \partial\Omega \cap B_\delta(\bar{x}) .$$

We can assume $\delta < r$. Let us define the function $\widetilde{g} : B_r^{N-1} \to \mathbb{R}$ (where B_r^{N-1} denotes the ball of $\mathbb{R}^{N-1}$)

$$\widetilde{g}(x') := g\big(x', \psi(x')\big) .$$

Moreover, for every function $\widetilde{\varphi} \in C_c^\infty(B^{N-1})$ we define the function $\varphi \in C^1(\bar{\Omega})$ as a regular extension of the function

$$\varphi\big(x', \psi(x')\big) := \widetilde{\varphi}(x') .$$

The existence of such an extension is easy to prove[3]. For such an extension φ, it holds

$$\int_{\partial\Omega} g(x)\varphi(x) \, \mathrm{d}\sigma = \int_{B^{N-1}} g\big(x', \psi(x')\big)\widetilde{\varphi}(x')\sqrt{1 + |\nabla\psi(x')|^2} \, \mathrm{d}\sigma(x') .$$

We have a little problem of regularity, since the function Ψ is only of class C^1, and thus we cannot conclude that the above quantity is always 0 for each function $\widetilde{\varphi} \in C_c^\infty(B^{N-1})$. To overcome this problem we can reason in two ways: we can notice that, actually, in our case we have

$$\int_{\partial\Omega} g\varphi \, \mathrm{d}\sigma = 0 , \qquad \forall \varphi \in C^1(\bar{\Omega}) ,$$

or we can approximate every C^1 function with C^∞ functions. Either e ways, we obtain that

$$\int_{B^{N-1}} g\big(x', \psi(x')\big)\widetilde{\varphi}(x')\sqrt{1 + |\nabla\psi(x')|^2} \, \mathrm{d}\sigma(x') = 0 ,$$

for all $\widetilde{\psi} \in C_c^\infty(B^{N-1})$. Since $\sqrt{1 + |\nabla\psi(x')|^2} > 0$, we can say that the above equality holds true for every test function in $C_c^\infty(B^{N-1})$. Thus, by the standard fundamental lemma we conclude. □

Thus, by applying the above lemma, we get the following result

THEOREM 4.5. *Let $\Omega \subset \mathbb{R}^N$ be an open bounded set with boundary of class C^1. Suppose that the lagrangian $f : \Omega \times \mathbb{R}^M \times \mathbb{R}^{M\times N} \to \mathbb{R}$ is of class C^2. Let $u \in C^2(\bar{\Omega}; \mathbb{R}^M)$ be a weak local minimizer of $\mathcal{F}$ among C^1 functions. Then it holds*

$$\begin{cases} L_f(u) = 0 & \text{in } \Omega , \\ \\ f_{\xi_\alpha^i}\nu_\alpha = 0 & \text{on } \partial\Omega , \quad \forall i = 1, \dots, M . \end{cases}$$

The last conditions are called natural boundary conditions.

[3]For example one can extend the function φ in a constant way in the vertical direction and then letting it go to zero.

4.2.0.2. *Some examples.*

(i) The natural boundary condition for the functional
$$\int_\Omega \left(\frac{1}{2}|\nabla u|^2 + fu\right)\,\mathrm{d}x\,,$$
is $\partial_\nu u = 0$ on $\partial\Omega$, where ν denotes the normal vector to $\partial\Omega$.

(ii) In the case of the functional
$$\int_\Omega \left(\sqrt{1+|\nabla u|^2} + Hu\right)\,\mathrm{d}x\,,$$
the natural boundary condition reads as
$$\nu \cdot \frac{\nabla u}{\sqrt{1+|\nabla u|^2}} = 0 \qquad \text{on } \partial\Omega\,.$$
Let us give a geometric interpretation of this condition: let us consider the cylinder $\Omega\times\mathbb{R}$, whose normal on $\partial\Omega\times\mathbb{R}$ is $(\nu, 0)$. Noticing that the normal to graph$u$ is given by
$$\left(\frac{\nabla u}{\sqrt{1+|\nabla u|^2}}, -1\right),$$
we obtain that the natural boundary condition for the above functional requires the graph of u to meet $\partial\Omega\times\mathbb{R}$ orthogonally.

4.3. Inner variations

We now want to derive the analogous of the Du Bois-Reymond equation in the general case. The idea is to perform inner variations, *i.e.*, variations of the independent variable. So, suppose $\Omega\subset\mathbb{R}^N$ is an open bounded set with boundary of class C^1 and take a vector field $\lambda\in C^\infty_c(\Omega;\mathbb{R}^N)$. We consider the function $\Psi:\mathbb{R}\times\bar\Omega\to\mathbb{R}^N$ given by
$$\Psi(\varepsilon, x) := x + \varepsilon\lambda(x)\,.$$
We will denote by Ψ^α its α's component. We claim that there exists $\varepsilon_0>0$ such that for all $\varepsilon\in(0,\varepsilon_0)$ the function
$$x\mapsto \Psi(\varepsilon, x)$$
is a diffeomorphism of $\bar\Omega$ onto itself. The proof of this fact is technical and will be postponed in the Appendix (see Appendix Section11.5). Define $v_\varepsilon := u\big(\Psi(\varepsilon,\cdot)\big)$ and let us consider the function
$$\begin{aligned}\Phi(\varepsilon) := \mathcal{F}(v_\varepsilon) &= \int_\Omega f\big(x, v_\varepsilon(x), Dv_e(x)\big)\,\mathrm{d}x\\ &= \int_\Omega f\big(x, u\big(\Psi(\varepsilon,x)\big), Du\big(\Psi(\varepsilon,x)\big)D\Psi(\varepsilon,x)\big)\,\mathrm{d}x\,.\end{aligned}$$
By using the change of variable $y=\Psi(\varepsilon,x)$ (recall that ε is fixed!) and denoting by $\eta(\varepsilon,\cdot)$ the inverse function of $\Psi(\varepsilon,\cdot)$, we can write
$$\Phi(\varepsilon) = \int_\Omega f\big(\eta(\varepsilon,y), u(y), Du(y)D\Psi\big(\varepsilon,\eta(\varepsilon,y)\big)\big)\det D\eta(\varepsilon,y)\,\mathrm{d}y\,.$$
If $u\in C^2(\bar\Omega;\mathbb{R}^M)$ and the lagrangian f is of class C^2, then we know that
$$\partial\mathcal{F}(u,\lambda) := \Phi'(0) = 0\,.$$

In order to write the derivative of the above object, we need some preliminary computations. Let us notice that

$$\eta(\varepsilon, y) = y - \varepsilon\lambda(y) + o(\varepsilon)\,.$$

Hence

$$\frac{\partial}{\partial\varepsilon}\eta(\varepsilon, y)_{|\varepsilon=0} = -\lambda(y)\,.$$

Moreover, since $D\Psi(\varepsilon, x) = \mathrm{Id} + \varepsilon D\lambda(x)$, we have

$$\frac{\partial}{\partial\varepsilon}D\Psi\big(\varepsilon, \eta(\varepsilon, y)\big)_{|\varepsilon=0} = D\lambda(y)\,.$$

This implies that

$$\begin{aligned}\frac{\partial}{\partial\varepsilon}\Big(\big(Du(y)D\Psi(\varepsilon, y)\big)_{i\beta}\Big)_{|\varepsilon=0} &= \frac{\partial}{\partial\varepsilon}\big(u^i_{x_\alpha}(y)D_\beta\Psi^\alpha\big(\varepsilon, \eta(\varepsilon, y)\big)\big)_{|\varepsilon=0}\\ &= u^i_{x_\alpha}(y)D_\beta\lambda^\alpha(y)\,.\end{aligned}$$

Finally, the derivative of the determinant turns out to be (see Appendix, Section 11.4)

$$\frac{\partial}{\partial\varepsilon}\det D\eta(\varepsilon, y)_{|\varepsilon=0} = -\mathrm{div}\lambda(y) = D_\alpha\lambda^\alpha(y)\,.$$

Thus we can write

$$\begin{aligned}0 = \int_\Omega \big[&-f_{x_\alpha}\big(x, u(x), Du(x)\big)\lambda^\alpha(x) + f_{\xi^i_\beta}\big(x, u(x), Du(x)\big)u^i_{x_\alpha}(x)D_\beta\lambda^\alpha(x)\\ &- f\big(x, u(x), Du(x)\big)D_\alpha\lambda^\alpha(x)\,\big]\,\mathrm{d}x\,.\end{aligned}$$

By integrating by parts the last two terms, and recalling that $\lambda = 0$ on $\partial\Omega$, we obtain

$$\begin{aligned}0 = \int_\Omega \big[&-f_{x_\alpha}\big(x, u(x), Du(x)\big) - D_\beta f_{\xi^i_\beta}\big(x, u(x), Du(x)\big)u^i_{x_\alpha}(x)\\ &+ D_\alpha f\big(x, u(x), Du(x)\big)\,\big]\lambda^\alpha(x)\,\mathrm{d}x\end{aligned}$$

Since this holds true for each vector field $\lambda \in C^\infty_c(\Omega; \mathbb{R}^N)$, we conclude that

$$-f_{x_\alpha}\big(x, u(x), Du(x)\big) - D_\beta f_{\xi^i_\beta}\big(x, u(x), Du(x)\big)u^i_{x_\alpha}(x) + D_\alpha f\big(x, u(x), Du(x)\big) = 0$$

in Ω, for all $\alpha = 1, \ldots, M$. It is possible to write the above equation in a more compact way by introducing the so called *energy-momentum tensor*

$$T^\beta_\alpha(x) := f_{\xi^i_\beta}\big(x, u(x), Du(x)\big)u^i_{x_\alpha}(x) - \delta^\beta_\alpha f\big(x, u(x), Du(x)\big)\,,$$

where $\delta^\beta_\alpha = 1$ if $\beta = \alpha$ and 0 otherwise. Then we have

THEOREM 4.6. *Let $\Omega \subset \mathbb{R}^N$ be an open bounded set with boundary of class C^1. Suppose that the lagrangian $f : \Omega \times \mathbb{R}^M \times \mathbb{R}^{M\times N} \to \mathbb{R}$ is of class C^1. Let $u \in C^2(\bar\Omega; \mathbb{R}^M)$ be a weak local minimizer of $\mathcal{F}$ among C^1 functions with the same boundary value. Then it holds*

$$D_\beta T^\beta_\alpha + f_{x_\alpha} = 0\,, \qquad \text{in } \Omega\,, \tag{4.2}$$

for all $\alpha = 1, \ldots, M$.

We now want to see how to write the inner variation $\partial\mathcal{F}(u, \cdot)$ by using the Euler operator.

LEMMA 4.7. *It holds:*

$$\partial\mathcal{F}(u, \lambda) = -\int_\Omega (L_f(u)\cdot D_\alpha u)\cdot\lambda\,\mathrm{d}x\,.$$

PROOF. We have that[4]

$$
\begin{aligned}
-\partial\mathcal{F}(u,\lambda) &= \int_\Omega \big[\, f_{x_\alpha}\lambda^\alpha + f D_\alpha\lambda^\alpha - f_{\xi^i_\beta} u^i_{x_\alpha}\lambda^\alpha_{x_\beta} \,\big]\,\mathrm{d}x = \int_\Omega \big[\, f_{x_\alpha} - D_\alpha f + D_\beta(f_{\xi^i_\beta} u^i_{x_\alpha})\,\big]\lambda^\alpha\,\mathrm{d}x \\
&= \int_\Omega \big[\, f_{x_\alpha} - (f_{x_\alpha} + f_{p^i}u^i_{x_\alpha} + f_{\xi^i_\beta} D_\beta u^i_{x_\alpha}) + (u^i_{x_\alpha} D_\beta f_{x^i_\beta} + f_{x^i_\beta} D_\beta u^i_{x_\alpha})\,\big]\lambda^\alpha\,\mathrm{d}x \\
&= \int_\Omega \big[\, D_\beta f_{x^i_\beta} - f_{p^i}\,\big] u^i_{x\alpha}\lambda^\alpha\,\mathrm{d}x\,.
\end{aligned}
$$

□

REMARK 4.8. By the above lemma we get that the condition

$$\partial\mathcal{F}(u,\lambda) = 0\,,$$

for all $\lambda \in C^\infty_c(\Omega;\mathbb{R}^N)$ implies that

$$L_f(u)\cdot D_\alpha u = 0$$

in Ω, that is the Euler operator is perpendicular to the surface

$$\{\,(x,u(x)) \in \mathbb{R}^N \times \mathbb{R}^M \,:\, x\in\Omega\,\}\,.$$

On the other hand the Euler-Lagrange equation (4.1) tells us that $L_f(u) = 0$ in Ω. So, a function u that satisfies the Euler-Lagrange equation will also satisfy $\partial\mathcal{F}(u,\lambda) = 0$ for all $\lambda \in C^\infty_c(\Omega;\mathbb{R}^N)$.

4.4. Isoperimetric problems

The same result for Lagrange multipliers in the one dimensional scalar case turns out to be true also for the general case. since the proof is the same, we just state it.

THEOREM 4.9 (Lagrange multiplier rule). *Let $f,g : \Omega\times\mathbb{R}^M\times\mathbb{R}^{MN} \to \mathbb{R}$ be two functions of class C^1. Let $u\in\mathcal{A}$ be a point of local minimum for $\mathcal{F}$, where we define the admissible class as*

$$\mathcal{A} := \{u \in C^1(\bar\Omega : \mathbb{R}^M) \,:\, u_{|\partial\Omega} = h,\ \mathcal{G}(u) = c\}\,,$$

for some function $h\in C^0(\partial\Omega)$ and some $c\in\mathbb{R}$. Assume there exists $\psi\in C^\infty_c(\Omega;\mathbb{R}^M)$ for which

$$\delta\mathcal{G}(u,\psi) \neq 0\,.$$

Then there exists $\lambda\in\mathbb{R}$ such that

$$\delta\mathcal{F}(u,\varphi) + \lambda\delta\mathcal{G}(u,\varphi) = 0\,,$$

for all $\varphi\in C^\infty_c(\Omega;\mathbb{R}^M)$. In particular, if in addition $u \in C^1(\bar\Omega;\mathbb{R}^M)\cap C^2(\Omega;\mathbb{R}^M)$, then

$$\operatorname{div}(f_\xi + \lambda g_\xi) = f_p + \lambda g_p\,.$$

[4] By omitting the arguments of the functions for sake of clearness, but you know what they are!

4.5. Holomic constraints

We now want to derive a first order necessary condition for minimizer of a variational integral that are constrained to lie on a submanifold. So, we consider the problem

$$\min_{v \in \mathcal{A}_G} \mathcal{F}(v)\,,$$

where the admissible class is defined as

$$\mathcal{A}_G : \{\, v \in C^1(\Omega; \mathbb{R}^M) \ : \ v_{|\partial\Omega} = h\,,\ G\big(x, v(x)\big) = 0\,,\ \text{for all } x \in \Omega\,\}\,,$$

for some function $h \in C^1(\partial\Omega; \mathbb{R}^M)$. Since we have in mind the case of manifolds (that, for the sake of generality, can vary from point to point) we ask G to be a function in $C^2(\Omega \times \mathbb{R}^M; R^k)$, where $k \in \{1, \dots, M-1\}$, with the matrix $G_p(x,p)$ of maximal rank k at each point $(x,p) \in \Omega \times \mathbb{R}^M$ such that $G(x,p) = 0$. This will ensure that the set

$$M(x) := \{\, p \in \mathbb{R}^M \ : \ G(x,p) = 0\,\}\,,$$

turns out to be a C^2 submanifold of $\mathbb{R}^M$ of dimension $M - k$. Let us denote by $T_pM(x)$ the tangent plane of $M(x)$ at the point $p \in \mathbb{R}^M$, and with $\Pi(x,p) : \mathbb{R}^M \to T_pM(x)$ the orthogonal projection on $T_pM(x)$.

As usual in the Calculus of Variations, we want to make variations! To do so, we notice that, let us take a function $u \in \mathcal{A}_G$ and perturb it with with a field $\varphi \in C_c^\infty(\Omega; \mathbb{R}^M)$, *i.e.*, let us consider the function

$$v_\varepsilon := u + \varepsilon\varphi\,.$$

We need to impose $v_e \in \mathcal{A}_G$, that is $G(x, v_e(x)) = 0$ for all $x \in \Omega$. To ensure the validity of the constraint, the idea is to project $v_e(x)$ on $M(x)$. For this reason, if $\varphi(x)$ is orthogonal to $T_{u(x)}M(x)$, the corresponding variation will have no effect on the projected function. This induces us to consider only *tangential* variations, introduced as follows

Definition 4.10. Let $u \in \mathcal{A}_G$. We say that a vector field $\varphi \in C_c^\infty(\Omega; \mathbb{R}^M)$ is *tangential along* u if $\varphi(x) \in T_{u(x)}M(x)$ for all $x \in \Omega$.

By standard results in differential geometry, we know that, fixed $u \in \mathcal{A}_G$ and $\varphi \in C_c^\infty(\Omega; \mathbb{R}^M)$ tangential along u, there exist $\varepsilon_0 > 0$ and a function $g : (-\varepsilon_0, \varepsilon_0) \times \mathbb{R}^M \to \mathbb{R}^M$ such that

$$\psi(\varepsilon, x) := u(x) + \varepsilon\varphi(x) + g(\varepsilon, x)\,,$$

is such that $G\big(\psi(\varepsilon, x)\big) = 0$ for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$ and $x \in \Omega$. Moreover it also holds that

$$\lim_{\varepsilon \to 0} \frac{g(\varepsilon, x)}{\varepsilon} = 0\,,$$

for all $x \in \Omega$. This means that the linear approximation of $M(x)$ around $u(x)$ made by the tangent plane $T_{u(x)}M(x)$ can be made uniform for all $x \in \Omega$ (to be extremely precise, this is true if we consider $\bar{\Omega}$.).

So, let us now suppose that u is a weak local minimizer of $\mathcal{F}$ over $\mathcal{A}_G$. By considering the function

$$\Phi(\varepsilon) := \mathcal{F}\big(\Psi(\varepsilon, \dot{)}\big)\,,$$

we have that $\exists \Phi'(0) = 0$. By the same computations you are now master at using, we obtain

$$\int_\Omega L_f(u) \cdot \varphi \, \mathrm{d}x = 0\,,$$

for all $\varphi \in C_c^\infty(\Omega; \mathbb{R}^M)$ that are orthogonal along u. In order to deduce a differential constrain from the above integral one, we need a suitable version of the fundamental lemma.

LEMMA 4.11. *Let $h \in C^0(\Omega;\mathbb{R}^M)$ be such that*

$$\int_\Omega g \cdot \varphi = 0\,,$$

for all $\varphi \in C_c^\infty(\Omega;\mathbb{R}^M)$ orthogonal along u. Then $\Pi(x,u(x))g(x) = 0$ for all $x \in \Omega$, that is $g(x)$ is orthogonal to $T_{u(x)}M(x)$.

PROOF. We prove the result locally. Let us fix a point $(\bar{x}, u(\bar{x}))$ and consider a neighborhood of it, that can be taken of the form $B_\delta(\bar{x}) \times U$, where $U \subset \mathbb{R}^N$ is an open neighborhood of $u(\bar{x})$. Let us consider C^2 functions

$$\tau_1, \ldots, \tau_k, \nu_1 \ldots, \nu_{M-k} : \Omega \times \mathbb{R}^M \to \mathbb{R}^M\,,$$

such that $\tau_1, \ldots, \tau_k$ is an orthonormal basis of $T_{u(x)M(x)}$ and $\nu_1 \ldots, \nu_{M-k}$ is an orthonormal basis of its complement. Then we can write the function g as

$$g(x) = \sum_{i=1}^{M-r} a_i(x)\tau_i(x,u(x)) + \sum_{j=1}^{k} b_j(x)\nu_j(x,u(x))\,,$$

for suitable $a_1, \ldots, a_k, b_1, \ldots, b_{M-k}$ of class C^0, and

$$\varphi(x) = \sum_{i=1}^{M-k} \varphi_i(x)\tau_i(x,u(x))\,.$$

Our hypothesis implies that

$$\int_\Omega \sum_{i=1}^{M-k} a_i\varphi_i \, \mathrm{d}x = 0\,,$$

Since the π_i's can be arbitrary functions in $C_c^\infty(B_\delta(\bar{(}x)))$, we conclude by using the standard fundamental lemma. □

Applying the above result to our case we obtain

$$\Pi(x,u(x))L_f(u(x)) = 0\,,$$

for all $x \in \Omega$. In particular we obtain that there exists C^2 functions $\lambda_1, \ldots, \lambda_{M-k}$ such that

$$L_f(u(x)) - \sum_{i=1}^{M-k} \lambda_i(x)G_p^i(x,u(x)) = 0\,,$$

for all $x \in \Omega$.

REMARK 4.12. Basically we obtain the same result as we did for the Lagrange multipliers. The difference is that, in this case, the multipliers are functions. And this is due to the dependence of G through p (not by its dependence through x!).

CHAPTER 5

Second order necessary conditions

So far, we've seen only necessary conditions related to the nullity of the first variation. We now want to investigate higher order necessary conditions. We'll see three of these: non-negativity of the second variations, the Legendre-Hadamard condition for weak local minimizers and the Weierstrass condition for strong local minimizers. The first one will be proved in full generality, while for the other two we will specialize to the case of curves, *i.e.*, functions $u : [a, b] \to \mathbb{R}^M$. The reason is because, in this case, the computation are easier and so one can better grab the main idea. All the proofs adapts to the more general case, up to technicalities.

5.1. Non-negativity of the second variation

The first one doesn't come as a surprise. We already know that, in the finite dimensional case, a point $\bar{x} \in \mathbb{R}^N$ of local minimum for $g : \mathbb{R}^N \to \mathbb{R}$ is a critical and stable point, that is $\nabla g(\bar{x}) = 0$ and $D^2 g(\bar{x}) \geq 0$, where with this second writing we mean that

$$A\eta \cdot \eta \geq 0\,,$$

where A denotes a (symmetric) $M \times M$ matrix. What we claim is that the same holds true also for variational integrals.

THEOREM 5.1. *Let us consider a lagrangian* $f \in C^2\big(\Omega \times \mathbb{R}^M \times \mathbb{R}^{M\times N}\big)$ *and suppose* $u \in C^1\big(\Omega; \mathbb{R}^M\big)$ *is weak local minimizer. Then*

$$\partial^2 \mathcal{F}(u)[\varphi] \geq 0\,,$$

for all $\varphi \in C_c^\infty(\Omega; \mathbb{R}^M)$, *where we define*

$$\begin{aligned}\partial^2 \mathcal{F}(u)[\varphi] := \int_\Omega \Big(& f_{pp}\big(x, u(x), Du(x)\big)\varphi(x) \cdot \varphi(x) \\ & + 2 f_{p\xi}\big(x, u(x), Du(x)\big)\varphi(x) \cdot D\varphi(x) \\ & + f_{\xi\xi}\big(x, u(x), Du(x)\big) D\varphi(x) \cdot D\varphi(x)\Big)\,\mathrm{d}x\,.\end{aligned}$$

PROOF. Let us take $\varphi \in C_c^\infty(\Omega; \mathbb{R}^M)$ and consider the function

$$\Phi(\varepsilon) := \mathcal{F}(u + \varepsilon\varphi)\,.$$

Then, we know that $\Phi'(0) = 0$ and $\Phi''(0) \geq 0$. The assertion follows by the latter condition. □

REMARK 5.2. The preceding necessary condition says that $\varphi \equiv 0$ is a minimizer of the *accessory integral* $\partial^2 \mathcal{F}(u)$.

5.2. The Legendre-Hadamard necessary condition

The previous necessary condition has an integral form. We would like to derive from it a pointwise condition, as we did for obtaining the Euler-Lagrange equation. In order to do so, as anticipated in the preamble of this chapter, we will specialize to the case of curves. The meaning of the following result is that, among the three terms appearing in the second variation, the leading one is $f_{\xi\xi}$.

THEOREM 5.3 (Legendre-Hadamard condition). *Let us consider* $f \in C^2\big((a,b)\times\mathbb{R}^M\times\mathbb{R}^M\big)$ *and assume* $u \in C^1\big((a,b);\mathbb{R}^M\big)$ *to be weak local minimizer of the functional* $\mathcal{F}$*. Then*

$$f_{\xi\xi}\big(x,u(x),Du(x)\big) \geq 0\,, \qquad \textit{for } x \in [a,b]\,.$$

PROOF. Fix a point $\bar{x} \in (a,b)$ (if $\bar{x} \in \{a,b\}$ proceed similarly) and consider, for $k \in \mathbb{N}$, the function

$$\psi_k(x) := \begin{cases} kx - \frac{\bar{x}}{k} + \frac{1}{k^2} & x \in \left[\bar{x} - \frac{1}{k}, \bar{x}\right], \\ \\ -kx - \frac{\bar{x}}{k} - \frac{1}{k^2} & x \in \left[\bar{x}, \bar{x} + \frac{1}{k}\right], \\ \\ 0 & \text{otherwise}\ . \end{cases}$$

Take a vector $\eta \in \mathbb{R}^M$, and consider the function

$$\varphi_k(x) := \psi_k(x)\eta\,.$$

This is not an admissible test function, because we have three corners. But the result will simply follow by perform the same computations with φ_k *smoothed out.* So, in what follows, we will assume φ_k to be smooth.

By the previous proposition, we know that $\partial^2\mathcal{F}(u)[\varphi_k] \geq 0$. Since

$$\varphi_k'(x) = \psi_k'(x)\eta\,, \qquad \varphi_k''(x)0\psi_k''(x)\eta\,,$$

the above condition writes as

$$\begin{aligned} 0 \leq& \int_{\bar{x}-\frac{1}{k}}^{\bar{x}} f_{xx}\big(x,u(x),u'(x)\big)\eta\cdot\eta\big(\psi(x)\big)^2\,\mathrm{d}x - 2k\int_{\bar{x}-\frac{1}{k}}^{\bar{x}} f_{x\xi}\big(x,u(x),u'(x)\big)\psi_k(x)\,\mathrm{d}x \\ +& \quad k^2\int_{\bar{x}-\frac{1}{k}}^{\bar{x}} f_{\xi\xi}\big(x,u(x),u'(x)\big)\eta\cdot\eta\,\mathrm{d}x + \int_{\bar{x}}^{\bar{x}+\frac{1}{k}} f_{xx}\big(x,u(x),u'(x)\big)\eta\cdot\eta\big(\psi(x)\big)^2\,\mathrm{d}x \\ +& \quad 2k\int_{\bar{x}}^{\bar{x}+\frac{1}{k}} f_{x\xi}\big(x,u(x),u'(x)\big)\psi_k(x)\,\mathrm{d}x + k^2\int_{\bar{x}}^{\bar{x}+\frac{1}{k}} f_{\xi\xi}\big(x,u(x),u'(x)\big)\eta\cdot\eta\,\mathrm{d}x\,. \end{aligned}$$

Dividing the above expression by $\frac{k}{2}$ and sending $k \to \infty$, we are left with

$$\lim_{k\to 0}\frac{k}{2}\int_{\bar{x}-\frac{1}{k}}^{\bar{x}+\frac{1}{k}} f_{\xi\xi}\big(x,u(x),u'(x)\big)\eta\cdot\eta\,\mathrm{d}x \geq 0\,. \tag{5.1}$$

In order to conclude, we recall that, if $g : [a,b] \to \mathbb{R}$ is a continuous function, then for every $\bar{x} \in [a,b]$ and $\varepsilon > 0$, there exists $\delta > 0$ such that $|g(x) - g(\bar{x})| < \varepsilon$ if $|x - \bar{x}| < \delta$. Thus

$$2\delta(g(\bar{x}) - \varepsilon) \leq \int_{\bar{x}-\delta}^{\bar{x}+\delta} g(x)\,\mathrm{d}x \leq 2(\bar{x} + \varepsilon)\delta\,.$$

By recalling that $f_{\xi\xi}\big(x,u(x),u'(x)\big)$ is a continuous function, we obtain the desired condition.

$\square$

5.3. The Weierstrass necessary condition

In this section, by assuming u to be a *strong* local minimizer, we will be able to find a necessary condition of a *global* type, in the following sense: the preceding condition is local in the sense that it depends on the local behavior of the function

$$\xi \mapsto f\big(x, u(x), \xi\big) .$$

The condition we are going to prove takes into consideration the behavior of the above map in the whole space $\mathbb{R}^M$.

