Lecture 1
Introduction and characterization of critical points

Luca Nenna

November 26, 2024

Contents

Important: this lecture is based on Antonin Chambolle's notes on Calculus of Variations.

1 Introduction

The calculus of variations is the study of the minimizers or critical points of "functionals," which are functions defined in spaces of infinite dimensions, typically functional spaces.

This needs to adapt the notions of differential calculus. The stationarity of a functional $\mathcal{E}(u)$ is "simply" characterized by the equation:

$$\mathcal{E}'(u) = 0 \tag{1.1}$$

which, in general, will be a partial differential equation (PDE) in u (or something more general).

A goal of the calculus of variations is to "solve" such PDEs: more precisely, to show that they actually have one or several solutions (or none...), study their properties, and possibly design numerical methods to compute these solutions or approximations.

A first important observation is that not all PDEs will be solved by a variational analysis: only the PDEs which are "variational," meaning that their equation is precisely of the form (1.1) for a particular $\mathcal{E}$.

1.1 Why is it interesting?

- It sometimes provides a very simple tool for showing the existence of (weak) solutions to a problem.
- Many PDEs come from problems in physics, mechanics, etc., and precisely from "variational" principles and are therefore (often minimizing) critical points of some physical energy.
- Many problems in the industry (or finance, etc.) are designed as finding the "best" state according to some criterion, and their solution is precisely a minimizer, or maximizer, of this criterion ("optimization").

1.2 Standard Examples

1. Laplace equation $-\Delta u = 0$ characterizes the critical points of the "Dirichlet energy":
$$\int |\nabla u|^2 \, dx.$$
In this case, since the energy is convex, critical points and minimizers are the same.
2. Horizontal elastic membrane subject to a vertical force f:
$$\min_{u=0 \text{ on } \partial\Omega} \frac{1}{2} \int_\Omega |\nabla u|^2 \, dx - \int_\Omega f u \, dx,$$
in this case equation (1.1) reads:
$$\begin{cases} -\Delta u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$
Variant with a nonlinear potential energy:
$$\min_{u=0 \text{ on } \partial\Omega} \frac{1}{2} \int_\Omega |\nabla u|^2 \, dx + \int_\Omega F(u) \, dx,$$
then the equation becomes $-\Delta u + F'(u) = 0$.
3. Nonlinear elasticity:
$$\mathcal{E}(u) = \int_\Omega W(\nabla u) \, dx,$$
where now $u : \Omega \subset \mathbb{R}^3 \to \mathbb{R}^3$ is vectorial-valued. W might depend on $\nabla u^T \nabla u$, $\det \nabla u$, etc. Equation (1.1) becomes now a system.

4. Minimal surfaces, geodesics:
$$\mathcal{E}(u) = \int_\Omega \sqrt{1 + |\nabla u|^2}\, dx \quad \text{(area of the graph of } u\text{)},$$
$$\mathcal{E}(u) = \int_0^1 |u'(x)|^2\, dx,$$
with $u(0) = x_0, u(1) = x_1, u(x) \in M$ a manifold for each x.

For each of these problems, the natural questions are: is there existence of a solution? Uniqueness? How can it be characterized? How can it be computed?

2 Characterization of the Critical Points

This section addresses the issue of the meaning of equation $\mathcal{E}'(u) = 0$.

2.1 Differentiability

Definition 2.1. Let X be a Banach space and U an open subset. A functional $\mathcal{E} : U \subset X \to \mathbb{R}$ is *Fréchet-differentiable* at $u \in U$ if there exists a continuous linear form $D\mathcal{E}(u) \in X^*$, called the *differential*, such that:
$$\boxed{\mathcal{E}(u + v) = \mathcal{E}(u) + D\mathcal{E}(u) \cdot v + o(\|v\|_X).}$$
The functional $\mathcal{E}$ is of class $C^1(U)$ if $u \mapsto D\mathcal{E}(u)$ is continuous.

Definition 2.2. $\mathcal{E}$ is said to be *Gâteaux-differentiable* at $u \in U$ in the direction $v \in X$ if:
$$D_v E(u) = \frac{d}{dt}\mathcal{E}(u + tv)\big|_{t=0}$$
exists. It is said to be Gâteaux-differentiable at u if this derivative exists for all $v \in X$.

Clearly, if $\mathcal{E}$ is (Fréchet-)differentiable at u, then it is Gâteaux-differentiable, and $D_v E(u) = D\mathcal{E}(u) \cdot v$ for all v. The converse is not true.

Example 2.3.
1. $u(x, y) = \frac{x^3}{x^2+y^2}$ (and $u(0) = 0$): the Gâteaux derivative in direction $(a, b) \neq 0$ is the function:
$$\frac{a^3}{a^2 + b^2},$$
which is not linear in (a, b).
2. $u(x, y) = \frac{x^2 y}{x^4+y^2}$ ($\sqrt{x^2 + y^2} \leqslant 1/2$, and $u(0) = 0$): the Gâteaux derivative is 0. However, if $(x, y) = (t, t^2) \to 0$ as $t \to 0$,
$$\frac{u(x, y)}{\sqrt{x^2 + y^2}} = \frac{t^4}{2t^4} = \frac{1}{2} \neq 0.$$

The Fréchet-differential $D\mathcal{E}(u)$ is also denoted $\mathcal{E}'(u)$, and by definition, a critical point u is a point where $\mathcal{E}'(u) = 0$.

2.2 First Variation of a Functional

In practice, to derive the stationarity conditions of an energy, and in particular the minimality conditions, it is enough to know how to compute directional derivatives for all "admissible" v. Assume that u is a minimizer of $\mathcal{E}$ over a set $K \subset X$. Then given v, provided $u+tv \in K$ for $t > 0$ small (which is the meaning of an "admissible variation" v), one always has $\mathcal{E}(u+tv) \geqslant \mathcal{E}(u)$. Thus:

$$\lim_{t \to 0^+} \frac{\mathcal{E}(u+tv) - \mathcal{E}(u)}{t} \geqslant 0.$$

If $\mathcal{E}$ is differentiable at u, one recovers that $D\mathcal{E}(u) \cdot v \geqslant 0$. If both v and $-v$ are admissible, one recovers $D\mathcal{E}(u) \cdot v = 0$.

For a general theory, we restrict ourselves to functionals of the form:

$$\boxed{\mathcal{E}(u) = \int_\Omega \mathcal{L}(x, u(x), Du(x))\, dx,} \tag{2.2}$$

where $\Omega \subset \mathbb{R}^n$ is a bounded open set, and $u : \Omega \to \mathbb{R}^m$ is a possibly vectorial-valued function (in some functional space). Here Du is the differential, identified as an $m \times n$ matrix with entries $\frac{\partial u_i}{\partial x_\alpha}$, $i = 1, \ldots, m$, $\alpha = 1, \ldots, n$. The function $\mathcal{L} : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m \times n} \to \mathbb{R}$, $(x, u, p) \mapsto \mathcal{L}(x, u, p)$, is called the *Lagrangian*, and is assumed to be smooth.

Given u, v, and $t > 0$ small, one has:

$$\mathcal{E}(u+tv) = \mathcal{E}(u) + t \int_\Omega \left(\sum_i \frac{\partial \mathcal{L}}{\partial u_i}(x, u, Du) v_i + \sum_{i,\alpha} \frac{\partial \mathcal{L}}{\partial p^i_\alpha}(x, u, Du) \frac{\partial v_i}{\partial x_\alpha} \right) dx + o(1).$$

Hence:

$$\frac{d}{dt} \mathcal{E}(u+tv)\big|_{t=0} = \int_\Omega \left(\sum_i \frac{\partial \mathcal{L}}{\partial u_i}(x, u, Du) v_i + \sum_{i,\alpha} \frac{\partial \mathcal{L}}{\partial p^i_\alpha}(x, u, Du) \frac{\partial v_i}{\partial x_\alpha} \right) dx.$$