Proposition 5.4 (Weiersrtass condition). *Let $f \in C^2\big([a,b] \times \mathbb{R}^M \times \mathbb{R}^M\big)$ and assume $u \in C^1\big([a,b]; \mathbb{R}^M\big)$ to be strong local minimizer. Define the so called* excess *function as*

$$\mathcal{E}(x, p, \xi\eta) := f(x,p,\eta) - f(x,p,\xi) - f_\xi(x,p,\xi) \cdot (\eta - \xi) .$$

Then, it holds that

$$\mathcal{E}\big(x, u(x), u'(x), \eta\big) \geq 0 ,$$

for all $x \in [a,b]$ and all $\eta \in \mathbb{R}^M$.

Proof. Fix $\bar{x}, \bar{x}_1 \in (a,b)$ (as before, if $\bar{x} \in \{a,b\}$ just adapt the following computations) and $\eta \in \mathbb{R}^M$. Let us consider the linear function

$$w(x) := u(\bar{x}) + (x - \bar{x})\eta .$$

The idea is to consider a variation of u is a right-neighborhood of $\bar{x}$ with the function w. For, we attach w with u at the point $\bar{x}$, and then we re-glue them together. Formally, let us define the function $z \in C^1\big([a,b] \times [0,1]; \mathbb{R}^M\big)$ as

$$z(x,s) := \begin{cases} u(x) & x \in [a, \bar{x}] , \\ w(x) & x \in [\bar{x}, s] , \\ v(x,s) & x \in [s, \bar{x}_1] , \\ u(x) & x \in [\bar{x}_1, b] . \end{cases}$$

where $s \in [\bar{x}, \bar{x}_1]$, and we define

$$v(x,s) := u(x) + \frac{x - \bar{x}_1}{s - \bar{x}_1}(w(x) - u(x)) .$$

Notice that v is used to pass from w to u in $[s, \bar{x}_1]$, that is

$$v(s,s) = w(s) , \qquad v(\bar{x}_1, s) = u(\bar{x}_1) ,$$

Notice also that

$$\|u - z(\cdot, s)\|_{C^0} \to 0 ,$$

as $s \to \bar{x}^+$. This is the important point where we need u to be a *strong* local minimizer, since otherwise we cannot (and we don't want to) have any control on the derivative of $z(\cdot, s)$. So, by strong local minimality of u, we deduce that

$$\mathcal{F}(u) \leq \mathcal{F}\big(z(\cdot, s)\big) ,$$

for all $s \geq 0$. Thus,

$$\frac{\mathrm{d}}{\mathrm{d}s} \mathcal{F}\big(z(\cdot, s)\big)_{|s = \bar{x}} \geq 0 .$$

So, let us compute the above object

$$
\begin{aligned}
\mathcal{F}\big(z(\cdot,s)\big) &= \int_a^{\bar{x}} f\big(x,u(x),u'(x)\big)\,\mathrm{d}x + \int_{\bar{x}}^{s} f\big(x,u(x),\eta\big)\,\mathrm{d}x \\
&\quad + \int_s^{\bar{x}_1} f\Big(x, v(x,s), \frac{\partial}{\partial x}v(x,s)\Big)\,\mathrm{d}x + \int_{\bar{x}_1}^{b} f\big(x,u(x),u'(x)\big)\,\mathrm{d}x \\
&=: I_1 + I_2 + I_3 + I_4\,.
\end{aligned}
$$

We now compute each term of $\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{F}\big(z(\cdot,s)\big)_{|s=\bar{x}}$ saparately. We have that

$$
\frac{\mathrm{d}}{\mathrm{d}s} I_1\big(z(\cdot,s)\big)_{|s=\bar{x}} = 0\,,
$$

$$
\frac{\mathrm{d}}{\mathrm{d}s} I_4\big(z(\cdot,s)\big)_{|s=\bar{x}} = 0\,,
$$

$$
\frac{\mathrm{d}}{\mathrm{d}s} I_2\big(z(\cdot,s)\big)_{|s=\bar{x}} = f(\bar{x}, u(\bar{x},\eta))\,,
$$

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s} I_2\big(z(\cdot,s)\big)_{|s=\bar{x}} &= -f\Big(\bar{x}, v(\bar{x},\bar{x}), \frac{\partial}{\partial x}v(\bar{x},\bar{x})\Big) \\
&\quad + \int_{\bar{x}}^{\bar{x}_1} f_p\Big(x, v(x,\bar{x}), \frac{\partial}{\partial x}v(x,\bar{x})\Big)\frac{\partial}{\partial s}v(x,\bar{x})\,\mathrm{d}x \\
&\quad + \int_{\bar{x}}^{\bar{x}_1} f_\xi\Big(x, v(x,\bar{x}), \frac{\partial}{\partial x}v(x,\bar{x})\Big)\frac{\partial^2}{\partial x\partial s}v(x,\bar{x})\,\mathrm{d}x \\
&=: f\Big(\bar{x}, v(\bar{x},\bar{x}), \frac{\partial}{\partial x}v(\bar{x},\bar{x})\Big) + J_1 + J_2\,.
\end{aligned}
$$

By noticing that

$$
\frac{\partial}{\partial x}v(x,\bar{x}) = u'(\bar{x})\,,
$$

and by using the Euler-Lagrange equation to write

$$
\begin{aligned}
&\int_{\bar{x}}^{\bar{x}_1} f_p\Big(x, v(x,\bar{x}), \frac{\partial}{\partial x}v(x,\bar{x})\Big)\frac{\partial}{\partial s}v(x,\bar{x})\,\mathrm{d}x \\
&\qquad = \int_{\bar{x}}^{\bar{x}_1} \frac{\mathrm{d}}{\mathrm{d}x}\Big(f_\xi\Big(x, v(x,\bar{x}), \frac{\partial}{\partial x}v(x,\bar{x})\Big)\Big)\frac{\partial}{\partial s}v(x,\bar{x})\,\mathrm{d}x\,,
\end{aligned}
$$

we get

$$
\begin{aligned}
J_1 + J_2 &= \int_{\bar{x}}^{\bar{x}_1} \frac{\mathrm{d}}{\mathrm{d}x}\Big(f_\xi\big(x,u(x),u'(x)\big)\frac{\partial v}{\partial s}(x,s)\Big)\,\mathrm{d}x \\
&= f_\xi\big(\bar{x}_1, u(\bar{x}_1), u'(\bar{x}_1)\big)\frac{\partial v}{\partial s}(\bar{x}_1,\bar{x}) - f_\xi\big(\bar{x}, u(\bar{x}), u'(\bar{x})\big)\frac{\partial v}{\partial s}(\bar{x},\bar{x})
\end{aligned}
$$

Finally, we have that

$$
\frac{\partial v}{\partial s}(x,\bar{x}) = \frac{x-\bar{x}_1}{\bar{x}-\bar{x}_1}\big(\eta - u'(\bar{x})\big)\,,
$$

and hence

$$
\frac{\partial v}{\partial s}(\bar{x}_1,\bar{x}) = 0\,, \qquad \frac{\partial v}{\partial s}(\bar{x},\bar{x}) = \eta - u'(\bar{x})\,.
$$

The result then follows by putting together all the pieces. $\square$

REMARK 5.5. The above condition has a deep geometric meaning: it says that $u'(x)$ must be a point where the function $g(\eta) := f(x, u(e), \eta)$ coincides with its *convex envelope*, that is with the greatest convex function that lies below g. In this sense the above condition is a *global* condition, since it takes into consideration the global behavior of the function g.

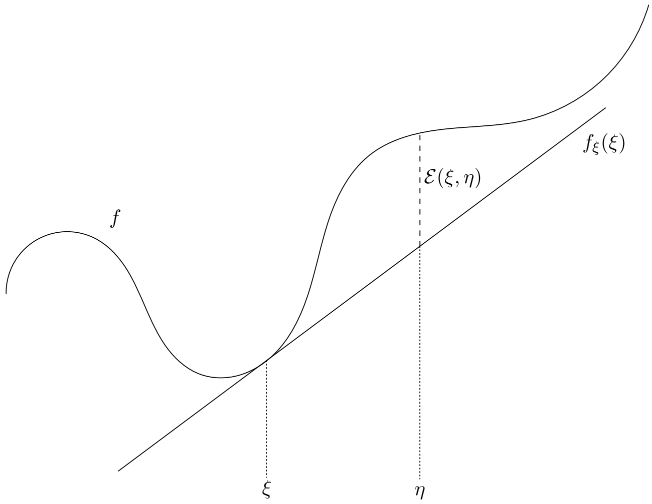

FIGURE 1. The geometrical meaning of the Weierstrass excess function for $f = f(\xi)$.

CHAPTER 6

Null lagrangians

In this chapter we would like to study lagrangians for which

$$L_f(u) = 0\,,$$

for all $u \in C^2(\Omega;\mathbb{R}^M)$ and functionals such that

$$\mathcal{F}(u) = \mathcal{F}(v)\,,$$

for all $u, v \in C^1(\Omega;\mathbb{R}^M)$ with $u_{|\partial\Omega} = v_{|\partial\Omega}$. This class of objects are interesting becase, in the first case, the Euler-Lagrange equation won't give us any information about minimum points of $\mathcal{F}$, while in the second case all points are minimum points (since the functional is constant!). This two pathologies are strictly connected and they will play a special role when we will develop the theory of sufficient conditions for strong local minimality.

Let us start by proving the connection between the two above phenomena.

Proposition 6.1. *The following are equivalent*

(i) $L_f(u) = 0$ *for all* $u \in C^1(\bar{\Omega};\mathbb{R}^M)$,
(ii) $\mathcal{F}(u) = \mathcal{F}(v)$ *for all* $u, v \in C^2(\Omega;\mathbb{R}^M)$ *with* $u_{|\partial\Omega} = v_{|\partial\Omega}$.

Proof. $(i) \Rightarrow (ii)$: let $u, v \in C^2(\Omega;\mathbb{R}^M)$ with $u_{|\partial\Omega} = v_{|\partial\Omega}$. Then

$$\begin{aligned}\mathcal{F}(v) - \mathcal{F}(u) &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}\big(u + t(v-u)\big)\,\mathrm{d}t \\ &= \int_0^1 L_f\big(u + t(v-u)\big) \times (v-u)\,\mathrm{d}t = 0\,.\end{aligned}$$

In the case $u, v \in C^1(\Omega;\mathbb{R}^M)$ the result follows by an approximation argument.
$(ii) \Rightarrow (i)$: trivial. □

Let us give some names.

Definition 6.2. Suppose $f \in C^2$ is such that one of the two above hold. Then we say that u is a *null-lagrangian.*

We now want to investigate the special structure of such a null-lagrangians.

Proposition 6.3. *The following are equivalent*

(i) f *is a null-lagrangian,*
(ii) $f\big(x, u(x), Du(x)\big) = D_\alpha\omega^\alpha\big(x, u(x), Du(x)\big)$, *for some functions* ω^α *of class* C^2.

Proof. $(i) \Rightarrow (ii)$: let us consider the function $g(\varepsilon) := f\big(x, \varepsilon u(x), \varepsilon Du(x)\big)$, for some fixed $x \in \Omega$. It holds that

$$f\big(x, u(x), Du(x)\big) - f(x, 0, 0) = g(1) - g(0)\,.$$

Since

$$g(1)-g(0)=\int_0^1 g'(t)\,\mathrm{d}t$$
$$=\int_0^1 \Big(f_p\big(x,\varepsilon u(x),\varepsilon Du(x)\big)\cdot u(x)+f_\xi\big(x,\varepsilon u(x),\varepsilon Du(x)\big)\cdot Du(x)\Big)\,\mathrm{d}t\,.$$

Now, notice that (let us be sloppy and forget about the arguments)

$$\mathrm{div}(u\cdot f_\xi)=f_\xi\cdot Du+u\cdot \mathrm{div} f_\xi\,.$$

So, we can continue the above computation by writing

$$g(1)-g(0)=\int_0^1 L_f(\varepsilon u)(x)\cdot u(x)\,\mathrm{d}t$$
$$+\int_0^1 \mathrm{div}\Big(u(x)\cdot f_\xi\big(x,\varepsilon u(x),\varepsilon Du(x)\big)\Big)\,\mathrm{d}t\,.$$

Notice the the first integral on the right-hand side is always zero. We can now give for grant the following fact: there exists $v\in C^3(\Omega;\mathbb{R}^M)$ such that $\mathrm{div}v(x)=f(x,0,0)$. If you believe this, let us just define the functions

$$\omega^\alpha(x,p,\xi):=v^\alpha(x)+\int_0^1 p^i\cdot f_{\xi^i_\alpha}(x,tp,t\xi)\,\mathrm{d}t\,.$$

$(ii)\Rightarrow(i)$: in this case we have that

$$\mathcal{F}(u)=\int_\Omega f\big(x,u(x),Du(x)\big)\,\mathrm{d}x=\int_\Omega D_\alpha\omega^\alpha\big(x,u(x),Du(x)\big)$$
$$=\int_{\partial\Omega}\omega^\alpha\big(x,u(x),Du(x)\big)\nu^\alpha(x)\,\mathrm{d}\sigma(x)\,,$$

where in the last step we used the Stokes theorem. In particular this means that

$$L_f(\varphi)=0\,,$$

for all $\varphi\in C^2_c(\Omega;\mathbb{R}^M)$. Thus f is a null-lagrangian. □

The previous result says that null-lagrangians are in a divergence form. It is worth noticing the following fact, that follows directly from the above proof.

Corollary 6.4. *Let f be a null-lagrangian. Then*

$$\int_{\partial\Omega}\omega^\alpha\big(x,u(x),Du(x)\big)\nu^\alpha(x)\,\mathrm{d}\sigma(x)=\int_{\partial\Omega}\omega^\alpha\big(x,v(x),Dv(x)\big)\nu^\alpha(x)\,\mathrm{d}\sigma(x)\,,$$

for all $u,v\in C^2(\Omega;\mathbb{R}^M)$ with $u_{|\partial\Omega}=v_{|\partial\Omega}$.

At a first sight it can seems surprising, since we are only asking u and v to agree on $\partial\Omega$, but we are asking nothing on the gradients of u and v on $\partial\Omega$. Since f depends also on the gradient of the function, we would expect the above result to holds true if we ask bot the functions and the gradients to coincide on $\partial\Omega$. But, for the case of null-lagrangians, we can only ask the functions to coincide on the boundary.

Finally, we would like to prove a more fine characterization of null-lagrangians in the one dimensional scalar case. For, we notice that, fixed a point $\bar{x}\in(a,b)$, and, for arbitrary $\alpha,\beta,\gamma\in\mathbb{R}$, let us consider the function

$$u(x):=\alpha+\beta(x-\bar{x})+\frac{\gamma}{2}(x-\bar{x})^2\,.$$

By hypothesis, we know that $L_f(u)(\bar{x}) = 0$. This means that

$$f_p\big(\bar{x}, u(\bar{x}), u'(\bar{x})\big) - f_{x\xi}\big(\bar{x}, u(\bar{x}), u'(\bar{x})\big) \\ - f_{p\xi}\big(\bar{x}, u(\bar{x}), u'(\bar{x})\big)\beta - f_{\xi\xi}\big(\bar{x}, u(\bar{x}), u'(\bar{x})\big)\gamma = 0\,,$$

that is

$$f_p\big(\bar{x}, \alpha, \beta\big) - f_{x\xi}\big(\bar{x}, \alpha, \beta\big) - f_{p\xi}\big(\bar{x}, \alpha, \beta\big)\beta - f_{\xi\xi}\big(\bar{x}, \alpha, \beta\big)\gamma = 0\,,$$

Since the above equations are true for every choice of α, β, γ, by using $\gamma = 0$, we get that f has to satisfy

$$\begin{cases} f_p\big(\bar{x}, \alpha, \beta\big) - f_{x\xi}\big(\bar{x}, \alpha, \beta\big) - f_{p\xi}\big(\bar{x}, \alpha, \beta\big)\beta = 0\,, \\ \\ f_{\xi\xi}\big(\bar{x}, \alpha, \beta\big) = 0\,, \end{cases}$$

for all $\alpha, \beta \in \mathbb{R}$. The second equation implies that

$$f(x, p, \xi) = A(x, p) + B(x, p)\xi\,.$$

By rewriting the first condition for such an f, we get

$$A_p(x, p) = B_x(x, p)\,,$$

In particular, this condition implies that

$$f\big(x, u(x), u'(x)\big) = A\big(x, u(x)\big) + B\big(x, u(x)\big)u'(x) = \Big(S\big(x, u(x)\big)\Big)'\,,$$

where S is such that $S_x(x, p) = A(x, p)$ and $S_p(x, p) = B(x, p)$. Thus, we've just proved the following

Proposition 6.5. *Let $f \in C^2\big((a, b) \times \mathbb{R} \times \mathbb{R}\big)$ be a null-lagrangian. Then*

$$f\big(x, u(x), u'(x)\big) = \Big(S\big(x, u(x)\big)\Big)'\,,$$

for some function S of class C^1.

In particular this means that null-lagrangians (in the case $N = M = 1$) do not depend explicitly by ξ. With this characterization in hand, we can see directly that

$$\mathcal{F}(u) = \int_a^b f\big(x, u(x), u'(x)\big)\,\mathrm{d}x = \int_a^b \Big(S\big(x, u(x)\big)\Big)' = S\big(b, u(b)\big) - S\big(a, u(a)\big)\,,$$

and thus $\mathcal{F}(u) = \mathcal{F}(v)$ for all $u, v \in C^1\big((a, b)\big)$ with the same boundary values.

Finally, we would like to mention that the same characterization of null-lagrangians holds true in the case $\min\{N, M\} = 1$, by using the same argument as above.

CHAPTER 7

Lagrange multipliers and eigenvalues of the Laplace operator

The motivation for our study comes from the following fact from linear algebra. Let us consider a (real) quadratic form Q on $\mathbb{R}^N$. We can write it as

$$Q(x) = \langle Ax, x \rangle \,,$$

where A is an $N \times N$ symmetric matrix and $\langle \cdot, \cdot \rangle$ denotes the scalar products on $\mathbb{R}^N$. Consider the following constrained minimization problem:

$$\min_{\|x\|^2=1} Q(x) \,.$$

Then, by the standard Lagrange multipliers rule we obtain that a critical points (that we know to exist thanks to the Weierstrass theorem: a continuous function defined on a compact set admits maximum and minimum) has to satisfy the equation

$$Ax = \lambda x \,, \tag{7.1}$$

for some $\lambda \in \mathbb{R}$. This is an eigenvalue problem for the linear operator defined by A. Since A is symmetric, we know that all its eigenvalues are real. In particular we know that the above equation admits a solution. For such a solution we have that

$$Q(x) = \langle Ax, x \rangle = \langle \lambda x, x \rangle = \lambda \|x\|^2 = \lambda \,.$$

Thus the problem of minimizing Q on the unit sphere reduces to finding the minimum eigenvalue of A. And this eigenvalue turns out to be the Lagrange multiplier given by the Lagrange multiplier rule. So, denoting by λ_1 the smaller eigenvalue of A, we have that

$$\lambda_1 = \min_{\|x\|^2=1} Q(x) \,.$$

We now want to get rid of the constrain. For, we notice that $Q(\mu x) = \mu^2 x$ for all $\mu \in \mathbb{R} \setminus \{0\}$, and thus

$$\frac{Q(\mu x)}{\|\mu x\|^2} = \frac{Q(x)}{\|x\|^2} \,.$$

We claim that:

$$\min_{\|x\|^2=1} Q(x) = \min_{0 \neq x \in \mathbb{R}^N} \frac{Q(x)}{\|x\|^2} \,.$$

Indeed we have that, by setting $g(x) := \frac{Q(x)}{\|x\|^2}$,

$$\nabla g(x) = \frac{2Ax}{\|x\|^2} - 2Q(x) \frac{x}{\|x\|^3} \,,$$

and thus $\nabla g(x) = 0$ if and only if x satisfies

$$Ax = \frac{Q(x)}{\|x\|} x \,,$$

that is x is an eigenvector of A. This gives the equivalence of the two minimum problems above. We can thus state the above result as follows:

$$\lambda_1 = \min_{0\neq x\in\mathbb{R}^N} \frac{Q(x)}{\|x\|^2}\,.$$

This is a variational characterization of the first eigenvalue. What for the other eigenvalues? Well, let us recall that if A is a symmetric operator and v is an eigenvector of A, then

$$\langle v, w\rangle = 0 \Rightarrow \langle Av, Aw\rangle = 0\,.$$

Indeed, by using the fact that $A^T = A$, we have that

$$\langle Av, Aw\rangle = \langle \lambda v, Aw\rangle = \lambda\langle A^T v, w\rangle = \lambda\langle Av, w\rangle = \lambda^2\langle v, w\rangle = 0\,.$$

In particular, this implies that we can find an orthogonal basis made by eigenvectors of A. Let $v_1, \ldots, v_N$ be such a basis relative to the eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$. Then the space[1]

$$v_1^\perp := \{\, v \in \mathbb{R}^N \;:\; \langle v, v_1\rangle = 0\,\}$$

is generated by $v_2, \ldots, v_N$. Thus, by applying the same argument as above, we get that

$$\lambda_2 = \min_{0\neq v\in v_1^\perp} \frac{\langle Av, v\rangle}{\|v\|^2}\,.$$

And so on:

$$\lambda_k = \min_{0\neq v\in S_{k-1}^\perp} \frac{\langle Av, v\rangle}{\|v\|^2}\,, \tag{7.2}$$

where we denote by S_k the space generated by $v_1, \ldots, v_k$, and $S_0 := \emptyset$ (where $\emptyset^\perp = \mathbb{R}^N$).

This is a variational characterization of all the eigenvalues of A.

But, you know, mathematicians are lazy, and they don't usually want to make computations! For this reason, the above characterization is not satisfactory, because, for computing the k^{th} eigenvalue, it requires to compute *explicitly* the first $k-1$ eigevectors. We would like to avoid this annoying computation. For, we notice that, fixed k, the eigenvalue λ_k is the highest eigenvalue of A restricted to the space S_k. Thus S_k is, among all spaces of dimension k, the *best* space, in the sense that it contains no directions with higher eigenvalues. Thus, we can say that

$$\lambda_k = \min_{\substack{V\subset\mathbb{R}^N\\ \dim(V)=k}} \max_{0\neq v\in V} \frac{\langle Av, v\rangle}{\|v\|^2}\,.$$

The above characterization is completely variational and it requires no knowledge of the first $k-1$ eigenvetors. This result is known as *Courant-Fisher theorem.*

The idea is to replicate all the above characterization in the case of the Laplacian,*i.e.*, we aim at obtaining a variational characterizaion as above of the values $\lambda \in \mathbb{R}$ for which there exists a solution of the Cauchy problem

$$\begin{cases} -\triangle u = \lambda u & \text{in } \Omega\,,\\ u = 0 \;\text{ on } \partial\Omega\,, \end{cases} \tag{7.3}$$

where $\Omega \subset \mathbb{R}^N$ is an open bounded set with smooth boundary $\partial\Omega$. The presentation we are going to provide is not self-contained, but relies on some results from functional analysis. In particular we will give for grant the following facts:

(i) the values λ for which the above minimum problem admits a solutions are a discrete set of $\mathbb{R}$, *i.e.*, are a sequence $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_k \leq \ldots$,

[1]To be very precise, we have to make the identification between the dual of $\mathbb{R}^N$ and $\mathbb{R}^N$ itself given by the scalar product. But you already know this technical detail from linear algebra.

(ii) the corresponding eigenfunctions (an eigenvector) are of class $C^2(\Omega)$.

But you can believe them! Just think to the one dimensional case, where $\Omega = (a,b)$: in this simple case, in order to obey the boundary condition $u(a) = u(b) = 0$, you know (since you can solve explicitly the problem!) that the λ's must be of the form $i\frac{\pi}{b-a}$, for some $i \in \mathbb{N}$.

First of all we need to make sense of all the objects we used in the finite dimensional case also in these context. Since it is not the aim of this chapter to treat formally the underlining spaces, we will be a bit sloppy, and we will just introduce them from an 'operative' point of view. We consider:

(i) a vector space
(ii) a scalar product on it (which will define a norm on the space)
(iii) a symmetric operator

We take:

(i) the space $C^2_0(\Omega)$,
(ii) as a scalar product on it we take
$$\langle v, w\rangle_{L^2} := \int_\Omega vw \,\mathrm{d}x\,.$$
If you want a reason to take it, is because it generates the norm
$$\|u\|^2_{L^2} := \langle v, v\rangle_{L^2} = \int_\Omega v^2 \,\mathrm{d}x\,,$$
that can be thought as a generalization of $\sum_{i=1}^N x_i^2$ when $N \to \infty$.
(iii) the symmetric operator is the laplacian $-\triangle$. First of all it is linear, since $-\triangle(v + w) = -\triangle v - \triangle w$. Let us check that it is symmetric:
$$\begin{aligned}\langle -\triangle v, w\rangle_{L^2} &= \int_\Omega (-\triangle v)w \,\mathrm{d}x = \int_\Omega \nabla v\cdot\nabla w \,\mathrm{d}x\\ &= \int_\Omega v(-\triangle w)\,\mathrm{d}x = \langle -\triangle w, v\rangle_{L^2}\,,\end{aligned}$$
where each time we integrate by parts we use the fact that we are using functions that are zero on $\partial\Omega$.

As before we consider the quadratic form
$$Q(v) = \langle -\triangle v, v\rangle_{L^2} = \int_\Omega |\nabla v|^2 \,\mathrm{d}x\,,$$
and the functional
$$\mathcal{G}(u) := \frac{\langle -\triangle u, u\rangle_{L^2}}{\|u\|^2_{L^2}} = \frac{\int_\Omega |\nabla u|^2 \,\mathrm{d}x}{\int_\Omega u^2 \,\mathrm{d}x}\,.$$
We want to compute the variational derivative of $\mathcal{G}$. For, let us take a function $\varphi \in C^\infty_c(\Omega)$ and compute the derivative of the function
$$\Phi(\varepsilon) := \mathcal{G}(u + \varepsilon u)\,,$$
at $\varepsilon = 0$. This turns out to be
$$\Big(\int_\Omega u^2 \,\mathrm{d}x\Big)\Big(\int_\Omega \nabla u\cdot\nabla\varphi \,\mathrm{d}x\Big) = \Big(\int_\Omega u\varphi \,\mathrm{d}x\Big)\Big(\int_\Omega |\nabla u|^2 \,\mathrm{d}x\Big)\,.$$
Thus, if u is a local minimizer for $\mathcal{G}$, it has to satisfy the equation
$$\int_\Omega \nabla u\cdot\nabla\varphi = \lambda\int_\Omega u\varphi \,\mathrm{d}x\,,$$

for all $\varphi \in C_c^\infty(\Omega)$, where $\lambda = \frac{\int_\Omega |\nabla u|^2\,\mathrm{d}x}{\int_\Omega u^2\,\mathrm{d}x}$. By integrating by parts the left-hand side we get

$$\int_\Omega (-\triangle u + \lambda u)\varphi\,\mathrm{d}x = 0\,,$$

for all $\varphi \in C_c^\infty(\Omega)$. Thus we get the equation

$$-\triangle u + \lambda u = 0\,.$$

As before we get that if u solves the above equation, then

$$\mathcal{G}(u) = \frac{\langle -\triangle u, u\rangle_{L^2}}{\|u\|_{L^2}^2} = \lambda\,.$$

So, as for the finite dimensional case, the problem of minimizing $\mathcal{G}$ (or equivalently, minimizing $u \mapsto \langle -\triangle u, u\rangle_{L^2}$ over the unit ball $\|u\|_{L^2}^2 = 1$) reduces to find the smallest eigenvalue for problem (7.3). Or, reading it in the opposite direction, we have the following variational characterization of the first eigenvalue of $-\triangle$:

$$\lambda_1 = \min_{0 \neq u \in C_0^2(\Omega)} \frac{\langle -\triangle u, u\rangle_{L^2}}{\|u\|_{L^2}^2}\,,$$

and the minimum is attained by an eigenfunction relative to λ_1. Let's prove a variational characterization for the k^{th} eigenvalue of (7.3) similar to (7.2).

Lemma 7.1. *Let us denote by S_{k-1} the space generated by the first $k-1$ eigenfunctions. Then*

$$\lambda_k = \min_{0 \neq u \in S_{k-1}^\perp} \frac{\int_\Omega |\nabla u|^2\,\mathrm{d}x}{\int_\Omega u^2\,\mathrm{d}x}\,,$$

and the minimun is attained by an eigenfunction relative to λ_k.

Proof. Let us take a test function $\varphi \in C_c^\infty(\Omega) cup S_{k-1}^\perp$. Then, by the same computation as above, we find that, if u_k is a minimum of the above functional, it has to satisfy the equation

$$\int_\Omega (-\triangle u + \lambda u)\varphi\,\mathrm{d}x = 0\,.$$

We want to say that the above equation is true for all $\varphi \in C_c^\infty(\Omega)$. For, given such a φ, we write it as

$$\varphi = \sum_{i=1}^{k-1} \langle \varphi, v_i\rangle_{L^2} v_i + \widetilde{\varphi}\,,$$

where $v_1, \cdot, v_{k-1}$ are the first $k-1$ eigenfunctions (that we suppose normalized, *i.e.*, with $\|v_i\|_{L^2} = 1$). Notice that, for all $j = 1, \dots, k-1$, it holds that

$$\langle \widetilde{\varphi}, v_j, \widetilde{\varphi}\rangle_{L^2} = \langle \varphi, v_j\rangle_{L^2} - \sum_{i=1}^{k-1} \langle \varphi, v_i\rangle_{L^2} \langle v_i, v_j\rangle_{L^2}\,.$$

By assuming that $\langle v_i, v_j\rangle_{L^2} = 0$ if $i \neq j$ (by using the standard Gran-Smith orthogonal procedure), we get $\langle \widetilde{\varphi}, v_j, \widetilde{\varphi}\rangle_{L^2} = 0$ for all $j = 1, \dots, k-1$. Moreover

$$\int_\Omega (-\triangle u + \lambda u) v_i\,\mathrm{d}x = \int_\Omega (u(-\triangle v_i) + \lambda u)\,\mathrm{d}x = (-\lambda_i + \lambda)\int_\Omega u v_i\,\mathrm{d}x = 0\,,$$

where in the last step we used the fact that u lies in the orthogonal space of $v_1, \dots, v_{k-1}$. So, we have that

$$\int_\Omega (-\triangle u + \lambda u)\varphi\,\mathrm{d}x = 0\,,$$

holds for all $\varphi \in c_c^\infty(\Omega)$, and thus, by the fundamental lemma we get the desired equation. The fact that the λ is the k^{th} eigenvalue follows immediately from the above equation and the fact that u lies in the orthogonal space of $v_1, \dots, v_{k-1}$. $\square$

As in the finite dimensional case, we would like to obtain the variational characterization without requiring the explicit knowledge of the eigenfunctions.

THEOREM 7.2. *It holds*

$$\lambda_k = \min_{\substack{V \subset C_0^2(\Omega) \\ \dim(V)=k}} \max_{0\neq v\in V} \frac{\int_\Omega |\nabla v|^2 \, \mathrm{d}x}{\int_\Omega v^2 \, \mathrm{d}x} .$$

PROOF. Let us denote by

$$\lambda_M(V) := \max_{0\neq u\in V} \frac{\int_\Omega |\nabla u|^2 \, \mathrm{d}x}{\int_\Omega u^2 \, \mathrm{d}x} .$$

Step 1: first of all we prove that $\lambda_k \leq \lambda_M(V)$, for all $V \subset C_0^2(\Omega)$ with $\dim(V) = k$. Let $w_1, \dots, w_k$ be a set of linear independent functions in V such that

$$\langle w_i, v_j \rangle_{L^2} = 0 ,$$

for all $i = 1, \dots, k$ and $j = 1, \dots, k-1$. Notice that this is possible because we have $k-1$ linear equations and k unknowns. Then, if $w \in V$, the above conditions implies that

$$\langle w, v_i \rangle_{L^2} = 0 ,$$

for all $i = 1, \dots, k-1$, and thus, by the previous lemma, we have that

$$\lambda_k \leq \frac{\int_\Omega |\nabla w|^2 \, \mathrm{d}x}{\int_\Omega w^2 \, \mathrm{d}x} \leq \max_{0\neq v\in V} \frac{\int_\Omega |\nabla v|^2 \, \mathrm{d}x}{\int_\Omega v^2 \, \mathrm{d}x} .$$

Step 2: let us now prove the reverse inequality. Let us consider the vector space $V \subset C_0^2(\Omega)$ generated by the first k eigenfunctions $v_1, \dots, v_k$. Let $v \in V$, and write

$$v = \sum_{i=1}^k c_i v_i ,$$

where $c_i = \langle v, v_i \rangle_{L^2}$. Then

$$\int_\Omega v^2 \, \mathrm{d}x = \sum_{i,j=1}^k c_i c_j \int_\Omega v_i v_j \, \mathrm{d}x = \sum_{i=1}^k c_i^2 ,$$

and

$$\begin{aligned} \int_\Omega |\nabla v|^2 \, \mathrm{d}x &= \int_\Omega \Big(\sum_{i=1}^k c_i \nabla v_i\Big) \cdot \Big(\sum_{j=1}^k c_j \nabla v_j\Big) \, \mathrm{d}x \\ &= \int_\Omega -\Big(\sum_{i=1}^k c_i \lambda_i v_i\Big)\Big(\sum_{j=1}^k c_j v_j\Big) \, \mathrm{d}x = \sum_{i=1}^k c_i^2 \lambda_i , \end{aligned}$$

where in the previous to last step we used integration by parts. Thus

$$\frac{\int_\Omega |\nabla v|^2 \, \mathrm{d}x}{\int_\Omega u^2 \, \mathrm{d}x} = \frac{\sum_{i=1}^k \lambda_i c_i^2}{\sum_{i=1}^k c_i^2} \leq \frac{\sum_{i=1}^k \lambda_k c_i^2}{\sum_{i=1}^k c_i^2} = \lambda_k .$$

$\square$

The above variational characterization has an important consequence. In order to underline the dependence of the eigenvalue from the domain, we write $\lambda_k(\Omega)$.

COROLLARY 7.3. *Let* $\Omega_1 \subset \Omega_2$. *Then* $\lambda_k(\Omega_1) \geq \lambda_k(\Omega_1)$ *for all* $k \in \mathbb{N}$.

PROOF. The result simply follows by noticing that $C_0^2(\Omega_1) \subset C_0^2(\Omega_2)$, and thus, by the variational characterizaion of each eigenvalue, when considering the eigenvalues for Ω_2 we are considering a bigger set of competitors, and thus the minimum will diminish. □

In particular it can be proved that

COROLLARY 7.4. *It holds*

(i) $\lambda_k \to \infty$ *as* $\lambda \to \infty$
(ii) $\lambda_k(B_r) \to \infty$ *as* $r \to 0$.

An important consequence of the variational characterization of the eigenvalues of the laplacian is that every element of $C_0^2(\Omega)$ can be approximated (with respect to the norm $\|\cdot\|_{L^2}$) with elements in the space generated by the eigenfunctions of the laplacian. The statement is something like: *you can approximate every real number with a sequence of rational numbers*, but it is amazing if you think about it, since here each real number plays the role of a new dimension. Indeed the space $C_0^2(\Omega)$ is an infinite dimensional vector space whose *'cardinality of its infinite dimensionality'* is the one of the continuum. But we can basically just work with a countable number of object, and obtain all of them by approximantion.

THEOREM 7.5. *Let* $(v_i)_{i\in\mathbb{N}}$ *be the sequence of orthonormal eigenfuntions for* $-\triangle$ *in* Ω. *Then, for every element* $v \in C_0^2(\Omega)$ *it holds*

$$r_n := v - \sum_{i=1}^n v_i \langle v, v_i \rangle_{L^2} \to 0\,, \qquad in\ L^2\,,$$

i.e.,

$$\|r_n\|_{L^2} \to 0\,.$$

PROOF. First of all we notice that

$$\langle r_n, v_i \rangle_{L^2} = 0\,,$$

for all $i = 1, \ldots, n$. Thus, by the variational characterization proved above, we know that

$$\lambda_k \leq \frac{\int_\Omega |\nabla r_n|^2\,\mathrm{d}x}{\int_\Omega r_n^2\,\mathrm{d}x}$$

from which we deduce that

$$\int_\Omega r_n^2\,\mathrm{d}x \leq \frac{\int_\Omega |\nabla r_n|^2\,\mathrm{d}x}{\lambda_k}\,.$$

The idea is to prove that the numerator is uniformly bounded for all $n \in \mathbb{N}$. For, we notice that

$$\begin{aligned}
\int_\Omega |\nabla r_n|^2 \, \mathrm{d}x &= \int_\Omega \Big| \nabla v - \sum_{i=1}^n c_i \nabla v_i \Big| \, \mathrm{d}x \\
&= \int_\Omega |\nabla v|^2 \, \mathrm{d}x + \Big| \sum_{i=1}^n c_i \nabla v_i \Big|^2 - 2 \sum_{i=1}^n c_i \int_\Omega \nabla v \cdot \nabla v_i \, \mathrm{d}x \\
&= \int_\Omega |\nabla v|^2 \, \mathrm{d}x + \sum_{i=1}^n c_i^2 \lambda_i - 2 \sum_{i=1}^n c_i^2 \lambda_i \\
&= \int_\Omega |\nabla v|^2 \, \mathrm{d}x - \sum_{i=1}^n c_i^2 \lambda_i \,,
\end{aligned}$$

where in the previous last step we used the fact that

$$\int_\Omega \nabla v_i \cdot \nabla v_j \, \mathrm{d}x = \lambda_i \int_\Omega v_i v_j \,.$$

Since $\lambda_i \geq 0$, we obtain that

$$\int_\Omega |\nabla r_n|^2 \, \mathrm{d}x \leq \int_\Omega |\nabla v|^2 \, \mathrm{d}x \,.$$

By using this uniform bound and the fact that $\lambda_k \to \infty$ as $k \to \infty$, we obtain the desired result. $\square$

CHAPTER 8

Sufficient conditions

In this chapter we would like to address the following question: which conditions do we have to add to the first and second order necessary conditions we derived in the previous chapters in order to ensure weak or strong local minimality?[1]

8.1. Coercivity of the second variation

We start with the investigation of the second variation of $\mathcal{F}$. We know that, in the finite dimensional case, if we have $g \in C^2(\mathbb{R}^N)$ and a point $\bar{x} \in \mathbb{R}^N$ such that

$$\nabla g(\bar{x}) = 0\,, \qquad D^2 g(\bar{x}) > 0\,,$$

then $\bar{x}$ turns out to be an isolated local minimizer of g. Thus, one can guess that the same holds true also for functionals of the calculus of variations. The answer is negative, as will be shown in the following example due to Sheefer.

Example: let us consider the following functional

$$\mathcal{F}(u) := \int_{-1}^{1} \left(x^2 \big(u'(x)\big)^2 + x\big(u'(x)\big)^3 \right) \mathrm{d}x\,,$$

defined over $u \in C_0^1\big((-1,1)\big)$. Then we have that

$$L_f(0) = 0\,, \qquad \delta^2\mathcal{F}(0)[\varphi] = \int_{-1}^{1} 2x^2 \big(\varphi'(x)\big)^2 \,\mathrm{d}x > 0\,,$$

for all $\varphi \in C_c^\infty((-1,1)) \setminus \{0\}$. If we now consider the sequence of functions $(u_k)_k$ defined as

$$u_k := \begin{cases} \frac{4}{3k}\left(x + \frac{1}{k}\right) & x \in \left[-\frac{1}{k}, 0\right], \\ -\frac{4}{3k}\left(x - \frac{1}{k}\right) & x \in \left[0, \frac{1}{k}\right], \\ 0 & \text{otherwise}\,, \end{cases}$$

we have that

$$\mathcal{F}(u_k) = -\frac{32}{27k^5}\,.$$

By smoothing out the corners we obtain a sequence $(v_k)_k$ such that $v_k \to 0$ in C^1 and $\mathcal{F}(u_k) < 0$. This implies that 0 cannot be a weak local minimizer.

[1]Notice that all the conditions we will present do not treat *global* minimality properties, but just local ones. In order to obtain global minimality results, more involved arguments are needed. But it is not a surprise, since, as for the finite dimensional case, these techniques are based on properties of 'second derivatives', that have, by definition, a *local* nature.

In order to understand which can of condition we need to ask in order to obtain some local minimality property, let us notice the following: let us consider a critical point $u \in C^2(\Omega;\mathbb{R}^M)$ for $\mathcal{F}$ such that

$$\delta^2\mathcal{F}(u)[\varphi] > 0\,,$$

for every $\varphi \in C_c^\infty(\Omega;\mathbb{R}^M)$. Actually we know that the above condition holds for all $\varphi \in C_0^2(\Omega;\mathbb{R}^M)$ (since this is the bigger space of admissible variations). Let $v \in C^2(\Omega;\mathbb{R}^M)$ with the same boundary value of u. Thus, by defining the function $g(\varepsilon) := \mathcal{F}\big(u+\varepsilon(v-u)\big)$, we have that

$$g'(0) = \delta\mathcal{F}(u)[v-u] = 0\,,$$

and hence we can write

$$\mathcal{F}(v) - \mathcal{F}(u) = g(1) - g(0) = \int_0^1 (1-t)g''(t)\,\mathrm{d}t\,.$$

Noticing that $v-u \in C_0^2(\Omega;\mathbb{R}^M)$ and that

$$g''(t) = \delta^2\mathcal{F}\big(u+t(v-u)\big)[v-u]\,,$$

we deduce that if we are able to prove that $\delta^2\mathcal{F}(w)[\varphi] > 0$ for all $\varphi \in C_0^2(\Omega;\mathbb{R}^M)$ and for all $w \in C^2(\Omega;\mathbb{R}^M)$ sufficiently close to u, we obtain the desired result. By noticing that

$$g''(0) = \delta^2\mathcal{F}(u)[v-u] > 0\,,$$

what we are asking for is a sort of continuity of the second variation. Clearly, such a property cannot be true for the functional of the previous example. But what is the reason why it cannot be true? The problem is that we are in an infinite dimensional space (C_c^∞...yes, it is infinite dimensional!) and we can have a degeneracy of the second variation. To explain better this concept, let us notice that, since is it quadratic, it holds that

$$\delta^2\mathcal{F}(u)[s\varphi] = s^2\delta^2\mathcal{F}(u)[\varphi]\,, \tag{8.1}$$

for every $s \in \mathbb{R}$. Now, given a function $\varphi \in C_c^\infty((a,b))$, we notice that in the second variation we are only interested in the behavior of

$$\int_a^b \big(\varphi(x)\big)^2\,\mathrm{d}x\,, \qquad \int_a^b \big(\varphi'(x)\big)^2\,\mathrm{d}x\,.$$

In particular we can define a *norm* of a function $\varphi \in C_c^\infty((a,b))$ as

$$\|\varphi\|_{H^1} := \left(\int_a^b \big(\varphi(x)\big)^2\,\mathrm{d}x + \int_a^b \big(\varphi'(x)\big)^2\,\mathrm{d}x\right)^{\frac{1}{2}}\,.$$

Thus, by (8.1), we have that we only have to control what happen for functions $\varphi \in C_c^\infty((a,b))$ with $\|\varphi\|_{H^1} = 1$. Now, if we were in finite dimension our test functions would be just vectors and our ball would be the standard Euclidean ball. If we had a continuous function $g : \mathbb{R}^N \to \mathbb{R}$ such that $g(x) > 0$ for all $\|x\| = 1$, that we could conclude that

$$\inf_{\|x\|=1} g(x) > 0\,. \tag{8.2}$$

This is exactly what goes wrong in the infinite dimensional case. Indeed, in our previous example, we have that if we consider as a test functions

$$\varphi_k := \frac{u_k}{\|u_k\|}\,,$$

we have

$$\partial^2\mathcal{F}(0)[\varphi_k] \to 0\,.$$

Let us look closely to this phenomena in comparison with the finite dimensional case. In the latter one the argument that allows us to conclude (8.2) is the following. Let $(x_n)_n$ be a minimizing sequence for g on the unit ball, *i.e.*, $\|x_n\| = 1$ and $g(n_n) \to \inf_{\|x\|=1} g(x)$. Since the unit ball is *compact*, up to a subsequence we can assume that $x_n \to \bar{x}$, where $\|\bar{x}\| = 1$. Then, by continuity of g, we would have $g(\bar{x}) = 0$, that is in contradiction with our initial hypothesis. The last step of the above argument brakes in the infinite dimensional case: indeed the unit ball of any infinite dimensional normed vector space turns out to be *not* compact and thus, in particular, we cannot extract the (norm)-converging subsequence as above. We cannot enter too much into these details, but just keep in mind that this is a peculiarity of infinite dimensional normed vector spaces that led to the introduction of the so called *weak convergence* and to a whole brunch of functional analysis.

Let us come back to our case. So, if we had a condition like

$$\inf_{\|\varphi\|_{H^1}=1} \partial^2\mathcal{F}(u)[\varphi] = \lambda > 0\,,$$

we can expect that, by continuity of the coefficients of the integrands of the second variation, the same property holds true also for $\partial^2\mathcal{F}(v)$, where v is sufficiently close to u in C^1. Luckily, this turns out to be true!

THEOREM 8.1. *Let f be a lagrangian of class C^2 and let $u \in C^2((a,b))$ be a critical point such that*

$$\inf_{\|\varphi\|_{H^1}=1} \partial^2\mathcal{F}(u)[\varphi] = \lambda > 0\,,$$

that is

$$\partial^2\mathcal{F}(u)[\varphi] \geq \lambda \int_a^b \big(\varphi(x)\big)^2 + \big(\varphi'(x)\big)^2\,\mathrm{d}x\,,$$

for all test functions $\varphi \in C_c^\infty((a,b))$. Then u is a weak local minimizer. In particular, u turns out to be an isolated weak local minimizer, and the following holds

$$\mathcal{F}(u) + \frac{\lambda}{4}\|u-v\|_{H^1}^2 \leq \mathcal{F}(v)\,,$$

for all $\|v-u\|_{C^1} < r$, for some $r > 0$.

PROOF. Let us define

$$\begin{aligned}
a(x,u) &:= f_{pp}\big(x, u(x), u'(x)\big)\,,\\
b(x,u) &:= f_{p\xi}\big(x, u(x), u'(x)\big)\,,\\
c(x,u) &:= f_{\xi\xi}\big(x, u(x), u'(x)\big)\,.
\end{aligned}$$

By continuity of all the second partial derivatives of f, we have that, fixed $\varepsilon > 0$ there exists $r > 0$ such that

$$\begin{aligned}
|a(x,u) - a(x,v)| &< \varepsilon\,,\\
|b(x,u) - a(x,v)| &< \varepsilon\,,\\
|c(x,u) - a(x,v)| &< \varepsilon\,,
\end{aligned}$$

for all $\|v-u\|_{C^1} < r$. So, we can estimate

$$\begin{aligned}
\Big|\partial^2\mathcal{F}(u)[\varphi] - \partial\mathcal{F}(v)[\varphi]\Big| &\leq \varepsilon \int_a^b \Big(\big(\varphi'(x)\big)^2 + 2|\varphi(x)||\varphi'(x)| + \big(\big(\varphi(x)\big)^2\Big)\,\mathrm{d}x\\
&= \varepsilon \int_a^b \big(\varphi(x) + \varphi'(x)\big)^2\,\mathrm{d}x\\
&\leq 2\varepsilon \int_a^b \Big(\big(\varphi'(x)\big)^2 + \big(\big(\varphi(x)\big)^2\Big)\,\mathrm{d}x\,,
\end{aligned}$$

where in the last step we have used the inequality $(a+b)^2 \leq 2(a^2+b^2)$. By choosing $\varepsilon > 0$ such that $2\varepsilon < \frac{\lambda}{2}$, we obtain that

$$\begin{aligned}\partial^2\mathcal{F}(v)[\varphi] &\geq \partial^2\mathcal{F}(u)[\varphi] - \Big|\partial^2\mathcal{F}(u)[\varphi] - \partial\mathcal{F}(v)[\varphi]\Big| \\ &\geq \frac{\lambda}{2}\int_a^b \left(\big(\varphi(x)\big)^2 + \big(\varphi'(x)\big)^2\right)\,\mathrm{d}x = \frac{\lambda}{2}\|\varphi\|^2_{H^1}\,.\end{aligned}$$

So, by defining $g(t) := \mathcal{F}\big(u + t(v-u)\big)$ we have (see argument above)

$$\begin{aligned}\mathcal{F}(v) - \mathcal{F}(u) = g(1) - g(0) &= \int_0^1 (1-t)g''(t)\,\mathrm{d}t \\ &= \int_0^1 (1-t)\partial^2\mathcal{F}\big(u + t(v-u)\big)[v-u]\,\mathrm{d}t \\ &\geq \frac{\lambda}{2}\|v-u\|^2_{H^1}\int_0^1 (1-t)\,\mathrm{d}t = \frac{\lambda}{4}\|v-u\|^2_{H^1}\,.\end{aligned}$$

□

Remark 8.2. Notice that the above argument cannot be applied to obtain *strong* local minimality, since we really need a control on the derivative of the function to estimates the second order derivatives of f. Unless, of course, they do not depend on the derivative of the function!

Moreover, if you read carefully the proof, you may notice that $v-u$ is just in $C^2_0((a,b))$ and not necessarily in $C^\infty_c((a,b))$. But it is possible to prove that an estimate from below of the second variation when applied to functions $\varphi \in C^\infty_c((a,b))$ implies a similar estimate for function $\varphi \in C^2_0((a,b))$.

Finally, the similar argument applies in the case $u \in C^2(\Omega;\mathbb{R}^M)$, but with just a lot of indexes!

8.2. Jacobi conjugate points

We now want to understand when it is possible to have an estimate from below of the second variation like the one in the hypothesis of the theorem. In this section we will consider only functions $u \in C^2((x_1,x_2))$. Fix a critical point $u \in C^2((x_1,x_2))$ and define

$$a(\cdot) := a(\cdot, u(\cdot))\,, \quad b(\cdot) := b(\cdot, u(\cdot))\,, \quad c(\cdot) := c(\cdot, u(\cdot))\,.$$

We notice that from the Legendre-Hadamard condition, if u is a weak local minimizer, then we have $a \geq 0$ in (a,b). Recall that we aim at proving the inequality

$$G(\varphi) := Q(\varphi) - \lambda\|\varphi\|^2_{H^1} \geq 0\,,$$

where $Q(\varphi) := \partial^2\mathcal{F}(u)[\varphi]$ is called the *accessory Lagrangian*. Notice that $\varphi \equiv 0$ is a minimizer for G and thus, by the Legendre-Hadamard condition applied to G, we obtain $a \geq \lambda > 0$. So, we can assume $a > 0$, *i.e.*, we have the *strict Legendre condition* in force.

The idea of Jacobi and Legendre was to complete the square of the accessory lagrangian by adding a null-lagrangian. We recall that, in our case, null-lagrangians are function of the form

$$\big(g(x,u(x))\big)'\,.$$

By choosing $g(x,p) := p^2 w(x)$, we obtain the null-lagrangian

$$G(x,p,\xi) := 2\xi p w(x) + p^2 w'(x)\,.$$

So

$$Q(\varphi) = Q(\varphi) + \int_{x_1}^{x_2} G\big(x, \varphi(x), \varphi'(x)\big)\, \mathrm{d}x$$
$$= \int_{x_1}^{x_2} \big(a(\varphi')^2 + 2(b+w)\varphi\varphi' + (c+w')(\varphi)^2 \big)\, \mathrm{d}x\,.$$

To complete the square we have to require that

$$\big(b+w\big)^2 - a(c+w') = 0\,. \tag{8.3}$$

If the above equation is in force, we have that

$$Q(\varphi) = \int_{x_1}^{x_2} a\Big(\varphi' + \frac{b+w}{a}\varphi\Big)^2\, \mathrm{d}x\,. \tag{8.4}$$

Let us take a closer look to equation (8.3). It's called *Legendre equation* and can be written as

$$w' = \frac{2b}{a}w + \Big(\frac{b^2}{a} - c\Big) + \frac{w^2}{a}\,. \tag{8.5}$$

The equation falls into the category of the so called Riccati equation. We see that we have a non linearity in w (the unknown) of order 2. The idea is to make a (clever) substitution in order to reduce it to something we are more comfortable with. Let us introduce the function $v > 0$ by

$$w = -a - b\frac{v'}{v}\,.$$

Then, from (8.5), we obtain that v must satisfy the equation

$$-(av')' + (c - b')v = 0\,. \tag{8.6}$$

This second order linear equation is called *Jacobi equation*, and we call a solution $v > 0$ a *Jacobi field*. Let us now suppose that there exists a Jacobi field v (recall that $v > 0$ on $[a, b]$). Then, from (8.4), we can write

$$Q(\varphi) = \int_{x_1}^{x_2} av^2\bigg(\Big(\frac{\varphi}{y}\Big)'\bigg)^2\, \mathrm{d}x\,.$$

In order to conclude we need a technical result.

Lemma 8.3 (Poincaré inequality). *Let $g \in C^1([x_1, x_2])$ such that $g(x_1) = 0$. Then*

$$\int_{x_1}^{x_2} \big(g(x)\big)^2\, \mathrm{d}x \le (x_2 - x_1)^2 \int_{x_1}^{x_2} \big(g'(x)\big)^2\, \mathrm{d}x\,.$$

Proof. Since $g(x_1) = 0$ we can write

$$g(x) = \int_{x_1}^{x} g'(t)\, \mathrm{d}t\,,$$

for $x \in [x_1, x_2]$. Now

$$\big(g(x)\big)^2 = \bigg(\int_{x_1}^{x} g'(t)\, \mathrm{d}t\bigg)^2 = (x_2 - x_1)^2\bigg(\frac{1}{(x_2 - x_1)}\int_{x_1}^{x} g'(t)\, \mathrm{d}t\bigg)^2$$
$$\le (x_2 - x_1)\int_{x_1}^{x} \big(g'(t)\big)^2\, \mathrm{d}t \le (x_2 - x_1)\int_{x_1}^{x_2} \big(g'(t)\big)^2\, \mathrm{d}t\,,$$

where in the previous to last step we used the Jensen inequality (see 2.16). By integrating both sides we obtain the desired result. □

We are now able to conclude:

$$
\begin{aligned}
Q(\varphi) &= \int_{x_1}^{x_2} av^2 \left(\left(\frac{\varphi}{y} \right)' \right)^2 \, \mathrm{d}x \\
&\geq \inf_{[x_1,x_2]} (av^2) \int_{x_1}^{x_2} \left(\left(\frac{\varphi}{y} \right)' \right)^2 \, \mathrm{d}x \\
&\geq \inf_{[x_1,x_2]} (av^2) \frac{1}{(x_2 - x_1)^2} \int_{x_1}^{x_2} \left(\frac{\varphi}{y} \right)^2 \, \mathrm{d}x \\
&\geq \inf_{[x_1,x_2]} (av^2) \frac{1}{(x_2 - x_1)^2} \inf_{[x_1,x_2]} \frac{1}{v^2} \int_{x_1}^{x_2} \varphi^2 \, \mathrm{d}x =: \mu \int_{x_1}^{x_2} \varphi^2 \, \mathrm{d}x \,,
\end{aligned}
$$

that is

$$
Q(\varphi) \geq \mu \int_{x_1}^{x_2} \varphi^2 \, \mathrm{d}x \,. \tag{8.7}
$$

By setting

$$
\alpha := \inf_{[x_1,x_2]} a > 0 \,, \quad \beta := \sup_{[x_1,x_2]} |b| \,, \quad \gamma := \sup_{[x_1,x_2]} |c| \,,
$$

we can perform the following estimate

$$
\begin{aligned}
Q(\varphi) &\geq \alpha \int_{x_1}^{x_2} (\varphi')^2 \, \mathrm{d}x - 2\beta \int_{x_1}^{x_2} |\varphi||\varphi'| \, \mathrm{d}x + \gamma \int_{x_1}^{x_2} (\varphi)^2 \, \mathrm{d}x \\
&\geq Q(\varphi) + \varepsilon \int_{x_1}^{x_2} (\varphi')^2 \, \mathrm{d}x + (\frac{\beta^2}{\varepsilon} + \gamma) \int_{x_1}^{x_2} (\varphi)^2 \, \mathrm{d}x \,,
\end{aligned}
$$

where in the last step we used the inequality

$$
2xy \leq \frac{1}{\varepsilon} x^2 + \varepsilon y^2 \,.
$$

By choosing $\varepsilon = \frac{\alpha}{2}$, we get

$$
\frac{\alpha}{2} \int_{x_1}^{x_2} (\varphi')^2 \, \mathrm{d}x \leq \Big(\frac{2\beta^2}{\varepsilon} + \gamma \Big) \int_{x_1}^{x_2} (\varphi)^2 \, \mathrm{d}x + Q(\varphi) \leq \Big[1 + \frac{1}{\mu} \Big(\frac{2\beta^2}{\varepsilon} + \gamma \Big) \Big] Q(\varphi) \,,
$$

where in the last step we used estimate (8.7).