At a critical point, one should have for all admissible v:

$$\int_\Omega \left(\sum_i \frac{\partial \mathcal{L}}{\partial u_i}(x, u, Du) v_i + \sum_{i,\alpha} \frac{\partial \mathcal{L}}{\partial p^i_\alpha}(x, u, Du) \frac{\partial v_i}{\partial x_\alpha} \right) dx = 0.$$

If all smooth $v = (v_1, \ldots, v_m)$ with compact support are admissible, one can integrate by parts the last term, yielding the following form for the equation $\mathcal{E}'(u) = 0$:

$$\frac{\partial \mathcal{L}}{\partial u_i}(x, u, Du) - \sum_{\alpha=1}^n \frac{\partial}{\partial x_\alpha} \left(\frac{\partial \mathcal{L}}{\partial p^i_\alpha}(x, u, Du) \right) = 0, \tag{2.3}$$

for all $i = 1, \ldots, m$, at least in the distributional sense.

Definition 2.4. The Euler-Lagrange equation (or system) associated with $\mathcal{E}$, given by:

$$\mathcal{E}(u) = \int_\Omega \mathcal{L}(x, u, Du)\, dx,$$

is:
$$\frac{\partial \mathcal{L}}{\partial u_i}(x,u,Du) = \sum_{\alpha=1}^{n} \frac{\partial}{\partial x_\alpha}\left(\frac{\partial \mathcal{L}}{\partial p_\alpha^i}(x,u,Du)\right) = \operatorname{div}\left(\frac{\partial \mathcal{L}}{\partial p_\alpha^i}(x,u,Du)\right),$$
for all $i = 1, \dots, m$.

Example 2.5.

1. For $\mathcal{E}(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 dx + \int_\Omega F(u)dx$, $\mathcal{L}(x,u,p) = \frac{|p|^2}{2} + F(u)$, and one recovers the equation:
$$\Delta u = F'(u).$$
2. For $\mathcal{L}(x,u,p) = (1/q)|p|^q, 1 < q < +\infty$, the equation is the q-Laplace equation:
$$\operatorname{div}\left(|\nabla u|^{q-2}\nabla u\right) = 0.$$
3. For $\mathcal{L}(x,u,p) = \sqrt{1+|p|^2}$, the equation is the minimal surface equation:
$$\operatorname{div}\left(\frac{\nabla u}{\sqrt{1+|\nabla u|^2}}\right) = 0.$$

2.3 Boundary Conditions

The Euler-Lagrange equations derived earlier are not complete. For example, if we consider the equation $-\Delta u = 0$, which characterizes the critical points of:
$$\int_\Omega |\nabla u|^2 dx, \tag{2.4}$$
it has infinitely many solutions: any harmonic function in Ω is a solution. However, if we consider the problem:
$$\min_{u \in X} \int_\Omega |\nabla u|^2 dx, \tag{2.5}$$
where $X = H^1(\Omega)$ or $X = H_0^1(\Omega)$, then most harmonic solutions are not solutions of the problem (neither minimizers nor critical points). Indeed, the minimal value for this problem is 0, reached for $u = 0$, while in general if u is harmonic, the integral will not vanish. This highlights the necessity of boundary conditions.

2.3.1 The Prescribed Boundary Conditions

The *Dirichlet* or prescribed boundary conditions arise from restrictions imposed on the space of definition of the functional $\mathcal{E}$. These conditions are, strictly speaking, not part of the Euler-Lagrange equation $\mathcal{E}'(u) = 0$ but are embedded in the condition $u \in X$.

Example 2.6. Consider the problem:
$$\min_{u \in g + H_0^1(\Omega)} \int_\Omega |\nabla u|^2 dx, \tag{2.6}$$

where $g \in H^1(\Omega)$. This ensures the existence of at least one function u with finite energy, otherwise the problem cannot have a solution.