Thus we've just proved the following result

THEOREM 8.4. *Let f be a lagrangian of class C^2 and let $u \in C^2((x_1, x_2))$ be a critical point. Suppose there exists a Jacobi vector field (relatively to u) on $[x_1, x_2]$. Then u is a strict weak local minimizer.*

8.3. Conjugate points and a necessary condition for weak local minimality

The above result gives us a sufficient condition for obtaining a weak minimality property of a critical point. We would like to understand whether this is just an *ad hoc* hypothesis or if it is something related to some kind of necessary condition. For, we need to study more in details equation (8.6). By setting

$$
p := \frac{a'}{a} \,, \qquad p := \frac{b' - c}{a} \,,
$$

we can rewrite it as

$$
v'' + pv' + qv = 0 \,. \tag{8.8}
$$

Since we assume $a > 0$ in $[x_1, x_2]$, we can also assume $a > 0$ in $I_0 := (x_1 - \delta, x_2 - \delta)$ for some small $\delta > 0$. Take a point $x_0 \in I_0$ and let us consider the following Cauchy problem

$$\begin{cases} v'' + pv' + qv = 0\,, \\ v(x_0) = \alpha\,, \\ v'(x_0) = \beta\,, \end{cases}$$

where $\alpha, \beta \in \mathbb{R}$. By the theory of ODE (read as: *you should know it!*) we know that there exists a unique solution of the above Cauchy problem, that we will denote by $\omega(x : x_0, \alpha, \beta)$. We also know that the space of solution of (8.8) turns out to be a vector space of dimension 2. In particular, if v_1, v_2 are linear independent solutions[2] then every solution v of (8.8) can be written as $\mu_1 v_1 + \mu_2 v_2$, for some $\mu_1, \mu_2 \in \mathbb{R}$,and that the *Wronskian*

$$W(x) := \det \begin{pmatrix} v_1(x) & v_2(x) \\ v_1'(x) & v_2'(x) \end{pmatrix} = v_1(x)v_2'(x) - v_2(x)v_1'(x)\,,$$

is always different from zero. In particular we obtain that if $v \not\equiv 0$ is a solution of (8.8), then the zeros of v are isolated. Indeed from $W(x) \neq 0$ we get that $v_1(x) = 0$ implies $v_1'(x) \neq 0$, and the result follows by writing v as a non trivial linear combination of v_1 and v_2. Finally we recall that[3]

$$a(x)W(x) \equiv C\,. \tag{8.9}$$

We now want to prove a version of the *Sturm's oscillation theorem.*

THEOREM 8.5. *Let v_1, v_2 are linear independent solutions of (8.8). The between two consecutive zeros of v_1 there exists one and only one zero of v_2.*

PROOF. Since $W(x) \neq 0$ and it is continuous, we can assume $W > 0$ in I_0. Let ξ_1, ξ_2 be two consecutive zeros of v_1. This implies that $v_1'(\xi_1)$ and $v_1'(\xi_2)$ have opposite signs. Then, by the conditions $W(\xi_1) > 0$ and $W(\xi_2) > 0$ we get that $v_2(\xi_1)$ and $v_2(\xi_2)$ must have opposite signs (and cannot be zero, otherwise W would be zero in some point). Since v_2 is continuous, we obtain that there exists at least one zero of v_2 in the interval (ξ_1, ξ_2). The uniqueness follows by arguing by absurd and by applying the above argument with v_1 and v_2 flipped. ☐

DEFINITION 8.6. Let us define

$$\triangle(x, x_0) := \omega(x; x_0, 0, 1)\,.$$

We will call the isolated zeros of $\triangle(\cdot, x_0)$ *conjugates points* to x_0.

LEMMA 8.7. *If there is no conjugate point to x_1 in $(x_, x_2]$, then there is no pair of conjugate points $\xi_1, \xi_2 \in (x_1, x_2]$.*

PROOF. By hypothesis we have that the function $\triangle(\cdot; x_1)$ never vanishes on $(x_1, x_2]$. Let us take any point $\xi \in (x_1, x_2]$ and consider the function $\triangle(\cdot; \xi)$. We claim that $\triangle(\,\cdot\,; x_1)$ and $\triangle(\,\cdot\,; \xi)$ are linear independent. Indeed, if $\mu_1 \triangle(\,\cdot\,; x_1) + \mu_2 \triangle(\,\cdot\,; \xi) = 0$, then from $0 = \mu_1 \triangle(\xi; x_1) + \mu_2 \triangle(\xi; \xi) = \mu_1 \triangle(\xi; x_1)$ we would obtain $\mu_1 = 0$ and consequently $\mu_2 = 0$. Now, by applying Sturm's oscillation theorem to $\triangle(\,\cdot\,; x_1)$ and $\triangle(\,\cdot\,; \xi)$, if there were a point $c \in [x_1, x_2]$ with $c \neq \xi$ and $\triangle(c; \xi) = 0$, we would obtain the existence of a zero of $\triangle(\,\cdot\,; x_1)$ in $(x_1, x_2]$. But this is impossible. ☐

[2]This means that if $\mu_1 v_1(x) + \mu_2 v_2(x) = 0$ for $x \in I_0$, then $\mu_1 = \mu_2 = 0$.

[3]To prove the claim, differentiate the function $x \mapsto a(x)W(x)$ and use equation (8.8).

Before proving the main result relating conjugate points and weak local minimality, we need to observe that the Legendre equation is the Euler-Lagrange equation of the accessory lagrangian q. Indeed, by recalling that

$$q(x,p,\xi) = a\xi^2 + 2bp\xi + cp^2\,,$$

we get that its Euler-Lagrange equation is

$$-(2a\xi)' + 2(v\xi + cp)\,,$$

and with some algebra we obtain that the above equation is equation (8.8).

We are now in position to prove a theorem regarding necessity and sufficiency of a Jacobi field (a positive solution to (8.6)) for weak local minimality.

THEOREM 8.8. *Let u be a critical point and suppose $a > 0$ in I_0. Then:*

(i) if there are no conjugate value to x_1 in $(x_1, x_2]$, then u is an isolated weak local minimizer,

(ii) if there exists $\xi \in (x_1, x_2)$ conjugated to x_1, then u is not *a weak local minimizer,*

(iii) if the first conjugate value to x_1 is x_2 then nothing can be said.

PROOF. (i): let us consider the functions $\triangle(\,\cdot\,;x_1)$ and $\triangle(\,\cdot\,;x_2)$. Since there are no conjugate values to x_1 in $(x_1,x_2]$ we know that $\triangle(\cdot;x_1) \neq 0$. We can assume $\triangle(\,\cdot\,;x_1) > 0$. we claim that $\triangle(\,\cdot\,;x_2) \neq 0$ in $[x_0, x_2)$ for some $x_0 \in (x_1 - \delta, x_1)$. Indeed, if not, we would have a point $\widetilde{x} \in [x_1,x_2)$ for which $\triangle(\widetilde{x};x_2) = 0$, and thus $\widetilde{x}$ would be conjugate to x_2. Let us suppose that $\widetilde{x}$ is the greater conjugate point of x_2. Then, since $\triangle(\,\cdot\,;x_1)$ and $\triangle(\,\cdot\,;x_2)$ are linear independent, from the Sturm's oscillation theorem, we would obtain that there exists a zero of $\triangle(\,\cdot\,;x_1)$ in the interval $(\widetilde{x}, x_2)$. But this is impossible because there are no conjugate values to x_1 in that interval.

Let us now take $\triangle(\,\cdot\,;x_2)$ and $\triangle(\,\cdot\,;x_0)$. They are linear independent, since $\triangle(x_0;x_2) \neq 0$. Thus we conclude that $\triangle(\,\cdot\,;x_0)$ doesn't change sign in $(x_0,x_2]$. And so we have found a Jacobi field ($\triangle(\,\cdot\,;x_0)$ or $-\triangle(\,\cdot\,;x_0)$).

(ii): the aim is to construct a test function $\varphi \in C_c^\infty\big((x_1,x_2)\big)$ such that

$$\partial^2\mathcal{F}(u)[\varphi] < 0\,.$$

Let $\xi \in (x_1,x_2)$ be the smallest value such that $\triangle(\xi;x1) = 0$, and let $\beta > \xi$ such that $\triangle(\,\cdot\,;x_1) \neq 0$ in $(\xi,\beta]$. Consider the function

$$v_2(x) := -\triangle(\,\cdot\,;\beta)\,.$$

Then $v_1(x) := \triangle(\,\cdot\,;x_1)$ and v_2 are linear independent (since $v_1(\beta) \neq 0$) and thus, by Sturm's oscillation theorem, there exists a zero $\alpha \in (x_1,\xi)$ of v_2. In particular that is the only zero of v_2 in the interval (x_1,β) (otherwise v_1 would have another zero in that interval). By noticing that both v_1 and v_2 are positive in (α,ξ), we get that there exists $\gamma \in (\alpha,\xi)$ such that $v_1(\gamma) = v_2(\gamma)$.

Notice that $v_2'(\beta) = -1$ implies that $v_2 < 0$ in $[x_1,\alpha)$ and $v_2 > 0$ in (α,β). Recalling equation (8.9), we have that

$$a(x)W(x) = a(x_1)W(x_1) = -a(x_1)v_2(x_1)v_1'(x_1) = -a(x_1)v_2(x_1) < 0\,. \tag{8.10}$$

Now, define the function

$$\varphi(x) := \begin{cases} v_1 & x \in [x_1,\gamma]\,, \\ v_2 & x \in [\gamma,\beta]\,, \\ 0 & x \in [\beta,x_2]\,, \end{cases}$$

Notice that φ is not a correct test function because it has two possible corners at γ and at β and it has no compact support. But it will be possible to approximate such a function with admissible test functions. The idea to conclude is the following: first of all we notice that we can write

$$Q(\varphi) = \int_{x_1}^{x_2} q\big(x, \varphi(x), \varphi'(x)\big)\,\mathrm{d}x\,,$$

where $q(x,p,\xi)$ is quadratic in (p,ξ) for all x. But we know that, for a quadratic form, it holds that

$$2q(x,p,\xi) = q_p(x,p,\xi)p + q_\xi(x,p,\xi)\xi\,,$$

and hence

$$\begin{aligned}
2Q(\varphi) &= \int_{x_1}^{x_2} \big[\, q_p\big(x,\varphi(x),\varphi'(x)\big)\varphi(x) + q_\xi\big(x,\varphi(x),\varphi'(x)\big)\varphi'(x)\,\big]\,\mathrm{d}x \\
&= \int_{x_1}^{x_2} L_q(\varphi)\varphi\,\mathrm{d}x + \varphi(\cdot)q_\xi\big(\cdot,\varphi(\cdot),\varphi'(\cdot)\big)\Big|_{x_1}^{x_2}\,.
\end{aligned}$$

So, by using the computations above, and recalling that (since they are Jacobi fields) $L_q(v_1) = L_q(v_2) = 0$, we have

$$\begin{aligned}
2Q(\varphi) &= \int_{x_1}^{\gamma} \big[\, q_p\big(x,\varphi(x),\varphi'(x)\big)\varphi(x) + q_\xi\big(x,\varphi(x),\varphi'(x)\big)\varphi'(x)\,\big]\,\mathrm{d}x \\
&\quad + \int_{\gamma}^{\beta} \big[\, q_p\big(x,\varphi(x),\varphi'(x)\big)\varphi(x) + q_\xi\big(x,\varphi(x),\varphi'(x)\big)\varphi'(x)\,\big]\,\mathrm{d}x \\
&= a(\gamma)\big(\, v_1(\gamma)v_1'(\gamma) + v_2(\gamma)v_2'(\gamma)\,\big) \\
&= a(\gamma)\big(\, v_2(\gamma)v_1'(\gamma) + v_1(\gamma)v_2'(\gamma)\,\big) \\
&= a(\gamma)W(\gamma) < 0\,,
\end{aligned}$$

where in the previous to last equality we used the fact that $v_1(\gamma) = v_2(\gamma)$, while the last inequality follows by (8.10).

(iii): in this case there are example in both directions, even if usually u turns out to be *not* a weak local minimizer. □

We prove that *critical points are local minimizers in small.*

Corollary 8.9. *Let u be a critical point and suppose that*

$$a(x_1) = f_{\xi\xi}\big(x_1, u(x_1), u'(x_1)\big) > 0\,.$$

Then there exists $\delta > 0$ such that u is an isolated weak local minimizer of the functional

$$\widetilde{\mathcal{F}}(v) := \int_{x_1}^{x_1+\delta} f\big(x, v(x), v'(x)\big)\,\mathrm{d}x\,.$$

Proof. We simply notice that we can choose $\delta > 0$ small enough such that $a(x) > 0$ in $[x_1, x_1+\delta]$ and there are no conjugate points to x_1 in $(x_1, x_1+\delta)$. This last condition can be obtained, since the zeros of the solutions of (8.8) are isolated. □

8.4. Geometric interpretation of conjugate points

We now want to understand the geometric nature of conjugate points. Fix $u \in C^3\big((x_1,x_2)\big)$ and $\varphi \in C^2\big((x_1,x_2)\big)$. Let us consider a general family of variations $\big(\Psi(t,\cdot)\big)_t$ such that

$$\Psi(0,\cdot) = u(\cdot)\,, \qquad \Psi_t(0,\cdot) = \varphi(\cdot)\,. \tag{8.11}$$

If we compute the Euler operator at $\Psi(t,\cdot)$ (*i.e.*, the Euler Lagrange equation of $\mathcal{F}$ at $\Psi(t,\cdot)$), we get

$$L_f\big(\Psi(t,\cdot)\big)(x) = f_p\big(x,\Psi(t,x),\Psi'(t,x)\big) - \Big(f_\xi\big(x,\Psi(t,x),\Psi'(t,x)\big)\Big)'.$$

We want to compute the derivative of the above expression with respect to the parameter t, that is

$$\begin{aligned}\frac{\partial}{\partial t}L_f\big(\Psi(t,\cdot)\big)(x) &= f_pp\big(x,\Psi(t,x),\Psi'(t,x)\big)\Psi_t(t,x) + f_{p\xi}\big(x,\Psi(t,x),\Psi'(t,x)\big)\Psi_t'(t,x)\\ &\quad - \Big(f_{p,\xi}\big(x,\Psi(t,x),\Psi'(t,x)\big)\Psi_t(t,x) + f_{\xi\xi}\big(x,\Psi(t,x),\Psi'(t,x)\big)\Psi_t'(t,x)\Big)',\end{aligned}$$

where in the last step we used the fact that, since f is of class C^3, the derivative with respect to x and to t commute. Then, by using conditions (8.11), we get

$$\begin{aligned}\Big(\frac{\partial}{\partial t}L_f\big(\Psi(t,\cdot)\big)\Big)_{|t=0}(x) &= f_{pp}\big(x,u(x),u'(x)\big)\varphi(x) + f_{p\xi}\big(x,u(x),u'(x)\big)\varphi'(x)\\ &\quad - \frac{\mathrm{d}}{\mathrm{d}x}\big(\, f_{\xi p}\big(x,u(x),u'(x)\big)\varphi(x) + f_{\xi\xi}\big(x,u(x),u'(x)\big)\varphi'(x)\,\big)\\ &= -\big(a(x)\varphi(x)\big)' + \big(c(x)-b'(x)\big)\varphi(x)\\ &= L_q(\varphi)\,.\end{aligned}$$

Denoting with $J_u(\varphi)$ the operator on the left-hand side, *i.e.*, the linearization of the Euler-Lagrange equation at u, we get that

Proposition 8.10. *$J_u(\varphi) = L_q(\varphi)$ for all $\varphi \in C^2\big((x_1,x_2)\big)$.*

The above observation leads us to the following

Lemma 8.11. *Let $\big(\Psi(t,\cdot)\big)_t$ be a family of solution of the Euler-Lagrange equation. Set $\varphi(\cdot) := \Psi_t(0,\cdot)$. Then*

$$J_u(\varphi) = 0\,,$$

that is, φ is a solution of the Legendre equation.

In order to give the desired geometric interpretation of the conjugate points, let us consider a family of critical points as in the above lemma, with the property that

$$\Psi(t,x_1) = z_1\,, \qquad \forall |t| < t_0\,,$$

for some $t_0 > 0$. Then we get that $\varphi(x_1) = 0$. Moreover, assume that $\Psi_{tx}(0,\cdot)$ is not identically zero in (x_1,x_2). In particular, since this implies that $\varphi(0,\cdot) \neq 0$, we get that $\varphi(0,x_1) \neq 0$. Now suppose to have x_1^*, the first conjugate point to x_1 in (x_1,x_2), that is $\varphi(x_1^*) = 0$ (since it is a Jacobi field, up to the multiplication to a constant, in order to have $\varphi'(0,x_1) = 1$). Recall this implies that $\varphi'(x_1^*) \neq 0$.

Let us now consider the following system of equations

$$\begin{cases}\Psi(x,t) - y = 0\,,\\ \Psi_t(x,t) = 0\,,\end{cases} \tag{8.12}$$

in the unknown (x, y, t). We know that a solution is given by the triple $(x_1^*, u(x_1^*), 0)$. Since we know that

$$0 \neq \varphi'(x_1^*) = \Psi(x_1^*, 0)\,,$$

we can apply the implicit function theorem to parametrize locally the set of solution of the above equation. In particular, we obtain a C^2 function $g : (-\widetilde{t}, \widetilde{t}) \to (x_1, x_2)$ with $g(0) = x_1^*$, such that

$$\Psi_t\big(g(t), t\big) = 0\,,$$

for all $|t| < \widetilde{t}$. Moreover, by setting

$$h(t) := \Psi\big(g(t), t\big)\,,$$

we have that

$$\{\, \big(g(t), h(t), t\big) \ : \ |t| < \widetilde{t}\,\}\,,$$

is a parametrization of the set of solutions of (8.12) around the point $(x_1^*, u(x_1^*), 0)$.

Let us now consider the set

$$\mathcal{E} := \{\, P(t) := \big(g(t), h(t)\big) \ : \ |t| < \widetilde{t}\,\}\,,$$

Let us suppose that $\Psi_{tt}\big(x_1^*, 0\big) \neq 0$. Then, up to taking a smaller $\widetilde{t}$, we obtain that the set $\mathcal{E}$ is of class C^2 and the point $P(t)$ is conjugate to x_1 along the critical point $\Psi(\cdot, t)$. Moreover, the set $\mathcal{E}$ is tangential to the curve $\Psi(\cdot, t)$ at the point $P(t)$, as can be seen from the definition of the function h. For this reason, the set $\mathcal{E}$ is the *envelope* of the family $\big(\Psi(\cdot, t)\big)_t$.

In the case $\Psi_{tt}\big(x_1^*, 0\big) = 0$ the set $\mathcal{E}$ can be degenerate. In particular it can be a single point (and in this case it is called a *nodal point*) or a cusp of $\mathcal{E}$. And it is in this last case that we have example where the first conjugate point to x_1 is x_2 and u is still a weak local minimizer.

8.5. Eigenvalues method for multiple integrals

The preceding theory of conjugate points applies to one dimensional scalar problems. We would like to understand if it is possible to develop a similar theory of sufficient conditions for weak local minimality also in the general case. The answer is affirmative, but the presentation won't be self-contained, because it relies on the theory of elliptic operators. So, we will just present the general ideas and state the most important results, without proof. Nevertheless, the reader will find everything we are going to say very likely to be true, according to the similarities with what we've seen in Chapter 7.

First of all we observe that, by performing the same computation as the ones in the beginning of the preceding section, we can prove that the Euler-Lagrange equation of the accessory lagrangian is the linearization of the Euler operator of the original lagrangian.

LEMMA 8.12. *$J_u(\varphi) = L_q(\varphi)$ for all $\varphi \in C^2(\Omega; \mathbb{R}^M)$.*

We now notice that the estimate we want to be true is the following

$$Q(\varphi) \geq \lambda \|\varphi\|^2_{H^1}\,.$$

And we want it to be true for some $\lambda > 0$. Clearly (by recalling that $Q(s\varphi) = s^2\varphi)$ this is possible if and only if

$$\min_{\|\varphi\|^2_{H^1}=1} Q(\varphi) > 0\,.$$

Thus, we are leading to consider the problem of minimizing the quadratic form Q among[4] all functions $\varphi \in C_0^2(\Omega;\mathbb{R}^M)$ such that $\|\varphi\|_{H^1}^2 = 1$. The minimum problem can be recast as following

$$\min_{\varphi \in C^2(\Omega;\mathbb{R}^M) \varphi \neq 0} \frac{Q(\varphi)}{\|\varphi\|_{H^1}^2}\,. \tag{8.13}$$

When we treated a similar problem in Chapter 7, we ended by studying the eigenvalues for the *derivative* of the quadratic form $Q(\varphi) := \int_\Omega |\nabla\varphi|^2\,\mathrm{d}x$. We will do the same also in this case. Notice that, if $\varphi \in C_0^2(\Omega;\mathbb{R}^M)$, then, by integrating by parts, we get that

$$Q(\varphi) = \frac{1}{2}\int_\Omega J_u(\varphi)\cdot\varphi\,\mathrm{d}x\,.$$

Thus, if we compute the Euler-Lagrange equation of (8.13), we obtain that a minimum has to satisfy

$$J_u(\varphi) = \lambda\varphi\,,$$

for some $\lambda \in \mathbb{R}$. Thus, we are led to consider the following eigenvalue problem

$$\begin{cases} J_u(\varphi) = \lambda\varphi & \text{in } \Omega\,, \\ \varphi = 0 & \text{on } \partial\Omega\,, \end{cases} \tag{8.14}$$

where $\varphi \in C^2(\bar\Omega;\mathbb{R}^M)$.

The basic theory for such a problem is summarized in the following

PROPOSITION 8.13. *Let us consider the problem (8.14) and assume $\partial\Omega$ is of class C^3. Then there exists a sequence of numbers $(\lambda_k)_k$ and a sequence of functions $(\varphi_k)_k \in C^2(\bar\Omega;\mathbb{R}^M)$ such that*

$$\begin{cases} J_u(\varphi_k) = \lambda_k\varphi_k & \text{in } \Omega\,, \\ \varphi_k = 0 & \text{on } \partial\Omega\,. \end{cases}$$

Moreover $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_k \le \ldots$ and $\lambda_k \to \infty$ as $k \to \infty$.

Moreover, we also have a monotonicity property for the eigenvalues.

LEMMA 8.14. *Let $\Omega_1 \subset \Omega_2$. Then $\lambda_k(\Omega_1) \ge \lambda_k(\Omega_2)$ for all $k \in \mathbb{N}$.*

By using the assumption that math is not such a mess, we can believe that it is possible to obtain a characterization of the eigenvalues of the operator J_u (by the way, recall that we are *always* assuming the strict Legendre-Hadamard condition!) similar to the one obtained in Chapter 7. In particular it holds that

$$\lambda_1 = \min\{Q(\varphi) \,:\, \varphi \in C_0^2(\Omega;\mathbb{R}^M)\,,\ \|\varphi\|_{H^1}^2 = 1\}\,.$$

This allows us to state the following result

THEOREM 8.15. *The following hold true:*

- *if $\lambda_1 > 0$, then u is an isolated weak local minimizer,*
- *if $\lambda_1 < 0$, then u is* not *an isolated weak local minimizer.*

[4] To be fair, this point is not completely correct, since the minimization should be take place in a larger space that C^2. But, for our purposes, we can ignore this fact, since it will be useful only for the existence of such a minimum. But you already know that the classical theory is *optimistic*, since we always give for grant the existence of a minimizer!

The above theorem is the analogous of Theorem 8.8. Indeed, let us consider in the one dimensional scalar case: take a Jacobi field on $\Omega := (x_1, x_2)$ and that there exists a conjugate value $x_1^* \in (x_1, x_2)$ to x_1. This means that $\lambda = 0$ is an eigenvalue of the Jacobi operator J_u in (x_1, x_1^*). What we proved in Theorem 8.8 is that, in this case, we have a negative eigenvalue of the Jacobi operator. The idea is to use Lemma 8.14 to obtain the strict inequality for λ_1. If no conjugate values to x_1 are present in $(x_1, x_2]$, then the result follows immediately by the variational characterization of the eigenvalues.

Finally, the analogous of Corollary 8.9, stating the minimality in small of critical points, is the following

COROLLARY 8.16. *Let $u \in C^2(\Omega; \mathbb{R}^M)$ be a critical point and suppose that*

$$f_{\xi^i_\alpha \xi^j_\beta}(\bar{x}, u(\bar{x}), Du(\bar{x}))\eta^i\eta^j\tau^\alpha\tau^\beta \geq \lambda|\eta|^2|\tau|^2\,,$$

for some $\lambda > 0$ and some $\bar{x} \in \Omega$. Then there exists $R > 0$ such that u is an isolated weak local minimizer of the functional

$$\widetilde{\mathcal{F}}(v) := \int_{B_R(\bar{x})} f\big(x, v(x), Dv(x)\big)\,\mathrm{d}x\,,$$

among all functions $v \in C^1(\Omega; \mathbb{R}^M)$ with $v_{|\partial B_R(\bar{x})} = u_{|\partial B_R(\bar{x})}$.

8.6. Weierstrass field theory

So far we develop sufficient conditions for having *weak* local minimality. In this section we would like to provide sufficient conditions ensuring a *strong* local minimality property of critical points. The theory we will develop will be restricted to the most simple case, that is the one dimensional scalar case. This is in order to catch the main ideas without being lost in the technical difficulties of the most general case (to be precise, the case is curves is not that difficult, but the most general case it is!).

In this section we'll do the opposite of what we've done so far: we'll start from the technical construction and only after having proved the big result, we'll come back and give an heuristic explanation of what we've done and why exactly in this way.

DEFINITION 8.17. A function $\varphi : [a, b] \times [\bar{\alpha}_1, \bar{\alpha}_2] \to \mathbb{R}^2$ given by

$$\varphi(x, \alpha) = \big(x, u_\alpha(x)\big)\,,$$

is called a *field of extremal* if

(i) φ is a C^2 diffeomorphism from $[a, b] \times [\bar{\alpha}_1, \bar{\alpha}_2]$ to its image, that we denote by G,
(ii) $u_\alpha(\cdot)$ is an extremal for all α.

In particular, it holds that

- the set G is *simply connected*,
- for each point $(x, p) \in G$ there exists a unique value of the parameter α, denoted by $\alpha^{-1}(x, p)$, such that $p = u_{\alpha^{-1}(x,p)}(x)$.

Of particular significance is the slope of the curve $u_\alpha(\cdot)$.

DEFINITION 8.18. Let h be a field of extremal. We define the so called *slope function* $P : G \to \mathbb{R}$ by

$$P\big(x, p\big) := u'_{\alpha^{-1}(x,p)}(x)\,.$$

That is, $P(x, p)$ is the slope at x of the extremal passing through (x, p).

Notation: in the following we will use the following notation

$$\bar f(x,p) := f\big(x,p,P(x,p)\big)\,.$$

REMARK 8.19. It is trivial to observe that the only solution of

$$\begin{cases} w'(x) = P\big(x,w(x)\big)\,, \\ w(\bar x) = \bar p\,, \end{cases} \tag{8.15}$$

is the function $u_{\alpha^{-1}(\bar x,\bar p)}$.

By using the fact that each u_α is a solution of the Euler-Lagrange equation, we obtain the following result.