In this case, the equation solved by a minimizer is:

$$\begin{cases} -\Delta u = 0 & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega. \end{cases} \tag{2.7}$$

Here, the first line comes from the stationarity of $\mathcal{E}$ at u, while the second arises because we have prescribed $u \in g + H_0^1(\Omega)$.

The admissible variations v are functions in $H_0^1(\Omega)$. Indeed for such functions, $u + tv$ is not modified on $\partial\Omega$. Thus, variations do not satisfy $v = g$, but rather $v = 0$ on the boundary.

Warning: If the Dirichlet condition $u = g$ is not the trace of an H^1 function on $\partial\Omega$, the integral $\int_\Omega |\nabla u|^2 dx$ becomes infinite, and the variational problem cannot be solved. However, this does not mean that the PDE (2.7) itself is unsolvable. For example, in the planar disk B_1, there exists an harmonic function $u \in C^\infty(B_1)$ with $\Delta u = 0$, $u(\cos\theta, \sin\theta) = \sin(\theta/2)$, $\theta \in [-\pi, \pi]$. But there is no function in $H^1(B_1)$ with a jump discontinuity on the boundary.

Example 2.7 (The Bi-Laplacian)**.** Consider:

$$\min_{u \in g + H_0^2(\Omega)} \int_\Omega |\Delta u|^2 dx, \tag{2.8}$$

for a given $g \in H^2(\Omega)$. Recall that $H_0^2(\Omega)$ is the closure of $C_c^\infty(\Omega)$ in $H^2(\Omega)$ The Euler-Lagrange equation is:

$$\begin{cases} \Delta^2 u = 0 & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega, \\ \nabla u = \nabla g & \text{on } \partial\Omega. \end{cases} \tag{2.9}$$

The latter condition arises since functions in H_0^2 have a vanishing gradient on the boundary (assuming the boundary smooth enough).

2.3.2 The Free Boundary Conditions

The *Neumann* or free boundary conditions are not derived from restrictions on u, but from the equilibrium equation $\mathcal{E}'(u) = 0$ written in the variational sense $D\mathcal{E}(u) \cdot v = 0$ and the fact that the variations v are allowed to vary on a part of the boundary.

Example 2.8. Consider the problem:

$$\min_{u \in H^1(\Omega)} \int_\Omega |\nabla u|^2 + (u - g)^2 dx. \tag{2.10}$$

Considering now variations $u + tv$, $u \in H^1(\Omega)$, we can easily find

$$\int_\Omega \nabla u \cdot \nabla v + (u - g) v \, dx = 0, \tag{2.11}$$

for any $v \in H_0^1(\Omega)$. Consider first $v \in C_c^\infty(\Omega)$ then integrating by parts the first term we get

$$\int_\Omega (-\Delta u + u - g) v \, dx = 0,$$

and since it must be true for all test function/variations we have that

$$-\Delta u + u = g, \text{ in } \Omega.$$

Now consider $v \in C^\infty(\overline{\Omega})$. Integrating by parts, we find:

$$\int_\Omega (-\Delta u + u - g) v \, dx + \int_{\partial\Omega} v \nabla u \cdot \nu \, d\sigma = 0, \tag{2.12}$$

where ν is the unit normal vector to $\partial\Omega$. Since we already know that $-\Delta u + u = g$ we find that $\nabla u \cdot \nu = 0$. Thus:

$$\begin{cases} -\Delta u + u = g & \text{in } \Omega, \\ \nabla u \cdot \nu = 0 & \text{on } \partial\Omega. \end{cases} \tag{2.13}$$

Notice that the second equation does not arise from the restrictions on the function space, but it is indeed encoded in the first variation condition $\mathcal{E}'(u) = 0$, that is the admissible variations v are free to vary on the boundary.

General Form. For a functional of the form (2.2)

$$\mathcal{E}(u) = \int_\Omega \mathcal{L}(x, u, Du) \, dx, \tag{2.14}$$

we find, in addition to (2.3), that for $v \in C^\infty(\overline{\Omega})^m$, integrating by parts

$$\begin{aligned} \int_\Omega \left(\frac{\partial \mathcal{L}}{\partial u^i}(x, u, Du) - \sum_{\alpha=1}^n \frac{\partial}{\partial x_\alpha} \left(\frac{\partial \mathcal{L}}{\partial p_\alpha^i}(x, u, Du) \right) \right) v^i \, dx \\ + \int_{\partial\Omega} v^i \sum_{\alpha=1}^n \left(\frac{\partial \mathcal{L}}{\partial p_\alpha^i}(x, u, Du) \right) \nu_\alpha = 0, \end{aligned}$$

for all $i = 1, \dots, m$. Since (2.3) holds the first integral vanishes and the second must vanish for all test functions we deduce for all $i = 1, \dots, m$

$$\left\langle \left(\frac{\partial \mathcal{L}}{\partial p_\alpha^i}(x, u, Du) \right)_{1 \leqslant \alpha \leqslant n} | \nu \right\rangle = 0. \tag{2.15}$$

2.4 A Remark on Null Lagrangian

Consider again the functional in the standard form (2.2).

Definition 2.9 (Null Lagrangian)**.** $\mathcal{L}$ is called a *Null Lagrangian* if, for any $u : \Omega \to \mathbb{R}^m$ smooth enough, the Euler-Lagrange equation is satisfied, that is

$$\frac{\partial \mathcal{L}}{\partial u_i}(x, u, Du) - \sum_{\alpha=1}^n \frac{\partial}{\partial x_\alpha} \left(\frac{\partial \mathcal{L}}{\partial p_\alpha^i}(x, u, Du) \right) = 0.$$

It means that any smooth u is a critical point of the energy. In this case, given u, v regular enough with $u = v$ on the boundary $\partial\Omega$, one has $\mathcal{E}(u) = \mathcal{E}(v)$. Take $f(t) = \mathcal{E}(u + t(v-u))$ for $t \in [0,1]$, then (rough abuse of notation)

$$f'(t) = D\mathcal{E}(u) \cdot u - v.$$

Using (2.3) and that $u = v$ on the boundary it follows that $f'(t) = 0$, hence f is constant and $f(0) = f(1)$.

Example 2.10 (Does it exist?)**.** If $m = 1$, the only Null Lagrangians are of the form:

$$\mathcal{L}(x, u, p) = a \cdot p + b(x),$$

where a is a constant vector. For $m > 1$, there exist less trivial examples.

2.5 Constrained Problems, Obstacles

When there are constraints in a minimization problem, the Euler-Lagrange equation is not generally valid. This section illustrates this with two significant cases: convex constraints and differentiable constraints.

2.5.1 Convex Constraints

Consider the *obstacle problem*:

$$\min_{u \geqslant \psi} \mathcal{E}(u) := \int_\Omega |\nabla u|^2 \, dx, \tag{2.16}$$

where the function u is constrained to belong to the convex set $C = \{u \in H_0^1(\Omega) : u \geqslant \psi\}$. Instead of considering variations of the form $u + tv$ (How to choose v?), it is more convenient to consider another candidate $v \in C$ and variations of the form $u + t(v-u) \in C$. For a minimizer u, this leads to:

$$D\mathcal{E}(u) \cdot (v - u) \geqslant 0, \quad \forall v \in C, \tag{2.17}$$

which is known as a *variational inequality*. For (2.16), this becomes:

$$\int_\Omega \nabla u \cdot \nabla (v - u) \, dx \geqslant 0, \quad \forall v \in C. \tag{2.18}$$

- In the set $\{u > \psi\}$, the constraint $u \geqslant \psi$ is inactive. Choosing $v = u \pm t\varphi$, where $\varphi \in C_c^\infty(\{u > \psi\})$, gives:

$$-\Delta u = 0 \quad \text{in } \{u > \psi\}. \tag{2.19}$$

- In the set $\{u = \psi\}$, the constraint $u \geqslant \psi$ is active. For any nonnegative $\varphi \geqslant 0$, the variational inequality implies:

$$\int_\Omega \nabla u \cdot \nabla \varphi \, dx \geqslant 0, \quad \forall \varphi \geqslant 0. \tag{2.20}$$

  In other words we have that

$$\boxed{-\Delta u \geqslant 0,} \tag{2.21}$$

  which means that u is superharmonic.