PROPOSITION 8.20. *Let h be a field of extremal. Then*

$$\frac{\partial}{\partial p}\big[\,\bar f(x,p) - P(x,p)\bar f_\xi(x,p)\,\big] = \frac{\partial}{\partial x}\bar f_\xi(x,p)\,,$$

for all $(x,p)\in G$.

PROOF. By recalling that $\bar f(x,p) := f\big(x,p,P(x,p)\big)$, we have that

$$\begin{aligned}\frac{\partial}{\partial p}\big[\,\bar f(x,p) - P(x,p)\bar f_\xi(x,p)\,\big] &= \bar f_p(x,p) - \bar f_{\xi p}(x,p)P(x,p) \\ &\quad - \bar f_{\xi\xi}(x,p)P(x,p)P_p(x,p)\,,\end{aligned} \tag{8.16}$$

and

$$\frac{\partial}{\partial x}\bar f_\xi(x,p) = \bar f_{x\xi}(x,p) + \bar f_{\xi\xi}(x,p)P_x(x,p)\,. \tag{8.17}$$

Since every function u_α satisfies the Euler-Lagrange equation, we have that

$$\begin{aligned} f_p\big(x,u_\alpha(x),u'_\alpha(x)\big) &= f_{x\xi}\big(x,u_\alpha(x),u'_\alpha(x)\big) + f_{p\xi}\big(x,u_\alpha(x),u'_\alpha(x)\big)u'_\alpha(x) \\ &\quad + f_{\xi\xi}\big(x,u_\alpha(x),u'_\alpha(x)\big)u''_\alpha(x)\,.\end{aligned} \tag{8.18}$$

Now, choose a point $(\bar x,\bar p)\in G$ and consider the function $u_{\alpha^{-1}(\bar x,\bar p)}$. The idea is to compute the Euler-Lagrange equation for the function $u_{\alpha^{-1}(\bar x,\bar p)}$ at the point $(\bar x,\bar p)\in G$. In order to deal with the term $u''_\alpha(x)$, we recall that the observation made in Remark 8.19 allows us to say that

$$u''_\alpha(x) = P_x\big(x,u_\alpha(x)\big) + P_p\big(x,u_\alpha(x)\big)u'_\alpha(x)\,.$$

In particular, by writing (8.18) for the function $u_{\alpha^{-1}(\bar x,\bar p)}$, and computing it at the point $(\bar x,\bar p)\in G$, we get

$$\begin{aligned}\bar f_p(x,p) - \bar f_{\xi p}(\bar x,\bar p))P(\bar x,\bar p)) &= \bar f_{x\xi}(\bar x,\bar p)) + \bar f_{p\xi}(\bar x,\bar p))P(\bar x,\bar p)) \\ &\quad + \bar f_{\xi\xi}(\bar x,\bar p))\Big(\,P_x(\bar x,\bar p)) + P_p(\bar x,\bar p))P(\bar x,\bar p))\,\Big)\,,\end{aligned}$$

that is the desired equality. $\square$

What the above result is telling us is that the differential form ω defined on G as

$$\omega(x,p) := \big(\bar f(x,p) - P(x,p)\bar f_\xi(x,p)\big)\,\mathrm{d}x + \bar f_\xi(x,p)\,\mathrm{d}p\,,$$

is closed. Since the set G is simply connected, by the Poincaré lemma (see Appendix Section 11.6) we get that ω is exact, that is, there exists a potential $S : G\to\mathbb{R}$ such that

$$\frac{\partial}{\partial x}S(x,p) = \bar f(x,p) - P(x,p)\bar f_\xi(x,p)\,, \qquad \frac{\partial}{\partial p}S(x,p) = \bar f_\xi(x,p)\,.$$

The above observation leads us to the following definition[5]

[5]For the definition of the integral of a differential form over a curve, see Appendix, section 11.1.

DEFINITION 8.21. We define the *Hilbert invariant integral* as

$$H(\gamma) := \int_\gamma \omega\,,$$

where $\gamma \subset G$ is a curve of class C^1. Moreover, if $v : [a,b] \to \mathbb{R}$ is a function of class C^1 whose graph is contained in G, we define

$$H(v) := H(\mathrm{graph}(v))\,.$$

The Hilbert invariant integral is the second ingredient of the ideas stated at the beginning of this section. Indeed, the following holds.

THEOREM 8.22. *Let $u \in C^2([a,b])$ be an extremal for the lagrangian f. Suppose that u is immersed in a field of extremal,* i.e., *there exists a field of extremal $\varphi : [a,b] \times (\bar\alpha_1, \bar\alpha_2) \to \mathbb{R}^2$ and $\alpha \in (\bar\alpha_1, \bar\alpha_2)$ such that $u = u_\alpha$ in $[a,b]$. Moreover, assume that*

$$f_{\xi\xi}(x,p,\xi) > 0\,,$$

for all $(x,p) \in [a,b] \times [u(x)-\delta, u(x)+\delta]$, for some $\delta > 0$ and for all $\xi \in \mathbb{R}$. Then u is an isolated strong local minimizer of $\mathcal{F}$.

PROOF. The proof is a sequence of simple observations:

(i) since ω is exact, the Hilbert invariant integral depends only on the boundary values of the curve. In particular, if $v \in C^2([a,b])$ has the same boundary value of u, then $H(v) = H(u)$.

(ii) The Hilbert invariant integral coincides with $\mathcal{F}$ when computed on an extremal immersed in a field of extremal. Indeed, we have that

$$\begin{aligned} H(v) &= \int_{\mathrm{graph}(v)} \omega \\ &= \int_{\mathrm{graph}(v)} \Big(\big(\bar f(x,p) - P(x,p)\bar f_\xi(x,p)\big)\,\mathrm{d}x + \bar f_\xi(x,p)\,\mathrm{d}p \Big) \\ &= \int_a^b \Big(\bar f\big(x,v(x)\big) - P\big(x,v(x)\big)\bar f_\xi\big(x,v(x)\big)\big) + \bar f_\xi\big(x,v(x)\big)v'(x) \Big)\,\mathrm{d}x \\ &= \int_a^b \Big(\bar f\big(x,v(x)\big) + \Big(v'(x) - P\big(x,v(x)\big)\Big)\bar f_\xi\big(x,v(x)\big)\big) \Big)\,\mathrm{d}x \end{aligned}$$

If u is immersed in a field of extremal, we have $P(x,u(x)) = u'(x)$. This implies that $\bar f\big(x,u(x)\big) = f\big(x,u(x),u'(x)\big)$ and that the second piece of the integral above disappear.

(iii) If $v \in C^2([a,b])$ has the same boundary values of u, then, by using (i) and (ii), we have that

$$\begin{aligned} \mathcal{F}(v) - \mathcal{F}(u) &= \mathcal{F}(v) - H(u) = \mathcal{F}(v) - H(v) \\ &= \int_a^b \mathcal{E}\big(x, v(x), P\big(x,v(x)\big), v'(x)\big)\,\mathrm{d}x\,. \end{aligned} \tag{8.19}$$

(iv) Let us consider the function $g : \mathbb{R} \to \mathbb{R}$ given by

$$g(\xi) := \mathcal{E}\big(x, v(x), P\big(x,v(x)\big), \xi\big)\,.$$

Then $g\big(P\big(x,v(x)\big)\big) = 0$ and $g'\big(P\big(x,v(x)\big)\big) = 0$. Thus, by the Taylor's formula, we have that

$$\mathcal{E}\big(x, v(x), P\big(x,v(x)\big), v'(x)\big) = \frac{1}{2} f_{\xi\xi}\big(x,v(x),\eta\big)\Big(v'(x) - P\big(x,v(x)\big)\Big)^2,$$

where η is between $v'(x)$ and $P\big(x, v(x)\big)$. By hypothesis we know that $f_{\xi\xi}\big(x, u(x), \xi\big) \geq \mu > 0$ for all $x \in [a, b]$ and $\xi \in \mathbb{R}$. By continuity, we can find $r > 0$ such that $f_{\xi\xi}\big(x, v(x), \xi\big) \geq \mu > 0$ for all $\|v - u\|_{C^0} < r$, $x \in [a, b]$ and $\xi \in \mathbb{R}$. Thus, by (8.19) and the above equation, we get the desired result.

□

Only one question remains, and is what conditions we have to impose in order to be able to embed a critical point into a field of extremal. Basically, we've already solved this problem when we proved the first assertion of Theorem 8.8. We indeed have the following

LEMMA 8.23. *Let u be a critical point and suppose that there are no conjugate points to x_1 in $(x_1, x_2]$. Then u can be embedded into a C^3 field of extremal.*

PROOF. The absence of conjugate points to x_1 in $(x_1, x_2]$ allowed us to obtain the existence of a point $x_0 \in (x_1 - \delta, x_0)$ such that $\Delta(\,\cdot\,; x_0) > 0$ in $(x_0, x_2]$. Now, let us consider the following initial value problem

$$\begin{cases} L_f(w) = 0 & \text{in } [x_0, x_2]\,, \\ w(x_0) = 0\,, \\ w'(x_0) = \alpha\,. \end{cases}$$

We know that u is a solution for $\alpha_0 := u'(x_0)$. By the theory of ODE we know that there exists $\rho_0 > 0$ such that the above initial value problem admits a unique solution $\varphi(\,\cdot\,; \alpha)$ for all $\alpha \in [\alpha_0 - \rho_0, \alpha_0 + \rho_0]$. Moreover, the function $\varphi : [x_0, x_2] \times [\alpha_0 - \rho_0, \alpha_0 + \rho_0] \to \mathbb{R}$ is of class C^3.

Let us now consider the function $v(\,\cdot\,) := \varphi_\alpha(\,\cdot\,; \alpha_0)$. We know that v is a solution of the Jacobi equation $J_u(v) = 0$. By the initial condition of the above problem, we also know that

$$v(x_0) = 0\,, \qquad v'(x_0) = 1\,.$$

Thus $v(\cdot) = \Delta(\,\cdot\,; x_0)$. So $v > 0$ on $(x_0, x_2]$. By continuity, we obtain that $\varphi_\alpha(\cdot; \alpha) > 0$ on $[x_1, x_2]$. This implies that φ is the desired field of extremal. □

REMARK 8.24. It is not a surprise that the conjugate points appear also in the Weierstrass field theory, since strong local minimality implies weak local minimality, and we know that, a necessary condition in the latter case is the absence of conjugate points to x_1 inside (x_1, x_2).

Thus, the result about sufficient conditions for having strong local minimality is the following

THEOREM 8.25. *Let $u \in C^2([a, b])$ be an extremal for the lagrangian f. Suppose there are no conjugate points to x_1 in $(x_1, x_2]$. Moreover, assume that*

$$f_{\xi\xi}\big(x, p, \xi\big) > 0\,,$$

for all $(x, p) \in [a, b] \times [u(x) - \delta, u(x) + \delta]$, for some $\delta > 0$ and for all $\xi \in \mathbb{R}$. Then u is an isolated strong local minimizer of $\mathcal{F}$.

It's now time to explain where all the previous idea come from. The explanation we are going to give is called *Carathéodory royal road to Weierstrass theory*, as goes as follows.

Let us suppose we are given a critical point u embedded in a field of extremal. Assume to be so lucky that our lagrangian f is such that

$$f\big(x, p, P(x, p)\big) = 0\,, \qquad f\big(x, p, \xi\big) > 0 \quad \text{if } \xi \neq P(x, p)\,.$$

Since $u'(x) = P(x, u(x))$ for all $x \in [a,b]$, we simply have

$$\mathcal{F}(u) < \mathcal{F}(v)\,, \qquad \text{for all } v \neq u\,.$$

So, problem solved! As a matter of fact, it's not so robust to rely on lucky! But the basic idea is the above one. We can try to fit into the above situation by adding to our lagrangian f a null-lagrangian g. Thanks to the characterization of null-lagrangians in the one dimensional scalar case, we know that g must be of the form

$$g(x,p,\xi) = S_x(x,p) + S_p(x,p)\xi\,,$$

for some function $S \in C^2(G)$. Let us consider the modified lagrangian

$$\widetilde{f}(x,p,\xi) := f(x,p,\xi) - S_x(x,p) - S_p(x,p)\xi\,.$$

We want this modified lagrangian to satisfy the following properties

$$\widetilde{f}\big(x,p,P(x,p)\big) = 0\,, \qquad \widetilde{f}\big(x,p,\xi\big) > 0 \quad \text{if } \xi \neq P(x,p)\,. \tag{8.20}$$

Thus, we are asking $\xi = P(x,p)$ to be the only global minimum of the function

$$\xi \mapsto \widetilde{f}\big(x,p,\xi\big)\,,$$

for every $(x,p) \in G$. So, we must have the following conditions in force

$$\widetilde{f}_\xi\big(x,p,P(x,p)\big) = 0\,, \qquad \widetilde{f}_{\xi\xi}\big(x,p,P(x,p)\big) \geq 0\,.$$

The first condition above gives us

$$S_p(x,p) = f_\xi\big(x,p,P(x,p)\big)\,, \tag{8.21}$$

and the second one is

$$f_{\xi\xi}\big(x,p,P(x,p)\big) \geq 0\,,$$

while the first condition of (8.20) gives us

$$S_x(x,p) = f\big(x,p,P(x,p)\big) - P(x,p) f_p\big(x,p,P(x,p)\big)\,. \tag{8.22}$$

By combining together (8.21) and (8.22) we obtain that the second requirement in (8.20) writes as

$$f(x,p,\xi) - f\big(x,p,P(x,p)\big) - (p - P(x,p)) f_\xi\big(x,p,P(x,p)\big) \geq 0\,,$$

that is the Weierstrass necessary condition.

Now the picture is complete! Namely, by asking (8.20) we end up by having all the conditions we came upon along the exposition of the Weierstrass field theory.

8.7. Stigmatic fields and Jacobi's envelope theorem

The aim of this section is to prove Jacobi's envelope theorem, that will allows us to say something more about the weak (and strong) minimality property of an extremal up to its first conjugate point.

We begin by introducing something we already used when we constructed a Jacobi field in Theorem 8.8. In that proof we started from a fixed point $P_0 := (x_0, p_0)$ and we construct a family of critical points emanating from P_0. We want to give a name to this object, also allowing each estremal to have a different interval of definition.

Definition 8.26. Take $x_0, p_0 \in \mathbb{R}$, $I \subset \mathbb{R}$ compact, and define the sets

$$\widetilde{\Gamma} := \{\,(x,\alpha) \in \mathbb{R} \times I : x_0 \leq x \leq x(\alpha)\,\}\,,$$

$$\Gamma := \{\,(x,\alpha) \in \mathbb{R} \times I : x_0 < x \leq x(\alpha)\,\}\,.$$

We say that $\varphi : \widetilde{\Gamma} \to \mathbb{R}^2$ given by

$$\varphi(x,\alpha) = \big(x, u_\alpha(x)\big)\,,$$

is a *stigmatic field* if $\varphi_{|_\Gamma}$ is a field of extremals as in Definition 8.18 and $\varphi(x_0,\alpha) = p_0$ for all $\alpha \in I$. The point $P_0 := (x_0, p_0)$ is called *nodal point.*

For such a fields, we have a very nice way to compute the eikonal S.

THEOREM 8.27. *Let S denotes the eikonal related to the field $\varphi_{|_\Gamma}$. Then*

$$S(x,p) = \Sigma\big(x, \alpha^{-1}(x,p)\big)\,,$$

where

$$\Sigma(x,\alpha) := \int_{x_0}^{x} f\big(y, h(y,\alpha), h'(y,\alpha)\big)\,\mathrm{d}y\,. \tag{8.23}$$

Moreover $\lim_{P\to P_0} S(P) = 0$, *for* $P \in \varphi(\Gamma)$.

PROOF. It is clear from the definition that $\Sigma \in C^1(\widetilde{\Gamma})$. In order to prove the first claim, we show that $T(x,p) := \Sigma\big(x,\alpha^{-1}(x,p)\big)$ satisfies the equations characterizing S. Then

$$T_p(x,p) = \Sigma_\alpha\big(x,\alpha^{-1}(x,p)\big)\alpha_p^{-1}(x,p)\,.$$

We want to rewrite the two terms on the right-hand side. Let us start from the second one. By recalling that

$$p = \varphi\big(x, \alpha^{-1}(x,p)\big)\,,$$

by differentiating with respect to p, we deduce that

$$1 = \varphi_\alpha\big(x,\alpha^{-1}(x,p)\big)\alpha_p^{-1}(x,p)\,,$$

and thus

$$\alpha_p^{-1}(x,p) = \varphi_\alpha^{-1}\big(x,\alpha^{-1}(x,p)\big)\,. \tag{8.24}$$

For the first term, we notice that

$$\begin{aligned}\Sigma_\alpha(x,\alpha) &= \int_{x_0}^{x} L_f\big(\varphi(y,\alpha)\big)\varphi_\alpha(y,\alpha)\,\mathrm{d}y + f_\xi\big(y,\varphi(y,\alpha),\varphi'(y,\alpha)\big)\varphi_\alpha(y,\alpha)\Big|_{x_0}^{x} \\ &= f_\xi\big(x,\varphi(x,\alpha),\varphi'(x,\alpha)\big)\varphi_\alpha(x,\alpha)\,,\end{aligned}$$

where in the last equality we used the fact that $\varphi(\cdot,\alpha)$ is an extremal, and that $\varphi(x_0,\alpha) = p_0$ for all $\alpha \in I$. By this equality and (8.24), we get

$$T_p(x,p) = f_\xi\big(x,\varphi(x,\alpha),\varphi'(x,\alpha)\big)\,.$$

Moreover, by differentiating the identity $T\big(x,\varphi(x,\alpha)\big) = Sigma(x,\alpha)$ with respect to x, we get

$$\Sigma_x(x,\alpha) = T_x(x,p) + T_p(x,\alpha)\varphi'(x,\alpha) = T_x(x,p) + T_p(x,\alpha)P(x,p)\,,$$

while, by differentiating (8.23) with respect to x, we obtain

$$\Sigma_x(x,p) = f\big(x,p,P(x,p)\big)\,.$$

Thus

$$T_x(x,p) = f\big(x,p,P(x,p)\big) - f_\xi\big(x,p,P(x,p)\big)P(x,p)\,.$$

This implies that $T = S$.

To prove the limit, we notice that, by the compactness of I, we have that

$$\lim_{x\to x_0} \Sigma(x,\alpha) = 0\,,$$

uniformly in $\alpha \in I$. This implies that $\lim_{P\to P_0} S(P) = 0$. $\square$

Before proving Jacobi's envelope theorem, we recall some facts about the locus of first conjugate points. Suppose the strict Legendre-Hadamard condition

$$f_{\xi\xi}\big(x, \varphi(x,p), \varphi'(x,p)\big) > 0\,,$$

is in force, as well as $\varphi'_\alpha(x_0,\alpha) \neq 0$, for all $\alpha \in I$. Then, for each $\alpha \in I$, the function $v(\cdot) := \varphi_\alpha(\cdot,\alpha)$ turns out to be a Jacobi field over the extremal $\mathcal{C}(\alpha) := \varphi\big((x_0, x(\alpha)), \alpha\big)$, *i.e.*, the image of $(x_0, x(\alpha))$ through the map $\varphi(\cdot,\alpha)$. Suppose the locus of first conjugate point

$$\mathcal{E} := \big\{\big(x, \varphi(x,\alpha) \;:\; x_0 < x < x(\alpha)\,,\, \varphi_\alpha(x,\alpha) = 0\big)\big\}\,,$$

to be non empty. Denote by $P(\alpha) = \big(g(\alpha), h(\alpha)\big)$ a point in this set, where $h(\alpha) = \varphi\big(g(\alpha),\alpha\big)$. We would like to give a description of $\mathcal{E}$. For, we recall that in Section 8.4 we proved that, if we require $\varphi_{\alpha\alpha}\big(g(\alpha),\alpha\big) \neq 0$, then $\mathcal{E}$ turns out to be a graph of a C^1 function $s : (x_1, x_2) \to \mathbb{R}$, for some $x_0 < x_1 < x_2$. Notice that the function s is given by

$$s(s) = h\big(g^{-1}(x)\big)\,,$$

since the hypothesis $\varphi_{\alpha\alpha}\big(g(\alpha),\alpha\big) \neq 0$ allows us to say that g has an inverse. From the above formula we get that

$$s'(x) = h'\big(g^{-1}(x)\big)(g^{-1})'(x) = \frac{h'\big(g^{-1}(x)\big)}{g'\big(g^{-1}(x)\big)} = \varphi'\big(x, g^{-1}(x)\big)\,,$$

where in the last step we used the definition of h. Thus, the graph of s intersect *tangentially* every curve of the family of extremals. This is why $\mathcal{E}$ is called the *envelope* of the family of extremals.

We are now in position to prove the main result of this section

THEOREM 8.28 (Jacobi's envelope theorem). *Let us suppose all the before-mentioned hypothesis to hold true. Take α_1 and α_2 such that $g(\alpha_1) < g(\alpha_2)$ Then hte following formula holds*

$$\begin{aligned}\int_{x_0}^{g(\alpha_2)} f\big(x, \varphi(x,\alpha_2), \varphi'(x,\alpha_2)\big)\,\mathrm{d}x &= \int_{x_0}^{g(\alpha_1)} f\big(x, \varphi(x,\alpha_1), \varphi'(x,\alpha_1)\big)\,\mathrm{d}x \\ &\quad + \int_{g(\alpha_1)}^{g(\alpha_2)} f\big(x, s(x), s'(s)\big)\,\mathrm{d}x\,.\end{aligned}$$

PROOF. The idea is just to extend by continuity the slope field up to graphs by

$$P\big(x, s(x)\big) := s'(x) = \varphi'\big(x, g^{-1}(x)\big) = P\big(x, s(x)\big)\,.$$

Then

$$\frac{\mathrm{d}}{\mathrm{d}x} S\big(x, s(x)\big) = S_x\big(x, s(x)\big) + S_p\big(x, s(x)\big)s'(x) = f\big(x, s(x), s'(x)\big)\,,$$

where in the last step we used the fact that S satisfies

$$S_x(x,p) = \bar{f}(x,p) - P(x,p)\bar{f}_\xi(x,p)\,, \qquad S_p(x,p) = \bar{f}_p(x,p)\,.$$

Thus

$$\int_{g(\alpha_1)}^{g(\alpha_2)} f\big(x, s(x), s'(s)\big)\,\mathrm{d}x = S\big(P(\alpha_2)\big) - S\big(P(\alpha_1)\big)\,,$$

and then the result follows by applying the previous theorem. $\square$

An interesting consequence of the above result, is that, under our hypothesis, a critical point looses its weak minimality property as soon as it reaches its first conjugate point. Indeed, let u be such a critical point, and assume that $u(\cdot)=\varphi(\cdot,\alpha_2)$, for some $\alpha_2\in I$ that is not the left-hand point of I. Then, pick $\alpha_1<\alpha_2$, and define the function

$$v(x):=\begin{cases}\varphi(x,\alpha_1) & x\in[x_0,g(\alpha_1)\,,\\ s(s) & x\in[\alpha_1,\alpha_2]\,.\end{cases}$$

Notice that $v\in C^2$, since $\mathcal{E}$ intersects each curve of the family tangentially. By the Jacobi's envelope theorem we have that

$$\mathcal{F}(u)=\mathcal{F}\big(\varphi(\cdot,\alpha_2)\big)=\mathcal{F}(v)\,.$$

Thus, if u is a weak local minimizer, for α_1 sufficiently closed to α_2, also v is a weak local minimizer. Then v must satisfy the Euler-Lagrange equation $L_f(v)=0$, that is a second order one. Since also $u=\varphi(\cdot,\alpha_2)$ satisfies the same equation, and noticing that

$$v\big(g(\alpha_2)\big)=u(\alpha_2)\,,\qquad v'\big(g(\alpha_2)\big)=u'(\alpha_2)\,,$$

we deduce that $u=v$. But this is in contradiction with the characterization we gave of $\mathcal{E}$, since we know that $\varphi_{\alpha\alpha}\neq 0$.

Thus, if $\mathcal{E}$ is a graph of a nice function, critical points loose their weak local minimality property as soon as they touch $\mathcal{E}$. It has been discover by Kneser that, in the case $\mathcal{E}$ has a cusp, it can be the case that extremals touching $\mathcal{E}$ at the cusp maintain their weak local minimality property.

8.8. Solution of the minimal surfaces of revolution problem

We now wan to apply the previous theory to solve the minimal surfaces of revolution problem. Let us recall that our energy is

$$\mathcal{F}(u):=2\pi\int_a^b u(x)\sqrt{1+\big(u'(x)\big)^2}\,\mathrm{d}x\,.$$

Let us fix a point $P_1=(x_1,p_1)$. Then, the solutions of the Du Bois-Reymond equation passing through the point P_1 can be parametrized by a parameter α as follows:

$$\varphi(x,\alpha):=\frac{p_1}{\cosh\alpha}\cosh\Big(\alpha+\frac{x-x_1}{p_1}\cosh\alpha\Big)\,.$$

Recall that solutions of Du Bois-Reymond equation with $\varphi'\neq 0$ are solutions of the Euler-Lagrange equation. Since we want to construct a family of extremal, we will have to make sure that $\varphi'\neq 0$ in that region. We will denote by $\mathcal{C}(\alpha)$ the catenoid relative to the parameter α. Its vertex is the point

$$P=\Big(x_1-p_1\frac{\alpha}{\cosh\alpha},\frac{p_1}{\cosh\alpha}\Big)\,.$$

The problem we want to study is the following: given a final point $P_2=(x_2,p_2)$, we want to solve the problem of the minimal surface of revolution passing trough P_1 and P_2. For, we want to apply the Weierstrass field theory. First of all we notice that the Weierstrass condition is satisfied. Indeed

$$\mathcal{E}(x,p,\xi,\eta)=f(x,p,\eta)-f(x,p,\xi)-f_\xi(x,p,\xi)(\eta-\xi)=\frac{1}{2}f_{\xi\xi}(x,p,\xi,v)\,,$$

for some v in between ξ and η. In the last step we used the fact that $f(x,p,\xi,\xi)=f_\xi(x,p,\xi,\xi)=0$. Since

$$f_{\xi\xi}(x,p,\xi,\eta)=\frac{2\pi p}{(1+\xi^2)^{\frac{3}{2}}}\,,$$

we conclude that

$$\mathcal{E}(x,p,\xi,\eta) > 0\,, \qquad \text{for } \eta \neq \xi\,.$$

You already proved in the homework that there exists a curve $\mathcal{E}$ (yes, this is a bad notation because you can get confused with the Weierstrass excess function. But from now on we will reserve the symbol $\mathcal{E}$ for denoting that curve!) such that

- if P_2 is above the curve we have two catenaries connecting P_1 to P_2,
- if $P_2 \in \mathcal{E}$ we just have one catenary joining it with P_1,
- if P_2 is below $\mathcal{E}$, none of the above solutions passe through it.

In the last case, we cannot have a solution to the minimum problem (among the class of C^1 graphs). In the other two cases, we need to find, for each catenary $\mathcal{C}(\alpha)$ the first conjugate point to P_1. For, we will use a geometric construction due to Lindelöf, that says the following:

draw the tangent line to $\mathcal{C}(\alpha)$ from P_1, and let $(\bar{x}_\alpha, 0)$ its intersection point with the x-axes. From it draw a line that is tangent to the curve (if possible) in a point P_1^*. Then, P_1^* is the first (and only) conjugate point to P_1 along $\mathcal{C}(\alpha)$.

We will prove this assertion for our case, but a similar construction can be carried out also for a general class of lagrangians showing some sort of geometrical invariance. Let us take a point $P = (x, \varphi(x,\alpha))$ on $\mathcal{C}(\alpha)$, and let consider the intersection $(\bar{x}(\alpha), \bar{y}(\alpha))$ of the tangent line passing through that point, with the one passing through P_1. That point has to satisfy

$$\left\{ \; \bar{x}(\alpha) = x_1 + \tfrac{\bar{y}(\alpha) - p_1}{\varphi'(x_1,\alpha)} \,, \bar{x}(\alpha) = x + \tfrac{\bar{y}(\alpha) - \varphi(x,\alpha)}{\varphi'(x,\alpha)} \,. \right.$$

The above equations imply that

$$\bar{y}(\alpha) = \frac{\varphi'(x,\alpha)\sinh\alpha}{\varphi'(x,\alpha) - \sinh\alpha}\left[x - x_1 - \frac{\varphi(x,\alpha)}{\varphi'(x,\alpha)} + \frac{p_1}{\sinh\alpha} \right].$$

By noticing that (after some algebra)

$$\varphi_\alpha(x,\alpha) = \frac{1}{\cosh\alpha}\left[-\varphi_\alpha(x,\alpha)\sinh\alpha + p_1\varphi'(x,\alpha) + x\varphi'(x,\alpha)\sinh\alpha - x_1\varphi'(x,\alpha)\sinh\alpha \right],$$

we can write

$$\bar{y}(\alpha) = \frac{\cosh\alpha\varphi_\alpha(x,\alpha)}{\varphi'(x,\alpha) - \sinh\alpha}\,,$$

where we recall that the denominator is always non zero! Thus

$$\varphi_\alpha(x,\alpha) = \frac{\varphi'(x,\alpha) - \sinh\alpha}{\cosh\alpha}\bar{y}(\alpha)\,.$$

Since the zeros of $\varphi_\alpha(\cdot,\alpha)$ are the conjugate points to P_1, we see that the only possibility to have it equal to zero is when $\bar{y}(\alpha) = 0$, that is precisely the claim we wanted to prove. Finally, we notice that the convexity of $\varphi(\cdot,\alpha)$ implies that there can be (at most) one such a conjugate point. In particular, if P_1 is the vertex of the catenary α, there is no conjugate point to it.

It is possible to perform a study of the behavior of the envelope We do not want to present here the computations needed to describe the behavior of the envelope $\mathcal{E}$, proving that it turns out to be a strictly convex increasing function, converging to 0 as $x \to 0$. We do not want to present here the computations for this kind of study.

We just summarize the complete solution of the minimal surface of revolution problem:

(i) the envelope $\mathcal{E}$ is the graph of a real analytic strictly convex function h with $\lim_{x\to 0^+} h(x) = 0$ and $\lim_{x\to\infty} h(x) = \infty$.

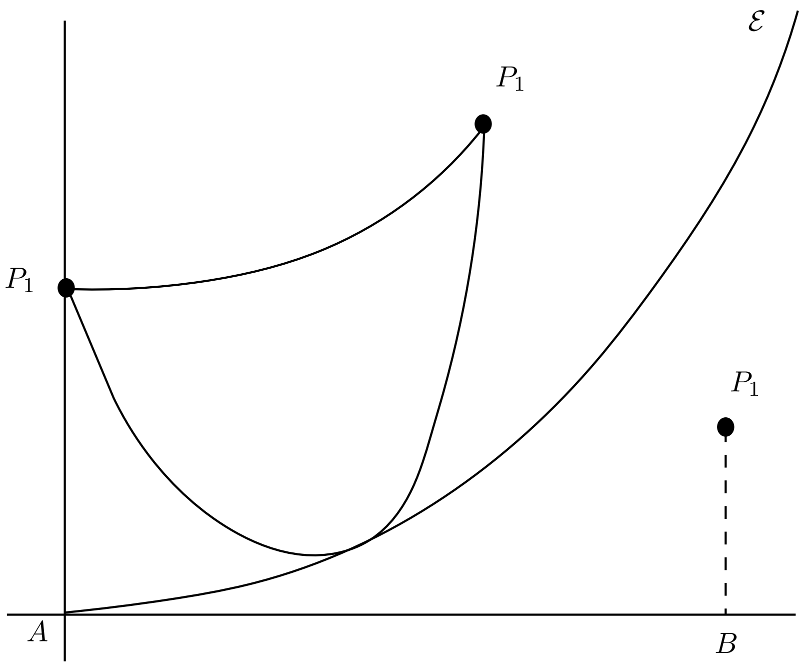

FIGURE 1. The envelope $\mathcal{E}$ separates the upper region where we have one or two catenary joining P_1 and P_2 and the lower one, where the *Goldsmith curve* (from P_1 down to A, straight to B and up to P_2) turns out to be the global minimizer.

(ii) If P_2 lies above $\mathcal{E}$, there are two catenaries joining it to P_1. If $P_2 \in \mathcal{E}$ we have only one such a catenary and if P_2 is below the envelope curve, no extremal connects it to P_1.

(iii) in the case $P_2 \in \mathcal{E}$, then P_2 is conjugate to P_1 and, by the result of the previous section, $\mathcal{C}(\alpha)$ is *not* a weak local minimizer.

(iv) If P_2 is above $\mathcal{E}$, the lower catenary has a conjugate point in between P_1 and P_2, and thus it is *not* a weak local minimizer. While the upper one has no conjugate point in between, so it turns out to be a strong local minimizer.

(v) If P_2 lies below $\mathcal{E}$, we have no classical solution. Nevertheless, a *solution* (among a more general class of admissible competitors) is given by the so called *Goldsmith curve* shown in the figure.

(vi) There exists a curve $\mathcal{G}$ that is above $\mathcal{E}$ such that if P_2 lies above $\mathcal{G}$, then the strong local minimizer catenary of the previous item is also a *global* solution. While, if it lies below, the global solution is give by the Goldsmith curve.

CHAPTER 9

The isoperimetric problem

This chapter is entirely devoted to the presentation of three proofs of the isoperimetric inequality in the plane. The reason[1] is that the techniques used to prove it have nothing in common with what we've presented so far. Thus, it is interesting to see how this problem has been tackled from different points of view.

9.1. A bit of history

In this chapter we want to present three proofs of the isoperimetric problem. This problem is of particular interest for the Calculus of Variations: not only it is the oldest one, but the attempts to prove it led to very important developments in mathematics. The story of this problem is very long and we do not want to report it here. We just limit ourselves to recall the two ends of the rope. The first result in the solution of the isoperimetric problem was given by Zenodoro, a greek mathematician who lived around 200 B.C.. He proved that the circle has a better isoperimetric constant among all polygons. In particular, he showed that among all polygons of a fixed number of edges, the regular one is better. The complete solution of the isoperimetric problem, *i.e.*, the solution of the problem in the most general class (where the enlargement of the class, as for the one of generalization of the notion of perimeter of a set, is due to the fact that we also need to prove an existence result for the problem) and for all dimensions, has been given in 1958 by the italian mathamatician De Giorgi. It is interesting to notice how long this problem has resisted to the attempt of mathematician of obtaining a complete solution to it.

We state the result we want to prove.

THEOREM 9.1 (Isoperimetric inequality). *For all regions in the plane that are enclosed by a curve of class C^1, the following inequality holds:*

$$4\pi A \leq L^2 \,,$$

where A denotes the area of the region and L the length of the curve.

The proofs we decided to present are of three different flavors: the first one is due to Steiner (19^{th} century) and can be cataloged as a *mechanical* proof. The second one, an analytic one, is due to Hurwitz (beginning of the 20^{th} century) and it is based on Fourier series and the Poincaré-Wirtinger inequality. The last one is more geometric and, we can say, that capture the best the peculiarity of the problem. It is due to Minkowski and uses the Brunn-Minkowski and the Steiner inequalities.

The proof we are presenting are taken from [11].

[1] The real reason is that I really like the isoperimetric problem!

9.2. Steiner's proof

Steiner gave five different proofs of the isoperimetric inequality. We can say that they are of a *mechanical* type. The one we want to present here goes as follows:

let E be a set that maximize the area among all planar figures enclosed by curves having a fixed length. Cut E with a segment S in such a way that the curve γ enclosing E is divided in two pieces of equal length. Call E_1 and E_2 the two regions E is divided into by S.

Then E_1 and E_2 must have the same area. Otherwise, if E_1 had more area than E_2, by considering the set obtained by E_1 and its reflection with respect to the segment S, we would obtain a set $\widetilde{E}$ with the same perimeter of E but with bigger area. This would contradict the maximality of E.

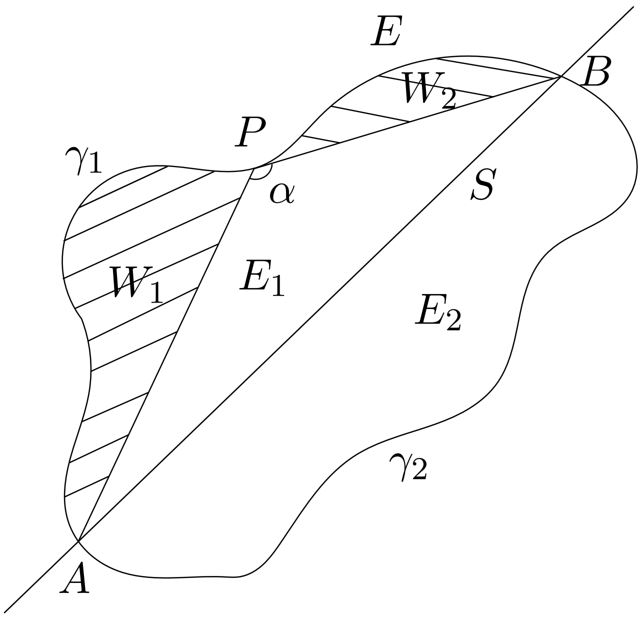

FIGURE 1. A picture showing the argument of Steiner.

Now, let us consider one of the two pieces, let's say E_1. If E_1 were not a half-circle, there would exists a point P on its boundary such that the angle α is not $\pi/2$ (see Figure 1). Think at W_1 and W_2 as two pieces of a scissor, that can move changing the angle α (and thus, the length of the segment AB). Notice that the area of the triangle APB is maximized when $\alpha = \pi/2$.

So, as before, if such a point P existed we could consider the set $\widetilde{E}_1$ obtained by E_1 by moving W_1 and W_2 in such a way that $\alpha = \pi/2$ and thus a set $\widetilde{E}$ obtained by reflection of $\widetilde{E}_1$ with the same perimeter of E (the segment AB disappears inside $\widetilde{E}$) but with bigger area. Contradiction.

Thus, E must be a circle.

9.3. Hurwitz's proof

We now present the proof by Hurwitz, that is based on the following

THEOREM 9.2 (Poincaré-Wirtinger inequality). *Let $f \in C^1(\mathbb{R})$ be 2π-periodic and denote by $\bar{f}$ the mean of f on a interval of length 2π. Then*

$$\int_0^{2\pi} \big(f(x) - \bar{f}\big)^2 \, \mathrm{d}x \le \int_0^{2\pi} \big(f'(x)\big)^2 \, \mathrm{d}x \,,$$

and equality holds if and only if $f(x) = \bar{f} + a\cos x + b\sin x$.

The proof of the above theorem can be done by writing f and f' in Fourier series and comparing term by term. By using the above result, we can prove the isoperimetric inequality as follows.

PROOF. (Hurwitz's proof of the isoperimetric inequality) let us consider the parametrization of the curve enclosing a region in the plane by arclength $s \mapsto \big(x(s), y(s)\big)$. If L denotes the length of the curve, then we can extend periocally the functions x and y. We can now consider the 2π periodic functions

$$f(t) := x\Big(\frac{Lt}{2\pi}\Big)\,, \qquad g(t) := y\Big(\frac{Lt}{2\pi}\Big)\,.$$

Then, we have that

$$(f')^2 + (g')^2 = \frac{L^2}{4\pi^2}\,. \tag{9.1}$$

Noticing that $\int_0^{2\pi} gi(t)\,\mathrm{d}t = 0$ (since the curve is closed!), we have that

$$\begin{aligned}
2A = 2\int_0^{2\pi} f(t)g'(t)\,\mathrm{d}t &= 2\int_0^{2\pi} \big(f(t) - \bar{f}\big)g'(t)\,\mathrm{d}t \\
&= \int_0^{2\pi} \Big[\big(f(t) - \bar{f}\big)^2 + \big(g'(t)\big)^2 - \big(f(t) - \bar{f} - g'(t)\big)^2\Big]\,\mathrm{d}t \\
&\le \int_0^{2\pi} \Big(\big(f'(t)\big)^2 + \big(g'(t)\big)^2\Big)\,\mathrm{d}t \\
&= \int_0^{2\pi} \frac{L}{4\pi^2}\,\mathrm{d}t = \frac{L^2}{2\pi}\,.
\end{aligned}$$

Thus the inequality is proved. Notice the the equality holds if we have equality in the Poincaré-Wirtinger inequality, *i.e.*, if

$$f(t) = \bar{f} + a\cos t + b\sin t\,,$$

for some $a, b \in \mathbb{R}$ and if

$$\int_0^{2\pi} \big(f(t) - \bar{f} - g'(t)\big)^2\Big]\,\mathrm{d}t = 0\,.$$

From the second one we get $g'(t) = f(t) - \bar{f}(t)$, that, using the first condition, writes as

$$g(t) = \bar{g} + a\sin t - b\cos t\,.$$

By using (9.1) we have that

$$a^2 + b^2 = \frac{L^2}{4\pi^2}\,.$$

Thus, we obtain that the curve enclosing our set is a circle. $\square$

9.4. Minkowski's proof

We now present the third proof. It is based on two geometric results. The underlining idea is to *add* to a set E a ball of radius r, and to estimate from above and from below the area of the resulting set as well as an estimate from above of the perimeter. The proofs we present will lack of the last step: this is because we do not want to introduce the notion of Lebesgue measure and of perimeter (for nice sets!). But the idea behind the approximation arguments is quite believable!

We first need to specify what we mean by adding a set to another one. we will present all the notions for the plane, but they can be generalized to higher dimensions.

DEFINITION 9.3. Let $A, B \subset \mathbb{R}^2$. We define the set $A+B$ as follows

$$A+B := \{ x+y \,:\, x \in A,\, y \in B \}.$$

We now make some examples:

- $A + B_r = \{ x \in \mathbb{R}^2 \,:\, \operatorname{dist}(x, A) \leq r \}$,
- if R_1 and R_2 are two rectangles, then $R_1 + R_2$ is a rectangle as well.

The estimates from above are given by the following result

THEOREM 9.4 (Steiner's inequalities). *Let $A \subset \mathbb{R}^2$ be a closed and bounded set with piecewise C^1 boundary. Let us denote with A and L its area and its perimeter respectively. Then*

$$\operatorname{Area}\big(E + B_r\big) \leq A + Lr + \pi r^2\,,$$
$$\operatorname{Length}\big(\partial(E + B_r)\big) \leq L + 2\pi r\,.$$

PROOF. *Step 1:* in the case E is a convex polygon, we actually have equality in both estimates, as can be easily seen from Figure 2.

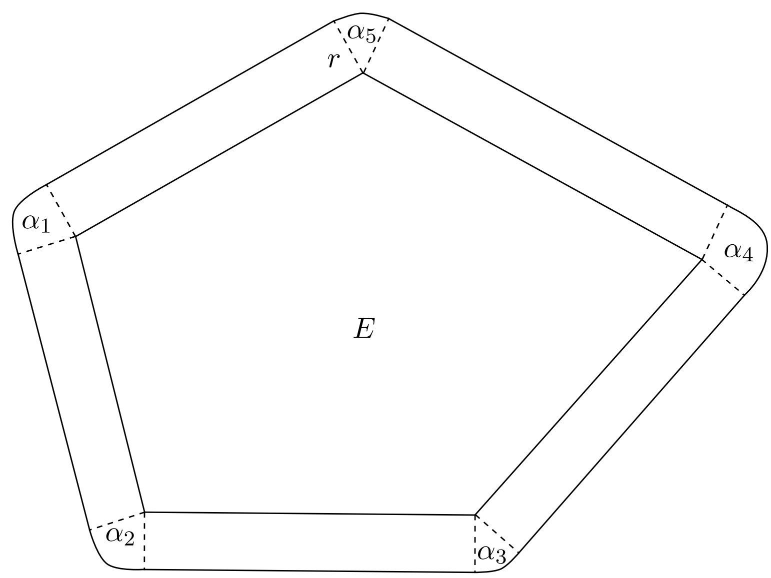

FIGURE 2. For a convex polygon we have that $\sum_i \alpha_i = 2\pi$.

Step 2: in the case E is a general polygon, we easily have the claimed inequalities.

Step 3: here we have to rely on the two following facts:

(i) every set E as in the theorem can be approximated by a sequence of polygons $(P_n)_n$ such that their area and their perimeter converge to those of E,

(ii) the addition of sets is continuous with respect to the above convergence, that is $P_n + B_r \to E + B_r$, and the area and the length converge as well.

By taking from grant the above statements, we proved the desired result. □

The lower bound for the area is more delicate.

THEOREM 9.5 (Brunn-Minkowski's inequality). *Let $A, B \subset \mathbb{R}^2$ be measurable sets* [2]. *Then*

$$\sqrt{\operatorname{Area}(A+B)} \geq \sqrt{\operatorname{Area}(A)} + \sqrt{\operatorname{Area}(B)}\,.$$

PROOF. *Step 1:* we prove the inequality when A and B are union of rectangles. We prove it by induction. Let us first suppose that $A = (a,b) \times (c,d)$ and $B = (e,f) \times (g,h)$. Then

$$\begin{aligned}
\operatorname{Area}(A+B) &= \operatorname{Area}\big((a+c, b+f) \times (c+g, d+h)\big)\\
&= (b-a+f-e)(d-c+h-g)\\
&= (b-a)(d-c) + (f-e)(h-g) + (b-a)(h-g) + (f-e)(d-c)\\
&\geq (b-a)(d-c) + (f-e)(h-g) + 2\sqrt{(b-a)(h-g)(f-e)(d-c)}\\
&= \Big(\sqrt{(b-a)(d-c)} + \sqrt{(f-e)(h-g)}\,\Big)^2\\
&= \Big(\sqrt{\operatorname{Area}(A)} + \sqrt{\operatorname{Area}(B)}\,\Big)^2
\end{aligned}$$

where we used the *arithmetic-geometric inequality*

$$\frac{|x|+|y|}{2} \geq \sqrt{|x|\,|y|}\,.$$

We now perform the induction step. Let us suppose the inequality is valid for all A and B that are union of rectangles with the total number of rectangles less than or equal to $l-1$. We want to prove it also when the number of rectangles is l. Take $A = \cup_{i=1}^n R_i$ and $A = \cup_{j=1}^m S_j$, where R_i and S_j are rectangles and $m+n = l$. Without loss of generality, we can suppose $n \geq 2$. We proceed as follows: choose two rectangles in A and place an horizontal or a vertical line between them. Suppose the line $\{x = x_1\}$ divides them (in the case of an horizontal line, proceed in a similar way). Consider the families of rectangles obtained by splitting the ones in A with the line $\{x = x_1\}$, that is, define the sets

$$\begin{aligned}
A_1 &:= \{\, R_i \cap \{x < x_1\} \;:\; i = 1, \ldots, n\,\}\,,\\
A_2 &:= \{\, R_i \cap \{x > x_1\} \;:\; i = 1, \ldots, n\,\}
\end{aligned}$$

By construction we have that the number or rectangles in each A_1 and A_2 is less than n, since there exist at least two rectangles that are on different half-spaces. Now perform the same procedure for the rectangles in B, where this time the line $\{x = x_2\}$ is chosen in such a way that

$$\theta = \frac{\operatorname{Area}(A_1)}{\operatorname{Area}(A)} = \frac{\operatorname{Area}(B_1)}{\operatorname{Area}(B)}\,.$$

Finally, notice that

$$A + B \supset A_1 + B_1 \,\cup\, A_2 + B_2\,,$$

[2] If you don't know what a measurable set is, I can ensure you that every set that comes to your mind is measurable. You need the Axiom of Choice to built a non measurable set!

and that the two Minkowski sums on the right-hand side are disjoint sets (on different sides of $x = x_1 + x_2$). We are now in position to conclude. Indeed, by using the induction hypothesis, we have that

$$\begin{aligned}\mathrm{Area}(A+B) &\geq \mathrm{Area}(A_1+B_1) + \mathrm{Area}(A_2+B_2)\\ &\geq \big(\sqrt{\mathrm{Area}(A_1)} + \sqrt{\mathrm{Area}(B_1)}\,\big)^2 + \big(\sqrt{\mathrm{Area}(A_2)} + \sqrt{\mathrm{Area}(B_2)}\,\big)^2\\ &= \theta\big(\sqrt{\mathrm{Area}(A)} + \sqrt{\mathrm{Area}(B)}\,\big)^2 + (1-\theta)\big(\sqrt{\mathrm{Area}(A)} + \sqrt{\mathrm{Area}(B)}\,\big)^2\\ &= \big(\sqrt{\mathrm{Area}(A)} + \sqrt{\mathrm{Area}(B)}\,\big)^2 .\end{aligned}$$

Step 2: we conclude by using an approximation argument similar as the one, we conclude for general sets by using the previous step. □

We can now present Minkowski's argument.

PROOF. (Minkowski's proof of the isoperimetric inequality) Let E be a set whose boundary is of class C^1. Then, by using Steiner's inequalities and the Brunn-Minkowski inequality, with just one line of computations[3] we obtain

$$\begin{aligned}A + L\pi + \pi r^2 \geq \mathrm{Area}(E+B_r) &\geq \Big(\sqrt{\mathrm{Area}(A)} + \sqrt{\mathrm{Area}(B_r)}\,\Big)^2\\ &= A + r\sqrt{4\pi A} + \pi r^2 .\end{aligned}$$

□

REMARK 9.6. It is worth noticing that, according to my little knowledge, all the proofs of the isoperimetric inequality (also the modern ones with more sophisticated tools and in a more general setting) use the arithmetic-geometric inequality.

[3] Actutally two, but just for aesthetic reasons!

CHAPTER 10

A general overview of the modern Calculus of Variations

The aim of what follows is the same of those of a book's backcover: make you so interested about it to convince you to download it or to wait 'till it comes out a movie based on it. Since the latter option is very unlikely to happen[1], I'm afraid that downloading articles and reading them is the only way to get more into the modern approach of the Calculus of Variation.

So, what we are going to do is to give a very brief and certainly not exhaustive overview of the modern developments on the subject, just to make the reader aware of why all the stuff we presented so far are called *classical* and what are the main ideas and issues of the modern Calculus of Variations. These latter will be introduced in a *non* rigorous way and no formal definitions or proofs will appear. Nevertheless, we will try to explain the heuristic leading to the formal definitions the reader will find in the *uncomplete* list of references we will provide.

10.1. Ain't talkin' 'bout unicorn

Everybody knows that the unicorn's[2] corn possesses unbelievable medical powers and that the dust obtained from it, opportunely mixed with a potion, can protect from deadly diseases.

FIGURE 1.
The gentle and pensive maiden has the power to tame the unicorn,
fresco probably by Domenico Zampieri, c. 1602

[1] Let's be realistic: a movie about the modern approach of the Calculus of Variation would be very, very boring!

[2] If you wish to catch a unicorn (or even get close to it), remember that unicorns can be tamed only by a virgin, traditionally naked sitting beneath a tree. But be aware: if a girl is merely pretending to be a virgin, the unicorn would tear her apart! So remember: don't mess with the unicorns!

We can really laugh about it. But why can we? Well, because we are sure (are we?) that unicorns do not exist. So, whatever we say say about them is clearly false, since they do not exist. And, on the other side, everything can be said about them. Indeed, if a sentence starts with *"If you take a unicorn..."*, then we can end it as we wish, since *ex falso sequitur quodlibet*[3].

Now, let us consider a typical statement in the classical Calculus of Variations we've encountered many times. It starts with *"Let u be a minimizer..."*. Then, we can ask ourselves: is u a unicorn? Namely, are we sure that all the theorems we proved didn't have as a primary hypothesis a false one? Well, it would be really embarrassing, after an entire course, to be forced to say: *"My dear students, do you remember all the theorems we've proved in the course? Well, it turns out that they are empty..."*. In that case unicorns with laugh at us! But we gave three examples of minimizers of variational problems. Hence, I'm afraid we'll never here the sound of a unicorn laughing.

Let's now forget about unicorns and their magic world, and let us come back to earth. It is very surprising how long the question of existence of minimizers has been underestimated and, in some cases, simply not taken into consideration. A famous case of the above situation happens with the so called *Dirichlet principle*. It states that, in order to find a solution of the minimum problem[4]

$$\min_{\substack{u\in C^2(\bar\Omega)\\ u=g \text{ on } \partial\Omega}} \frac{1}{2}\int_\Omega |\nabla u|^2 \, \mathrm{d}x\,,$$

we just have to solve the equation

$$\begin{cases} \triangle u = 0 & \text{in } \Omega\,, \\ u = g & \text{on } \partial\Omega\,. \end{cases}$$

Namely, the Dirichlet's principle says that the solution of the Euler-Lagrange equation furnishes a minimizer of our problem. But the Euler-Lagrange equation is just a necessary condition, and do not ensure existence of minimizers! Nevertheless, this principle was accepted without questioning it by Riemann.

But not all the mathematicians were comfortable with it. Indeed, Weierstrass criticized it a lot, but he wasn't able to provide a counterexample (that does not exist!). In his seeking to disprove the Dirichlet principle, Weierstrass started to study the existence for variational problems, and showed that the minimum problem

$$\min_{\substack{u\in C^2([-1,1])\\ u(-1)=-1,u(1)=1}} \int_{-1}^1 x^2\big(u'(x)\big)^2 \, \mathrm{d}x\,,$$

has no solution. Years later, the Dirichlet principle was proved true by Hilbert, but mathematicians started to be skeptic about the optimistic belief that all minimum problems admit a solution, and started to investigate more in detail conditions ensuring the existence of such a solution. The existence of minimizers of variational problems became a central issue in the Calculus of Variations: how can we be sure that a minimum problem admits a solution?

[3] From falsehood, anything (follows).

[4] Here $\Omega \subset \mathbb{R}^N$ is an open set, and g is defined on $\partial\Omega$. We do not want to enter into the detailed assumptions we ask in order to ensure the well-posedness of the problem.

10.2. The end

To understand how the existence problem has been tacked, let us consider a more general setting. Take $F : X \to [-\infty, +\infty]$ to be our energy, defined over a normed space $(X, \|\cdot\|)$. Forget about the integral form of the energy. Consider the problem

$$\min_X F\,.$$

Set $m := \inf_X F$. If $m = +\infty$, then nothing interesting happens, since it means[5] that our functional is $F \equiv +\infty$. So, suppose $m < \infty$. Then, by definition of infimum, we know that there exists a sequence $(x_n)_n \subset X$ such that

$$\lim_{n\to\infty} F(m_n) = m\,.$$

Such a sequence is called a *minimizing sequence*, and we only know that $\sup_n F(x_n) < \infty$. What we would like to be able to say is that the following two conditions hold true:

(i) **compactness:** up to a (not relabeled subsequence) $x_n \to x$, with respect to the metric d induced by the norm $\|\cdot\|$, for some $x \in X$,

(ii) **lower semi-continuity[6]:** for all $y \in X$ and all $y_n \to y$ with respect to d, it holds

$$F(y) \le \liminf_{n\to\infty} F(y_n)\,.$$

With these two conditions in hand, we can conclude. Indeed, let us suppose that we are able to say that $x_n \to x$, for $x \in X$. Then

$$m \le F(x) \le \liminf_{n\to\infty} F(x_n) = \lim_{n\to\infty} F(y_n) = m\,.$$

That is, $F(x) = m$, and thus x is a solution to our minimization problem.

Let us try to apply the above general procedure to our specific problem. Let $(u_n)_n \subset \mathcal{A}$, where $\mathcal{A} := \{\, u \in C^2(\bar{\Omega}),\ u = g \text{ on } \partial\Omega \,\}$, be a minimizing sequence for the functional $\mathcal{F}$ over $\mathcal{A}$. Suppose

$$\sup_n \mathcal{F}(u_n) = \sup_n \frac{1}{2}\int_\Omega |\nabla u_n|^2 \,\mathrm{d}x < +\infty\,.$$

The fact that $u_n = g$ on $\partial\Omega$ allows us to use a Poincaré inequality[7] to say that

$$\sup_n \int_\Omega |u_n|^2 \,\mathrm{d}x < \infty\,.$$

Thus, we have that

$$\sup_n \|u_n\|^2_{H^1(\Omega)} = \sup_n \Big(\int_\Omega |u_n|^2 \,\mathrm{d}x + \int_\Omega |\nabla u_n|^2 \,\mathrm{d}x \Big) < \infty\,.$$

From this bound we would like to deduce that, up to a subsequence, $u_n \to u$ (in some sense), where $u \in \mathcal{A}$.