One can show that it is a nonnegative Radon measure and write the Euler-Lagrange equation:

$$\begin{cases} -\Delta u = \mu, & \mu \geqslant 0, \\ u = 0, & \text{on } \partial\Omega, \\ \mu(\{u > \psi\}) = 0. \end{cases} \tag{2.22}$$

Here, μ is a nonnegative Radon measure concentrated on the set where $u = \psi$.

A second typical example is the following variational problem

$$\boxed{\min_{\mu \in \mathcal{P}(\Omega)} \mathcal{E}(\mu) := \int_{\Omega\times\Omega} w(x,y)\, d\mu(x)\, d\mu(y),} \tag{2.23}$$

where $\mathcal{P}(\Omega)$ is the set of probability measures on Ω and w is nonnegative, symmetric interaction (e.g. $w(x,y) = \frac{1}{|x-y|}$). If μ is the minimizer one finds that for any ν

$$\int_{\Omega\times\Omega} w(x,y)\, d\mu(x)\, d(\nu - \mu)(y) \geqslant 0$$

and deduce that there exists a constant c such that $\int_{\Omega\times\Omega} w(x,y)\, d\mu(x) = c$.

2.5.2 Differentiable Constraints

Now consider a problem with a *differentiable constraint*:

$$\boxed{\min_{\mathcal{F}(u)=0} \mathcal{E}(u),} \tag{2.24}$$

where $\mathcal{F} : X \to \mathbb{R}$ is a smooth functional representing the constraint. A minimizer should satisfy $D\mathcal{E}(u) \cdot v = 0$ in all directions which are tangent to the manifold $\mathcal{F}(u) = 0$ that is with $D\mathcal{F}(u) \cdot v = 0$. In other words there exists a Lagrange multiplier $\lambda \in \mathbb{R}$ such that

$$\boxed{D\mathcal{E}(u) = \lambda D\mathcal{F}(u).} \tag{2.25}$$

Example 2.11. Consider:

$$\min \left\{ \int_\Omega |\nabla u|^2\, dx \mid u \in H_0^1(\Omega),\ \int_\Omega u^2\, dx = 1 \right\}. \tag{2.26}$$

Here:

$$\mathcal{E}(u) = \int_\Omega |\nabla u|^2 dx, \quad \mathcal{F}(u) = \int_\Omega u^2 dx - 1.$$

Notice that for a general $v \in H_0^1(\Omega)$, there is no reason for which one would have $\int_\Omega (u + tv)^2\, dx = 0$. However one can consider

$$t \mapsto \frac{u + tv}{\|u + tv\|_{L^2}}$$

and since u is a minimizer we get

$$\frac{\int_\Omega |\nabla u + t\nabla v|^2\, dx}{\int_\Omega |u + tv|^2\, dx} \geqslant \int_\Omega |\nabla u|^2\, dx$$

for all $v \in H_0^1(\Omega)$. Developing we get

$$f(t) := \frac{\int_\Omega |\nabla u|^2\,dx + t^2 \int_\Omega |\nabla v|^2\,dx + 2t \int_\Omega \nabla u \cdot \nabla v\,dx}{1 + 2t \int_\Omega uv\,dx + t^2 \int_\Omega v^2\,dx} \geqslant \int_\Omega |\nabla u|^2\,dx$$

and it is clear that the left-hand side term is minimized for $t = 0$, that is $f'(0) = 0$ and we finally obtain

$$\int_\Omega \nabla u \cdot \nabla v\,dx - \left(\int_\Omega |\nabla u|^2\,dx \right) \int_\Omega uv\,dx = 0$$

The Euler-Lagrange equation becomes:

$$-\Delta u = \lambda u, \tag{2.27}$$

where $\lambda = \int_\Omega |\nabla u|^2 dx$.