---

[5]Notice that it is possible that our functional is $F \equiv +\infty$ even if we didn't define it in a different way. For example, it is possible to construct a continuous function g on the unitary circumference such that the harmonic function in the unit ball having g has a boundary value has infinite Dirichlet energy. Since this function is the minimizer, then the whole functional turns out to be $+\infty$.

[6]This is weaker than asking the continuity of the energy. This is because we are only interested in minimum problem.

[7]A Poincaré inequality is when you control the L^2-norm of a function (or the function minus its mean value) with the L^2-norm of its gradient.

We now present two good reasons why we cannot hope that to happen. The first reason is peculiar to the norm on which we know bounds. Let us consider the sequence

$$v_n := \begin{cases} |x| & |x| > \frac{1}{n}\,, \\ \\ x\alpha_n(x) & |x| \le \frac{1}{n}\,, \end{cases}$$

where α_n is a smooth increasing functions with $\alpha_n(-\frac{1}{n}) = -1,\ \alpha_n(\frac{1}{n}) = 1$. Then, it is easy to see that

$$v_n \to v\,, \qquad \text{that is } \|v_n - v\|_{H^1} \to 0\,,$$

where

$$v(x) := |x|\,.$$

Thus, it is possible to exit the class of C^2 functions. In higher dimension it is also possible to converge to more (but not too much) ugly functions!

The second reason has a more general flavor:

Fact of life: the unit ball of a normed vector space is compact (with respect to the topology induced by the norm), *if and only if* the space if finite dimensional.

All the (interesting) spaces of functions are infinite dimensional. The end.

10.3. Always look at the bright side of math

Something that one should always remind is:

if you don't like something, just change it[8].

In the previous statement, we've seen that, if the topology used in order to check the compactness of the ball is the one induced by the norm, we have no hope to have compactness of the unit ball in infinite dimension.

Let's ask ourselves this: are we forced to used the same topology?
Of course no!
So, let's just change it!
You may ask: how?
Well, of course in such a way that the compacteness
and the lower semi-continuity properties hold true.
Is it always possible to do it?
It clearly depends on the problem.

For many variational problems, there is a *natural* topology, called the *weak topology*, weaker that the one induced by the norm (that, in contrast, is called *strong topology*) that makes the magic comes true, *i.e.*, for which the desired properties happen to hold true. We won't enter into the details of the weak convergence, but we want to stress that the choice of *a* weaker topology that is suitable for the problem is *not* a data of the problem, but it is something that you have to *choose*.

The strategy sketched above for proving the existence of a solution for a minimum problem is called *direct method*, and has been developed by Tonelli in the beginning of the 20^{th} century. Basically, it is the Weierstrass' theorem ensuring the existence of a maximum and a minimum

[8]Disclaimer: I'm not sponsoring surgical aesthetics! You're perfect as you are!

of a continuous function on a compact set of a finite dimensional space adapted to a more general context and specialized for the existence of minima (without caring about maxima).

It goes as follows: let $F : X \to [-\infty, +\infty]$ and *choose* on X a topology such that the following hold

- **compactness:** for any $t \in \mathbb{R}$, any sequence in the set $\{F \leq t\}$ admits a converging subsequence,
- **l.s.c.:** for any $x \in X$ and any $x_n \to x$ (w.r.t. the chosen topology), we have

$$F(x) \leq \liminf_{n\to\infty} F(x_n)\,.$$

Then the problem

$$\min_X F\,,$$

admits a solution.

Remark 10.1. We want notice that the requirements of compactness and lower semi-continuity (in brief l.s.c.) ask for *opposite* features of the weaker topology: compactness is more happy when there are few open sets (so it is easier to extract a finite covering), while lower semi-continuity requires a lot of open sets (indeed, a function F is l.s.c. if its lower level sets $\{F < t\}$ are open for all $t \in \mathbb{R}$). Thus, the topology *we have to choose* for the problem, has to be a good balance between the two requirements.

Remark 10.2. Notice that when we ask for compactness of bounded sequences and lower semi-continuity of the energy, we ask too much for our purposes. Indeed, what we really need in order to prove the existence of a minimizer is that, for *a* minimizing sequence $(x_n)_n$ we have $x_n \rightharpoonup x$ for some $x \in Y$ and that $F(x) \leq \liminf_n F(x_n)$. The problem is that, in general, we do not know explicitly minimizing sequences, and so we are force to ask the above conditions for *all* possible sequences. On the other hand, if for a *specific* problem we are able to prove that the above properties hold for a minimizing sequence, then we have done, since we can directly prove the existence of a minimizer. This strategy requires to have/prove a lot of additional information about the minimizing sequence.

In the sequel we will write $x_n \rightharpoonup x$ to denote that the sequence $(x_n)_n \subset Y$ converges with respect to the weak topology to x.

The choice of a weaker topology (even with the strong one) forces us to consider a more general space of admissible functions, since we've seen that the usual control of the functions of a minimizing sequence is not with the C^1 norm, but with a weaker one, and thus we can converge to a function that is no more C^1. Thus, we are forced to dealing with more general objects. This implies that also the meaning of a functional has to be generalized, as well as the concept of satisfying the boundary conditions.

In the example of the Dirichlet integral, we obtained a bound on the H^1 norm of any minimizing sequence. Thus, the space we have to consider is the completion (or closure) of $C^2(\Omega)$ with respect to the metric given by the norm $\|\cdot\|_{H^1}$. The space of functions obtained in this way is the so called *Sobolev space*[9] $H^1(\Omega) = W^{1,2}(\Omega)$. More in general, for $p \geq 1$, if

[9]The spaces on the sides of this equality are defined in different ways: H^1 is defined as the closure of C^2 (or C^∞) with respect to the H^1-norm, while the right-hand side is defined as the space of functions having some weak notion of derivatives (see [7]). The fact that these two spaces are the same has been proved in [8].

we want to consider the existence problem for a functional of the form

$$\int_\Omega |\nabla u|^p \, \mathrm{d}x \,,$$

we end up with the Sobolev space[10] $W^{1,p}(\Omega)$ if $p > 1$ and the space of functions of bounded variations $BV(\Omega)$ in the case $p = 1$.

The theory of Sobolev and BV spaces allows to give a meaning to

$$\int_\Omega |\nabla u|^2 \, \mathrm{d}x \,,$$

for functions $u \in H^1$ that are not in C^2, as well as what we mean by saying that $u = g$ on $\partial\Omega$ (that is, Sobolev and BV functions posses a *trace* on the boundary of nice sets).

10.4. (Don't) fly me to the moon

One requirement we ask for a variational functional $\mathcal{F}$ (or better, to the underlining topology) in order to behave well with respect to out minimal purposes, is to be lower semi-continuous. A big effort has been put in the study of conditions ensuring the l.s.c. with respect to the weak topology for functionals of the form

$$\mathcal{F}(u) := \int_\Omega f\big(x, u(x), Du(x)\big) \, \mathrm{d}x \,,$$

where $u : \Omega \to \mathbb{R}^M$, and for lagrangians depending on something else (*e.g.*, the determinant). It turns out that sufficient and necessary conditions for l.s.c. are different in the scalar ($M = 1$) and in the vectorial ($M > 1$) case. In the former one, the key condition on the lagrangian f is the *convexity*[11] in the last variable, *i.e.*, in the derivative. But in the vectorial case, despite it is still sufficient for having weak l.s.c. of $\mathcal{F}$, convexity is no more a necessary condition. A *weaker form* of this geometrical condition has been discover to be the correct notion for the vectorial case. It is called *quasi-convexity.* The idea behind this notion is the following: weak convergence can be thought as convergence of oscillating sequences. For example, the sequence $u_n(x) := \sin(nx)$ does not converge strongly to anything, but it converges weakly to $u \equiv 0$ (the *mean* of u_n). Quasi-convexity says that the lagrangian f prefers to 'kill oscillations of the derivatives'. That is, if we take an affine function u (whose gradient is constant!) and we (locally) perturb it with a regular function, then the energy of the perturbation cannot be lower than the energy of the affine one. This kind of phenomenon is related to what happen for convex lagrangians. Indeed, take $f : \mathbb{R}^N \to \mathbb{R}$ a convex function, fix $\xi \in \mathbb{R}^N$ and consider a perturbation $\varphi \in C_c^\infty(\Omega)$. Then, we have that

$$⨍_\Omega f(\xi + \nabla\varphi) \, \mathrm{d}x \geq f\Big(\xi + ⨍_\Omega \nabla\varphi \, \mathrm{d}x\Big) = f(\xi) \,,$$

where in the first step we used Jensen's inequality, while in the last one we used the fact that $\int_\Omega \nabla\varphi \, \mathrm{d}x = 0$ for all $\varphi \in C_0^1(\bar{\Omega})$. The quasi-convexity required the above inequality to holds true for all test functions φ (in a slightly larger space). It turns out that there are functions satisfying the above integral inequality without being convex (*e.g.*, the determinant). Moreover, further notions that can be seen as a weaker version of convexity, and that play an important role in many problems, have been studied for vectorial lagrangians (see [4]).

---

[10]Here 1 stands for the number of derivatives and p for the order of integrability.

[11]We've already encounter such a condition when we studied necessary and sufficient conditions for local minimizers in the scalar scale.

What should we do in case we are given a functional $F : X \to [-\infty, \infty]$ that is *not* weak l.s.c.?

Example. Let us consider the the case $X = [-1, 1]$ endowed with the Euclidean topology, and let us define the functional

$$F(x) := \begin{cases} 1 & x \in [0,1] \\ -x & x \in [-1,0) \,. \end{cases}$$

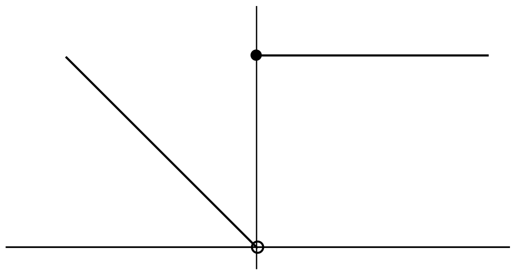

FIGURE 2. The function F of the example.

The 'problem' of the above function is that we have *spelled wrong* its value at $x = 0$. The *expected* value, from an 'existence of minimizers' point of view should have been $F(0) = 0$. Why? Well, because

$$\lim_{x \to 0^-} F(x) = 0 < F(0) = 1 \,.$$

Thus, among all way to approach $x = 0$ from a topological point of view (*i.e.*, with sequences $x_n \to x$), the *best* way from the *energetic* point of view is with a sequence $x_n \to x$ for which $\liminf_{n\to\infty} F(x_n)$ is the lower possible.

Keeping in mind the above example, we come back to the general case. Take a functional $F : X \to [-\infty, \infty]$, where X is the space of 'standard' objects, and suppose we want to define a functional $\bar{F} : Y \to [-\infty, \infty]$, where $Y \supset X$ is the space of 'generalized' objects, in such a way that $\bar{F}$ is the *best* functional obtained by F, that is lower semi-continuous [12]. So, we require that

$$\bar{F}(y) \leq \liminf_{n\to\infty} \bar{F}(y_n)\,, \tag{10.1}$$

for all $y \in Y$ and all $y_n \rightharpoonup y$. The problem of the above expression is that both sides are defined only when the points belong to X! So, it is difficult to check the above condition.

To over this difficulty, let us recall that the space Y is obtained by completing[13] the space X with respect to a norm $\|\cdot\|$, *i.e.*, Y is obtained by considering all the objects y for which there exists a sequence $(x_n)_n \subset X$ such that

$$\|x_n - y\| \to 0\,,$$

[12]Notice that this includes also the case when $X = Y$.

[13]In the same way as the space of real numbers can be obtained as the completion of the rational numbers with respect to the Euclidean norm.

Since the other topology we choose on Y is weaker than the one induced by $\|\cdot\|$, from the above convergence we also have $x_n \rightharpoonup y$. In particular, it is possible to view Y as the completion of X with respect to the weaker topology. We now have a way to rewrite condition (10.1). Let $y \in Y$ and take $(y_n)_n \subset Y$ such that $y_n \rightharpoonup y$. Since every object y_n in Y can be approximated in the weak topology with an object $x_n \in X$, condition (10.1) can be stated as

$$\bar{F}(y) \leq \liminf_{n\to\infty} \bar{F}(x_n) = \liminf_{n\to\infty} F(x_n)\,,$$

for all $y \in Y$ and for all $(x_n)_n \subset X$ such that $x_n \rightharpoonup y$. Notice that now, on the right-hand side, we have some known object.

Recalling the above example, we want the above condition to be sharp, in the sense that the value of $\bar{F}(y)$ has to satisfy (10.1) in the best way, *i.e.*, without being too small, without being too low with respect to all possible limits that can appear to the right-hand side. Thus, for $y \in Y$, we define the functional

$$\bar{F}(y) := \inf\left\{ \liminf_{n\to\infty} F(x_n) \;:\; (x_n)_n \subset X \text{ is such that } x_n \rightharpoonup y \right\}.$$

This functional is called the *relaxed functional* of F. By construction, $\bar{F}$ is l.s.c. with respect to the weak convergence, and turns out to be the *highest* l.s.c. functional that is below F.

Clearly, with the above expression is basically impossible, or rather, extremely difficult to perform any kind of analysis[14]. Thus, after the notion of relaxed functional has been developed, it should be good to be able to express $\bar{F}$ in a more manageable way. Maybe, one can hope that, if the starting functional F is of an integral form, also its relaxed functional $\bar{F}$ should maintained the same form, *i.e.*, can be expressed as an integral of some function. This problem goes under the name of *integral representation*.

10.5. U can't write this

We are happy! We faced the problem of existence and we developed a general strategy, the direct method, to face it. There's only a small, little issue to deal with: in order to prove existence we were forced to enlarge the space of admissible competitors and the solution of the (generalized) problem we are able to find lives in this bigger space. Usually, this is a space of objects that can be very wild. But we wonder whether a function in this space that happens to be a minimizer of some variational problem should present some sort of *regularity*.

Example. We've already encounter an example of an analytic lagrangian whose minimizers present singularities. This is what happen in the case $f(\xi) := (\xi^2 - 1)^2$. So, the regularity of the lagrangian is not enough in order to ensure regularity of the minimizers. Another kind of condition is required.

How to prove that a minimizer of a variational problem has some sort of regularity? Let us consider our Dirichlet energy

$$\mathcal{F}(u) := \int_\Omega \frac{|\nabla u|^2}{2}\,\mathrm{d}x\,.$$

One attempt to study properties of the minimizer is to use the fact that minimizers satisfy the Euler-Lagrange equation, that in our case is

$$\triangle u = 0 \qquad \text{in } \Omega\,. \tag{10.2}$$

[14] Try, for example, to write an Euler-Lagrange equation for the functional $\bar{F}$!

One should study the properties of solutions of the above equation in order to obtain some regularity properties for the minimizer of the Dirichlet energy.

There's only one small drawback: we cannot write the above equation! Indeed, recall that the space we are working in is the Sobolev space $H^1(\Omega)$. And, for such a space, we have no notion of second derivatives! But let's look at how we ended up with that equation. We took a local minimizer $u \in H^1(\Omega)$, a test function $\varphi \in C_c^\infty(\Omega)$, and we considered the variation $u + \varepsilon\varphi$. Then

$$0 = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathcal{F}(u+\varepsilon\varphi)_{|_{\varepsilon=0}} = \int_\Omega \nabla u \cdot \nabla\varphi \,\mathrm{d}x\,.$$

Now, for a function $u \in H^1$, the integral on the right-hand side makes sense. We only two steps missing in order to get (10.2) is integrate by parts and use the fundamental lemma. But integration by parts makes is forbidden in our case, since we need our function u to possess some sort of second derivatives.

The good news is that this is not such a big deal! Indeed, we can simply say that a function $u \in H^1(\Omega)$ satisfies (10.2) in the *weak sense* if

$$\int_\Omega \nabla u \cdot \nabla\varphi \,\mathrm{d}x = 0\,,$$

for all $\varphi \in C_c^\infty(\Omega)$. We can now study the properties of such a weak solutions in order to get regularity for minimizers of the Dirichlet energy.

This approach can be used also for dealing with general lagrangians

$$\mathcal{F}(u) := \int_\Omega f\big(x, u(x), Du(x)\big)\,\mathrm{d}x\,,$$

If $u : \Omega \to \mathbb{R}$ is a local minimizer of the minimum problem we know that, for any $\varphi \in C_c^\infty(\Omega)$, the followign integral equation holds

$$\int_\Omega f_p\big(x, u(x), Du(x)\big)\varphi \,\mathrm{d}x = -\int_\Omega f_\xi\big(x, u(x), Du(x)\big)\cdot\nabla\varphi\,\mathrm{d}x\,.$$

Thus, similarly to what we've done above, we say that an u (belonging to a Sobolev space that we do not specify here!) satisfying the above identity for all $\varphi \in C_c^\infty(\Omega)$ is a *weak solution* in Ω of the equation

$$f_p\big(x, u(x), Du(x)\big) = \mathrm{div}\big(\, f_\xi\big(x, u(x), Du(x)\big)\,\big)\,,$$

This is the other way Sobolev spaces were introduced: we say that a function $v \in L^p(\Omega)$ belongs to $W^{1,p}(\Omega)$ if there exists a function $w \in L^p(\Omega;\mathbb{R}^N)$ such that

$$\int_\Omega v\,\mathrm{div}\varphi\,\mathrm{d}x = -\int_\Omega w\cdot\varphi\,\mathrm{d}x\,,$$

for all $\varphi \in C_c^\infty(\Omega:\mathbb{R}^N)$. The function w is denoted by $\nabla u = (\partial_1 u, \ldots, \partial_N u)$, the *weak gradient*, and $\partial_i u \in L^p(\Omega)$ is called the i^{th} *weak derivative*.

But there are case where we *cannot* write the Euler-Lagrange equation, simply because the lagrangian f is not regular enough. So, different kind of techniques have to be used in order to prove regularity of minimizers. The milestone for this kind of investigation was the work of De Giorgi [2], related to the regularity of solutions to linear elliptic equation in divergence form. He proved the fact that solutions of the equation have to satisfy a 'reverse' Poincaré inequality (the so called *Caccioppoli type inequalities*) and used this fact to gain regularity of them. This kind of techniques has been extended by several authors to the case of non-linear elliptic equation in divergence form and thus to the study of regularity of minimizers of variational problems.

All we've said so far holds for *scalar* variational problems, *i.e.*, when $u : \Omega \to \mathbb{R}$. In the case of vectorial problems (everywhere) regularity is in general not expected, since vectorial problems naturally give rise to singularities. In this case, the aim is to prove *partial regularity*, that is, that the solution is regular everywhere but is a negligible set. Fine investigations on the *smallness* of this negligible sets are part of the current research.

A good overview of the regularity results in the Calculus of Variations is [9].

10.6. Let it go

At the beginning of the course we saw that it seems that Nature wants to minimize the energy, *i.e.*, prefers states with a low energy. But what happen when we have a system in a state whose energy is not the (locally) lowest possible? Well, according to our principle, Nature will lead the system to the state with (locally) lower energy possible. How?

Example. Let us consider a very simple example: take a cup a place a ball inside is at some height. Then, as soon as you'll let the ball to move, it will go directly to the lower point of the cup (ignore oscillations!), in the *fastest* way possible. That is, among all the ways to go from point P to point Q (see Figure 3), the ball will follow that path that minimize the energy *as fast as possible.*

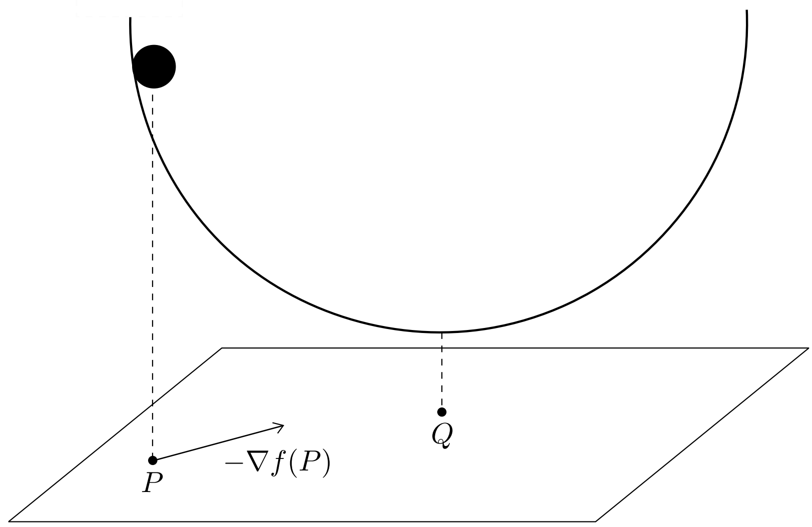

FIGURE 3. The motion of the ball is energy driven: at each point x, $-\nabla F(x)$ will be the velocity of the path.

Suppose that the profile of the cup is described by a function $f \in C^1(\mathbb{R}^2)$, and the energy we consider is the potential one (*i.e.*, the height of the ball, that ism the energy is f itself).

Then, since at each point $-\nabla f(x)$ is the direction of locally fasts decreasing of the energy f, the path that is followed is the one that, at each $x \in \mathbb{R}^2$ points in the direction of that vector (assuming it is not zero!). Let $\gamma : [0,1] \to \mathbb{R}^2$ be the path followed by the ball. Then, according to the above argument, we have that

$$\dot{\gamma}(t) = -\nabla f\big(\gamma(t)\big) \,. \tag{10.3}$$

Thus, the motion of the state $x \in \mathbb{R}^2$ is said to be *driven* by the energy.

The previous example can be easily generalized to a more general setting. In order to write (10.3), we need to give a meaning to ∇F when the space where F is defined is not $\mathbb{R}^N$. For, we better have a brief recalling in Analysis. When you have a function $g : \mathbb{R}^N \to \mathbb{R}$, you know how to define its *differential* $\mathrm{d}g(x) : \mathbb{R}^N \to \mathbb{R}^N$ at the point $x \in \mathbb{R}^N$. It is the *linear functional* such that $\mathrm{d}g(x)[v]$ is the directional derivative of g at x in the direction v. You know that you can *represent* the differential $\mathrm{d}g(x)$ with a vector $\nabla g(x)$ (called the *gradient*). That is:

$$\mathrm{d}g(x)[v] = \langle \nabla g(x), v \rangle \,,$$

for all $v \in \mathbb{R}^N$, where $\langle \cdot, \cdot \rangle$ denotes the usual scalar product on $\mathbb{R}^N$. Moreover, the above property characterize the gradient. Indeed, if $w \in \mathbb{R}^N$ is such that

$$\mathrm{d}g(x)[v] = \langle w, v \rangle \,,$$

for all $v \in \mathbb{R}^N$, then $w = \nabla g(x)$.

Now, consider a functional

$$\mathcal{F}(u) := \int_\Omega f\big(x, u(x), Du(x)\big) \,\mathrm{d}x \,,$$

defined over a space X. When we compute the variations in order to write the Euler-Lagrange equation, we are computing the *directional derivatives*. Than is, fixed a point $u \in X$, we consider a direction $\varphi \in C^\infty_c(\Omega)$, and we write

$$\delta\mathcal{F}(u)[\varphi] = \int_\Omega L_f(u)\varphi \,\mathrm{d}x \,.$$

We've seen in Chapter 7 that is possible to define a scalar product on the space of functions[15] $L^2(\Omega)$. It is defined as follows

$$\langle f, g \rangle_{L^2} := \int_\Omega fg \,\mathrm{d}x \,.$$

Thus, we obtain that

$$\delta\mathcal{F}(u)[\varphi] = \langle L_f(u), \varphi \rangle_{L^2} \,,$$

for all $\varphi \in C^\infty_c(\Omega)$. That is[16] we obtain that $L_f(u)$ is the *gradient* of $\mathcal{F}$ at u!

So, fix an initial point $u_0 \in X$. We define the *gradient flow* of F starting from u_0, as the curve[17] $\gamma : \mathbb{R} \to X$ solution of

$$\begin{cases} \dot{\gamma}(t) = -L_f\big(\gamma(t)\big) \,, \\ \gamma(0) = u_0 \,. \end{cases}$$

[15] We recall that a function u is said to be in $L^2(\Omega)$ whenever $\int_\Omega |u|^2 \,\mathrm{d}x < \infty$.

[16] Recall that we want to be sloppy!

[17] Here, for simplicity, we assume that the solution exists for all the times, *i.e.*, the curve γ is defined on the whole $\mathbb{R}$.

In a more fashion way, we can write the above system as

$$\begin{cases} \partial_t u = -L_f\big(u\big)\,, \\ u(0) = u_0\,. \end{cases}$$

Notice that the above motion stops when we reach a critical point, that is, when $L_f\big(u\big) = 0$. If we are lucky, this critical point will also be a minimum of the energy (true if the energy is convex). Usually, the above system happens to be *parabolic*.

A famous example of gradient flow is given when we want to minimize the Dirichlet energy:

$$\begin{cases} \partial_t u = -\triangle u\,, \\ u(0) = u_0\,, \end{cases}$$

that is the well-known heat equation. That is, when we solve the heat equation, we find a curve of functions $(u_t)_t$ that will (should) converge to a minimum point of the Dirichlet energy.

We now want to justify the meaning of the expression: the path that minimizes the energy *as fast as possible*. Consider a curve $(\gamma_t)_t$ in X and the function

$$t \mapsto \mathcal{F}\big(\gamma(t)\big)\,.$$

Then

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}\big(\gamma(t)\big) &= \langle L_f\big(\gamma(t)\big), \dot\gamma(t)\rangle_{L^2} \\ &\geq -\|L_f\big(\gamma(t)\big)\|_{L^2}\|\dot\gamma(t)\|_{L^2} \\ &\geq -\frac{1}{2}\|L_f\big(\gamma(t)\big)\|^2_{L^2} - \frac{1}{2}\|\dot\gamma(t)\|^2_{L^2}\,, \end{aligned}$$

where in the first inequality we used the Cauchy-Schwarz inequality, and in the last step we used $2ab \leq a^2 + b^2$. Recalling that the case of equality in the Cauchy-Schwarz inequality happens if and only if the two vectors are in the same direction, we see that if we take $\dot\gamma(t) = -L_f\big(\gamma(t)\big)$, we have equality in both the above inequalities. This justifies the fact that if we choose at each point the direction given by the gradient of the function, the energy will decrease faster than in all other directions. The choice of choosing exactly $-L_f\big(\gamma(t)\big)$ is dictated by the requirement (that we impose) to have

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}\big(\gamma(t)\big) = \|\dot\gamma(t)\|^2_{L^2} = \|L_f\big(\gamma(t)\big)\|^2_{L^2}\,.$$

This means that the speed with which we travel along the curve $\gamma([0,1])$ (the so called support of the curve) agrees with the behavior of the energy. This requirement can be justified as follows: in the above example, if the ball starts very closed to a minimum point, we do not expect it to move fast.

Finally, just to let you know: it is possible to develop the theory of gradient flows also in *metric* spaces. This study is motivated by the fact that many interesting energies are defined in such a spaces. Moreover, the connection between PDE and Wasserstein gradient flows, firstly pointed out in [6], has as a natural setting the one of metric spaces. The treatment is more involved.

10.7. Gamma mia

Motivation 1. Usually, mathematicians are happy just knowing that a solution to a minimum problem exists. Sometimes it is also possible to prove analytically some properties of these minimizers. And very rarely an explicit[18] solution is available. Thus, since the are communities of scientists that really need to work with these objects, numerical simulations or approximations are needed. The fact is that some of the energies used by mathematicians are difficult to implement numerically. *Approximated* energies are required, in such a way that their minimizers approximate the ones of the sharp one.

Motivation 2. There are situations when the energy of a physical system depends on a small parameter. For example, consider a system made by two immiscible fluids in a container. A simple version of the energy of such a system (without gravity) is the one that takes into consideration the interaction of the two kind of molecules happening at the interface between the two liquid. Usually, this interface is not a sharp one, but a diffuse one, in the sense that the two liquids are completely separated everywhere but in a region of size ε, where transition happens. From a mathematical point of view the energy F_ε of this system is not easy to study. Thus, it would be preferable to have a *limiting* energy that is easier to study, but still capturing all the relevant physical features we are interested in.

The two situations presented above motivate the study of a notion of variational convergence for functionals. Uniform convergence of the functionals is a too strong requirement for our purposes, and thus something weaker is needed. In the 70s De Giorgi developed the theory of Γ-convergence[19], that is the notion of variational convergence used nowadays. The idea is that, given a sequence of functionals F_n, on one hand we want the 'profile' of the limiting functional F to be energetically better that the limit of the profiles of the F_n's, while on the other hand we want the profile of F to be energetically reachable with the ones of the F_n's. This is because we want convergence of the minima. So, what we want is the the value of F at a point x is the *limiting value of the infima of the F_n's close to x*. That is, we want[20]

$$F(x) = \inf\left\{ \liminf_{n\to\infty} F_n(x_n) \,:\, x_n \to x \right\}.$$

If the above is true, we say that the sequence F_n Γ-convergence to F.

Remark 10.3. The above requirements are very similar to the ones we asked for the relaxed functional, that is indeed a particular case of Γ-convergence, where we consider the constant sequence $F_n := F$. Yes, with this notion of convergence it may happen that a constant sequence has a limit that is different from itself[21]. Get used to it!

The Γ-convergence allows for phenomena that are neglected by other stronger notions of convergence. For instance, it is possible for the sequence $(F_n)_n$ to oscillates and also to explode, and still Γ-converging to some (bounded) energy . For example, the sequence of functions

$$F_n := \begin{cases} n\chi_{[0.1]} & n \text{ odd} \,, \\ \chi_{[1.2]} & n \text{ even} \,, \end{cases}$$

will Γ-converge to $F \equiv 0$, as well as the sequence $F_n(x) := 1 + \sin(nx)$ does.

[18] Philosophical question: do you think that π is an *explicit* number?

[19] The classical reference for Γ-convergence is [1]

[20] To be precise, the notion of Γ-convergence must be defined in relation to an underlining metric.

[21] In particular, this implies that the Γ-convergence cannot be induced by a norm.

10.8. Everything's gonna be alright

Let us ask a natural question. Consider a sequence of energies $(F_\varepsilon)_\varepsilon$ that Γ-converges to an energy F. Let $(u_t^\varepsilon)_t$ be a solution to the gradient flow generated by F_ε, and $(u_t)_t$ be a solution of the gradient flow generated by F. Do they converge?

$$\begin{tikzcd}
F_\varepsilon \arrow[r, "\Gamma\text{-converge}"] \arrow[d, leftrightarrow] & \mathrm{F} \arrow[d, leftrightarrow] \\
u_t^\varepsilon \arrow[r, "?"] & u_t
\end{tikzcd}$$

This is easily seen to be false! Indeed, Γ-convergence is a convergence at the level of the functions, but not at the level of its derivatives/gradient. Since gradient flows are defined using the gradient of the function, we have no reason why to expect such a convergence.

Example. Since we did not introduce properly the notion of Γ-convergence, we cannot present the following example from the Γ-convergence point of view. But we want to give an idea of what is going on. For, we will use the uniform convergence[22] in place of the Γ-convergence. Let us consider the function $F(x) := x^2$. Choose a point $\bar{x} \neq 0$ and consider the following perturbation of f:

$$F_n(x) := \begin{cases} F(x) & x \notin \left[\bar{x} - \frac{2}{n}, \bar{x} + \frac{2}{n}\right], \\ \bar{x}^2 + (x - \bar{x})^2 & x \in \left[\bar{x} - \frac{1}{n}, \bar{x} + \frac{1}{n}\right], \\ \text{linear interpolation} & \text{otherwise} . \end{cases}$$

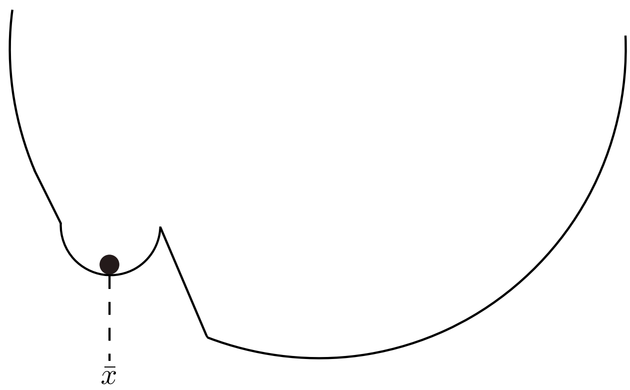

FIGURE 4. The gradient flow of F_n starting from $\bar{x}$ is stationary.

[22] That, in turn, implies the Γ-convergence.

Then, clearly f_n converges to f uniformly. The problem of these f_n's is that the slope of f at $\bar{x}$ is not zero, while it is zero for all the f_n's. Thus, if we consider the gradient flow generates by each f_n starting from $\bar{x}$ it will not move, since the derivative is zero. On the other hand, the gradient flow generated by f starting from $\bar{x}$ will actually move.

The idea is that we also need, in some sense, the gradients to converge, that is, an higher order version of the Γ-convergence. This theory has been developed in [10]. The main requirements for having the Γ-convergence of the gradient flows are (basically) two: that we cannot approximate the slope of F at a point x with lower slopes of F_ε at x, and that something similar happens for the time derivatives of u_t^ε and u_t.

10.9. Is this just a game?

We would like to conclude these notes with a legitimate question:

does anybody use this stuff?

If you study in the US, you are even more motivated to ask this question, since Calculus of Variations is the red headed step child[23] from Europe.

The answer is: yes and no! Let's explain myself in a better way[24]. The aim of Science is to understand Nature[25]. Mathematics is not a science, but it plays its (big) part in the game. Since there is no *a priori* correct mathematical tool to study a problem, you can choose the one you prefer. And the Calculus of Variations is one among them. It has pro and cons, as every tool does. In my opinion, when correctly used, it has a very beautiful way of describing the equilibrium states of a system, as well as in the way of proving existence for evolution equations (with the so called *minimizing movements* method). Of course, not everybody likes it and/or use it. It's a matter of personal taste and, sometimes, of degree of rigorousness you want/need.

So, there is no will of convincing the reader that Calculus of Variations is the most beautiful way to study physical problems[26]. We just want to conclude by saying that these theory has been used in a variety of different situations: from problems in material science to imaging theory, from the isoperimetric problem to the variational formulation of quantum mechanics[27].

[23]I have to thanks Ian for letting me be aware of this beautiful expression!

[24]What follows is my personal point of view, so not all the readers may agree with it.

[25]This is a sort of *Classical knowledge theory*, since we are giving from grant that we can understand Nature. Maybe one day there will be a *Direct method to knowledge theory*!

[26]Even if it is!

[27]See [3]

CHAPTER 11

Appendix

11.1. The Gauss-Green theorem

We would like to recall the Gauss-Green theorem. This theorem is a particular case of the more general Stokes theorem, that can be seen as a generalization of the fundamental theorem of calculus. Before stating the theorem we start by recalling some definitions.

DEFINITION 11.1. Let $\gamma : [a, b] \to \mathbb{R}^2$. We call $\gamma([0, 1])$ a *curve*, and γ a *parametrization* of the curve.

Sometimes, with an abuse of notation, we will talk about the curve γ, but remember that the object we are interested in is the image of the parametrization, not the parametrization itself. Recall that, when we talk about the regularity of a curve we mean that there exists a parametrization of the curve having the stated regularity.

DEFINITION 11.2. We say that a curve is *closed* if $\gamma(a) = \gamma(b)$. We say that a curve is *simple* if γ is an injective function on $[a, b]$, with possible exception at b (in the case the curve is closed).

DEFINITION 11.3. For a curve $\gamma : [a, b] \to \mathbb{R}^2$ of class C^1 and a function $f : \mathbb{R}^2 \to \mathbb{R}$, we define the *line integral* of f over γ as

$$\int_\gamma f := \int_a^b f\big(\gamma(t)\big)|\gamma'(t)| \, \mathrm{d}t \, .$$

Now suppose to have a particle moving along a curve γ, and suppose that F is a force field on the plane. We would like to compute the work done by F on the particle. This physical requirement motivates the following definition.

DEFINITION 11.4. Let $\gamma : [a, b] \to \mathbb{R}^2$ be a curve of class C^1 and let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be a vector field. We define the *integral of F over γ* as

$$\int_\gamma F := \int_\gamma (F_1 \, \mathrm{d}x + F_2 \, \mathrm{d}y) := \int_a^b F\big(\gamma(t)\big) \cdot \gamma'(t) \, \mathrm{d}t \, .$$

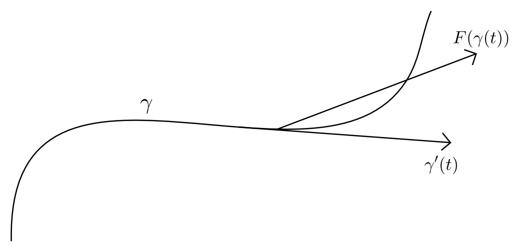

FIGURE 1. At each point of the curve γ we consider the scalar product between the tangent vector and the vector field at that point

Remark 11.5. Notice that the above objects are invariant under reparametrization of the curve γ, *i.e.*,

$$\int_a^b F\big(\gamma_1(t)\big)\cdot\gamma_1'(t)\ \mathrm{d}t=\int_a^b F\big(\gamma_2(t)\big)\cdot\gamma_2'(t)\ \mathrm{d}t\,,$$

$$\int_a^b f\big(\gamma_1(t)\big)|\gamma_1'(t)|\ \mathrm{d}t=\int_a^b f\big(\gamma_2(t)\big)|\gamma_2'(t)|\ \mathrm{d}t\,,$$

whenever $\gamma_1:[c,d]\to\mathbb{R}^2$ and $\gamma_2:[e,f]\to\mathbb{R}^2$ are two parametrizations of the curve $\gamma([a,b])$. Thus, the line integral is well defined.

It is easy to see that all the two above definitions can be extended to the case of a piecewise-C^1 curve γ.

Remark 11.6. Let γ be a simple closed curve. We call such a curve a *Jordan* curve. It is easy to believe (but not trivial to prove - try!) that a Jordan curve divides the plane $\mathbb{R}^2$ in two connected components, an interior one and an exterior one.

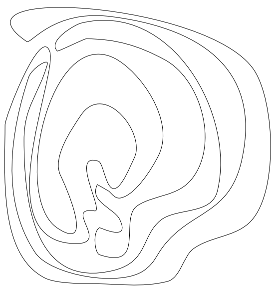

Figure 2. An example of a Jordan curve

We are now in position to state the Gauss-Green theorem.

Theorem 11.7. *Let $\gamma:[a,b]\to\mathbb{R}^2$ be a simple closed curve of class C^1 and take a vector field $F:\mathbb{R}^2\to\mathbb{R}^2$ of class C^1. Let us denote by E the interior region delimited by γ. Then it holds*

$$\int_E\Big(\frac{\partial F_2}{\partial x}-\frac{\partial F_1}{\partial y}\Big)\ \mathrm{d}x\ \mathrm{d}y=\int_\gamma(F_1\ \mathrm{d}x+F_2\ \mathrm{d}y)\,.$$

Proof. *Step 1*: we first prove that

$$\int_E\frac{\partial F_1}{\partial y}\ \mathrm{d}x\ \mathrm{d}y=-\int_\gamma F_1\ \mathrm{d}x\,,$$

for simple regions E of the type

$$E:=\{(x,y)\in\mathbb{R}^2\ :\ x\in[c,d]\,,\ f_1(x)\le y\le f_2(x)\}\,,$$

where $f_1,f_2:[c,d]\to\mathbb{R}$ are two functions of class C^1, with $f_2>f_1$ on (c,d). In this case we can split the curve γ into four pieces γ_i, $i=1,2,3,4$ as in the figure.

Thus, by using Fubini's theorem, we have that

$$\int_E\frac{\partial F_1}{\partial y}\ \mathrm{d}x\ \mathrm{d}y=\int_c^d\int_{f_1(x)}^{f_2(x)}\frac{\partial F_1}{\partial y}\ \mathrm{d}y\ \mathrm{d}x=\int_c^d\big(F_1(x,f_2(x))-F_1(x,f_1(x))\big)\ \mathrm{d}y\,.$$

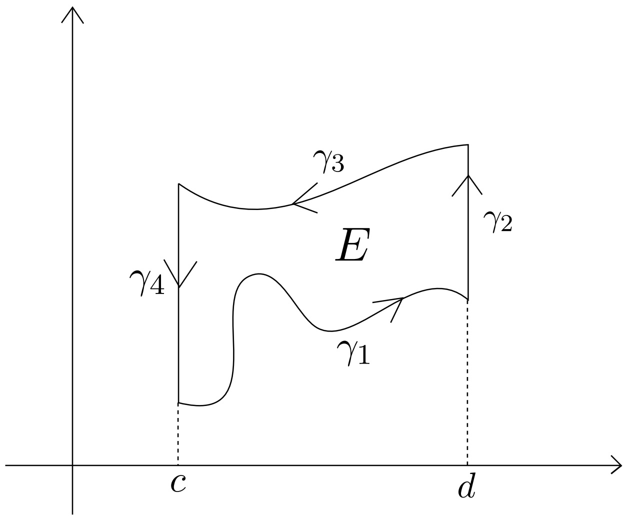

FIGURE 3. An example of a simple region

On the other hand, by using the following parametrizations for the curve γ:

$$\begin{array}{ll} \gamma_1 : [c,d] \to \mathbb{R}^2 & \gamma_1(t) := (t, f_1(t))\,, \\ \gamma_2 : [f_1(d), f_2(d)] \to \mathbb{R}^2 & \gamma_2(t) := (d, t)\,, \\ -\gamma_3 : [c,d] \to \mathbb{R}^2 & -\gamma_3(t) := (t, f_2(t))\,, \\ \gamma_4 : [f_1(c), f_2(c)] \to \mathbb{R}^2 & \gamma_4(t) := (c, f_2(c) - t(f_1(c) - f_2(t)))\,, \end{array}$$

where $-\gamma_3$ is the curve γ_3 oriented in the opposite way, we have that

$$\int_{\gamma_1} F_1 \ \mathrm{d}x = \int_c^d F_1(t, f_1(t)) \ \mathrm{d}t\,,$$

$$\int_{\gamma_2} F_1 \ \mathrm{d}x = \int_{\gamma_4} F_1 \ \mathrm{d}x = 0\,,$$

$$\int_{\gamma_3} F_1 \ \mathrm{d}x = -\int_c^d F_1(t, f_2(t)) \ \mathrm{d}t\,.$$

Step 2: in a similar way it is possible to prove that

$$\int_E \frac{\partial F_2}{\partial x} \ \mathrm{d}x \ \mathrm{d}y = \int_\gamma F_2 \ \mathrm{d}y\,,$$

for simple regions E of the type

$$E := \{(x,y) \in \mathbb{R}^2 \,:\, y \in [c,d]\,,\ f_1(y) \le x \le f_2(y)\}\,,$$

where $f_1, f_2 : [c,d] \to \mathbb{R}$ are two functions of class C^1, with $f_2 > f_1$ on (c,d).

Step 3: now we can conclude by dividing the region E in simple region of one of the previous type, obtaining the result just by summing up all the terms. □

COROLLARY 11.8. *Let γ be a simple closed curve of class C^1. Then the area of the interior region E delimited by γ is*

$$\mathcal{A}(E) = \int_\gamma x \ \mathrm{d}y = -\int_\gamma y \ \mathrm{d}x = \frac{1}{2}\int_\gamma (-y \ \mathrm{d}x + x \ \mathrm{d}x)\,.$$

11.2. A characterization of convexity

We want to prove a characterization of C^1 convex functions.

LEMMA 11.9. *Let $f : \mathbb{R}^N \to \mathbb{R}$ be a function of class C^1. Then f is convex if and only if*

$$f(y) \geq f(x) + \nabla f(x) \cdot (y - x)\,,$$

for all $x, y \in \mathbb{R}^N$.

PROOF. *Necessity*: we divide the proof of necessity in two steps.
Step 1. We claim that

$$\frac{f(y) - f(x)}{|y - x|} \geq \frac{f(z) - f(x)}{|z - x|} \tag{11.1}$$

for all $x, y \in \mathbb{R}^N$ and each $z \in \mathbb{R}^N$ that lies in the segment from x to y. Geometrically the above inequality is clear[1] (see Figure 4).

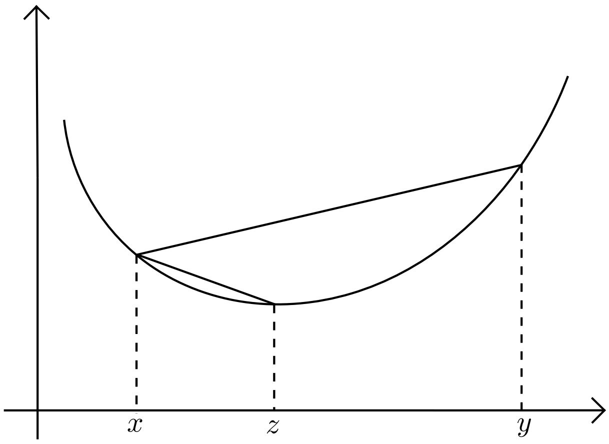

FIGURE 4. The claim of Step 1 is easily seen to be true by looking at the figure

It is clear that we can reduce ourselves to the one dimensional case. So, let us take $x, y \in \mathbb{R}$ and let us consider a point $z \in [x, y]$. It is possible to write z as follows

$$z = \frac{z - x}{y - x} y + \Big(1 - \frac{z - x}{y - x}\Big) x\,. \tag{11.2}$$

By convexity

$$f(z) \leq \frac{z - x}{y - x} f(y) + \Big(1 - \frac{z - x}{y - x}\Big) f(x)\,.$$

This is the desired inequality.

Step 2. By (11.1) we have that

$$f(y) \geq f(x) + \frac{f(z) - f(x)}{|z - x|} |y - x|\,.$$

[1]This is one case where a picture is definitely more direct than words!

By letting $z \to x$ we have that

$$\frac{f(z) - f(x)}{|z - x|} \to \nabla f(x) \cdot \frac{y - x}{|y - x|},$$

and thus we conclude.

Sufficiency: as before, we can assume $x, y, z \in \mathbb{R}$ and that $x < z < y$. We know that

$$f(x) \geq f(z) + \nabla f(z)(x - z),$$

and that

$$f(y) \geq f(z) + \nabla f(z)(y - z).$$

By writing

$$y - z = -\frac{x - z}{|x - z|}|y - z|,$$

from the above inequalities, we get

$$f(x) \geq f(z) + (f(z) - f(y))\frac{|x - z|}{|y - z|}.$$

Thus, by using (11.2), we conclude. □

11.3. Regularity of the boundary of a set

In this section we introduce the main objects we will use during the course.

We start by introducing the notion of regularity for the boundary of a set.

Definition 11.10. We say that a set $\Omega \subset \mathbb{R}^N$ has a boundary $\partial\Omega$ *of class* C^k, for $k \in \mathbb{N} \setminus \{0\}$, if for every point $\bar{x} \in \partial\Omega$ there exists $r > 0$, $\delta > 0$ and a function $\Psi : \mathbb{R}^{N-1} \to \mathbb{R}$ of class C^k such that, up to a rotation and a translation, the following two conditions are satisfied:

(i) $\partial\Omega \cap B_r(\bar{x}) = \left\{ \left(x', \Psi(x')\right) \in \mathbb{R}^{N-1} \times \mathbb{R} \ : \ |x'| < \delta \right\}$,

(ii) $\Omega \cap B_r(\bar{x}) = \left\{ (x', x_N) \in \mathbb{R}^{N-1} \times \mathbb{R} \ : \ |x'| < \delta, y > \Psi(x') \right\}$.

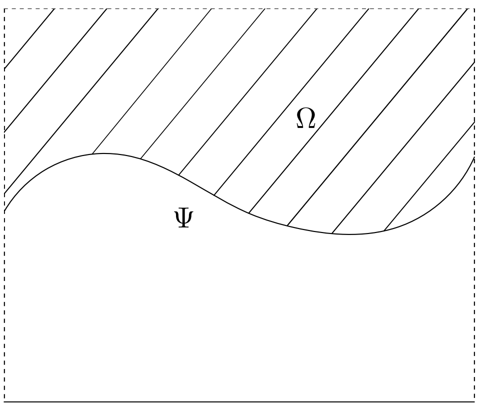

Figure 5. The situation described in Definition 11.10.

11.4. Derivative of the determinant

Let us consider the following situation: we are given, for $i = 1, \dots, N$, functions $A_i : (-\varepsilon, \varepsilon) \to \mathbb{R}^N$ of class C^1 and we consider the $N \times N$ matrix function A whose i^{th} column is given by A_i. We define the function

$$f(t) := \det A(t)\,.$$

We would like to compute the derivative of this function. To do so, let us recall that the determinant is a multi-linear function, *i.e.*,

$$\det(v_1, \dots, v_i + w_i, \dots, v_N) = \det(v_1, \dots, v_i, \dots, v_N) + \det(v_1, \dots, w_i, \dots, v_N)\,,$$

for any vectors $v_1, \dots, v_N, w_1 \dots, w_N \in \mathbb{R}^N$. Thus

$$f'(\bar{t}) = \sum_{i=1}^{N} \det\big(A_1(\bar{t}), \dots, A_i'(\bar{t}), \dots, A_N(\bar{t})\big)\,.$$

In particular, if $A_i(0) = \mathrm{Id}$ for all $i = 1, \dots, N$, we have that

$$A_i(t) = \mathrm{Id} + tB_i + o(\varepsilon)\,,$$

where $B = A_i'(0)$, and thus

$$f'(0) = \sum_{i=1}^{N} B_i^i\,.$$

11.5. Small perturbations of the identity

In this section we want to prove that small variations of the identity are diffeomorphisms. Let us state it in a clear way

THEOREM 11.11. *Let $\Omega \subset \mathbb{R}^N$ be an open bounded set and let $f : \bar{\Omega} \to \bar{\Omega}$ be a C^1 function such that $f = \mathrm{Id}$ in a neighborhood of partialΩ. Then there exists $\varepsilon_0 > 0$ such that if $\|f - \mathrm{Id}\|_{C^1(\bar{\Omega};\mathbb{R}^N)} < \varepsilon_0$, then f is a diffeomorphism from $\bar{\Omega}$ onto itself.*

PROOF. *Injectivity:* we have that

$$\begin{aligned}
|f(x) - f(y)| &= \Big|\Big(\int_0^1 Df\big(x + t(y-x)\big)\,\mathrm{d}t\Big) \cdot (y-x)\Big| \\
&\geq \Big(\Big|\int_0^1 \mathrm{Id}\,\mathrm{d}t\Big| - \Big|\int_0^1 \big(Df\big(x+t(y-x)\big) - \mathrm{Id}\big)\,\mathrm{d}t\Big|\Big)\,|y-x| \\
&\geq \Big(1 - \int_0^1 \big|Df\big(x+t(y-x)\big) - \mathrm{Id}\big|\,\mathrm{d}t\Big)\,|y-x| \\
&\geq (1 - \|Df - \mathrm{Id}\|_{C^0})|x-y|\,.
\end{aligned}$$

Thus, injecticity is true provided $\|Df - \mathrm{Id}\|_{C^0} < 1$.

Invertibility of the Jacobian matrix: by the continuity of the determinant (recall that $\bar{\Omega}$ is compact), it holds that there exists $\varepsilon_0 > 0$ such that if $\|Df - \mathrm{Id}\|_{C^0} < \varepsilon_0$, then $Df(x)$ is invertible for each $x \in \bar{\Omega}$.

Surjectivity: let us suppose by contradiction that there exists $z \in \bar{\Omega}$ that does not belong to the image of $\bar{\Omega}$ through f. By hypothesis $z \in \Omega$, because close to the boundary of Ω, f is the identity.

Claim: there exists $r > 0$ such that $B_r(z) \notin f(\Omega)$.
Indeed, if by contradiction there exists $(z_n)_n \subset \Omega$ with $z_n = f(x_n)$, $x_n \in \Omega$, and $z_n \to z$, then (up to a subsequence), $x_n \to x \in \bar{\Omega}$ and thus, by continuity of f, $z = f(x)$.

We are now in position to conclude, since by letting $\bar{z} \in \bar{\Omega}$ be such that[2]

$$\|\bar{z} - z\| = \min\{\, \|f(x) - z\| \,:\, x \in \bar{\Omega}\,\}\,,$$

we would get that $Df(\bar{z})$ is not surjective, because $\bar{z} - z \neq 0$ cannot be in its image. □

11.6. The Poincaré lemma

Here we just briefly recall the basic notions about 1-forms. You can think to 1-forms as the generalization of the concept of differential: given a function $f \in C^1(R^N)$, we know that $\mathrm{d}f(x)$, the *differential* of f at $x \in \mathbb{R}^N$, is a linear map on $\mathbb{R}^N$. Thus, the map $x \mapsto \mathrm{d}f(x)$ is a continuous map from $\mathbb{R}^N$ to the set of linear maps on $\mathbb{R}^N$. By recalling that every linear map on $\mathbb{R}^N$ can be identify with a vector of $\mathbb{R}^N$ itself, we can think to the differential as a map from $\mathbb{R}^N$ into itself $x \mapsto \omega(x)$. We generalize this notion in the following

Definition 11.12. A 1-*form* on $\mathbb{R}^N$ is a map ω from $\mathbb{R}^N$ into the space of linear functions on $\mathbb{R}^N$.

Thanks to the discussion above, we can identify the value of a 1-form ω at a point x with a vector $w(x) \in \mathbb{R}^N$. In this way, we have that

$$\omega(x)[v] = w(x) \cdot v\,.$$

Since 1-forms are generalization of the notion of differential, we can ask ourselves how much we enlarged the space of differentials. That is, when is true that a 1-form turns out to be a differential?

Definition 11.13. A 1-form ω is said to be *exact* if $\omega = \mathrm{d}f$, for some function f.

Let us now take a function $f \in C^2(\mathbb{R}^N)$. Its differential is identified at each point with the gradient ∇f, that is

$$\mathrm{d}f(x)[v] = \nabla f(x) \cdot v\,.$$

We know that $\partial_i \partial_j f = \partial_j \partial_i f$ for all $i, j = 1, \ldots, N,$. In particular, this implies that a *necessary* condition for a 1-differential form to be exact is to have

$$\partial_i w_j = \partial_j w_i\,, \tag{11.3}$$

for all $i, j = 1, dots, N$, where w is the vector identifying ω.

Definition 11.14. A 1-form ω is said to be *closed* if (11.3) holds true.

The relation between being closed and convex is clarified in the following result.

Lemma 11.15 (Poincaré lemma). *Let ω be a 1-form defined on an open and* simply connected *set of $\mathbb{R}^N$. If ω is closed, then ω is also exact.*

[2]Notice that such a minimum exists, by compactness of $\bar{\Omega}$.

Bibliography

[1] G. Dal Maso, *An introduction to Γ-convergence*, Progress in Nonlinear Differential Equations and their Applications, 8, Birkhäuser Boston, Inc., Boston, MA, 1993.

[2] E. De Giorgi, *Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari*, Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. (3), 3 (1957), pp. 25–43.

[3] R. P. Feynman, *Space-time approach to non-relativistic quantum mechanics*, Rev. Modern Physics, 20 (1948), pp. 367–387.

[4] I. Fonseca and G. Leoni, *Modern methods in the calculus of variations: L^p spaces*, Springer Monographs in Mathematics, Springer, New York, 2007.

[5] M. Giaquinta and S. Hildebrandt, *Calculus of variations. I*, vol. 310 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 1996. The Lagrangian formalism.

[6] R. Jordan, D. Kinderlehrer, and F. Otto, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal., 29 (1998), pp. 1–17.

[7] G. Leoni, *A first course in Sobolev spaces*, vol. 105 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2009.

[8] N. G. Meyers and J. Serrin, $H = W$, Proc. Nat. Acad. Sci. U.S.A., 51 (1964), pp. 1055–1056.

[9] G. Mingione, *Regularity of minima: an invitation to the dark side of the calculus of variations*, Appl. Math., 51 (2006), pp. 355–426.

[10] E. Sandier and S. Serfaty, *Gamma-convergence of gradient flows with applications to Ginzburg-Landau*, Comm. Pure Appl. Math., 57 (2004), pp. 1627–1672.

[11] A. Treibergs, *Inequalities that imply the isoperimetric inequality*.