University of Padua
Department of Mathematics “Tullio Levi-Civita”

Lecture Notes on Calculus of Variations 2025-2026

Orlando Ferrari
University of Padua

Version May 26, 2026

"Learn from those who came before, and lay the trail **for those who come after**."
— Gustave and Lune (Clair Obscur: Expedition 33)

Preface

These lecture notes were created as a study companion for my final exam in the Calculus of Variations course, taught by Prof. Guido De Philippis. Honestly, I found the lectures challenging to follow at times, but his enthusiasm kept my curiosity alive, motivating me to compile these notes for my own future reference.

Please forgive any inconsistencies throughout the text. I compiled this entire document under a strict two-week time constraint, and I simply did not have the bandwidth to polish every single detail.

In the spirit of full transparency: my original handwritten notes were not always perfectly sharp on the underlying assumptions. Due to the limited preparation time, several parts of this document were written with the assistance of Gemini 3.1 Pro. For instance, the statement of the Riemann-Lebesgue Lemma in Lecture 5 was corrected by Gemini, and much of the raw LaTeX code was generated by scanning my handwritten drafts. However, I deliberately wrote out the proofs myself (with some structural polishing). Letting an AI generate the proofs would have defeated the entire purpose of this project, which is for me to truly master the material.

As a disclaimer, I was dealing with some health issues while writing the latter half of these notes (from Lecture 14 onward). Because I was running out of time to prepare for my other exams, the LaTeX coding and the structural flow in those later sections rely much more heavily on Gemini's assistance as a desperate measure. Even so, I take full responsibility for any errors in this document, as everything is ultimately based on my own original handwritten notes. Additionally, I only include the notes from 22 lectures given since the last two lectures about the Plateau problem is optional and I simply unable to write them down here. However, I attach the handwritten note in my homepage for the sake of completion.

I would be very happy to discuss this topic further or receive corrections for any errata you might find. You can reach me via the contact details on my homepage at refrainfr.github.io.

Padova
6 May 2026 — Orlando Ferrari

Table of Contents

1 Some Calculus of Variations Problems

This lecture serves as a starting point for the topic of Calculus of Variations, we started from listing classical variational problems which serves as the motivation of the topic itself. The formulation of the problems itself are mostly derived from physical phenomenon.

Isoperimetric Problems

Among all sets $E \subset \mathbb{R}^d$ with a fixed volume $|E| = V_1$, find the one that minimizes the perimeter.

Instance of the Problem in $\mathbb{R}^2$: Let $\partial E = \gamma(S^1)$, where $\gamma : S^1 \to \mathbb{R}^2$ is a C^1 closed curve. We define the Area $A(\gamma)$ and Length $L(\gamma)$ as:

$$L(\gamma) = \int |\dot{\gamma}| dt$$

Problem:

$$\min \left\{ L(\gamma) : \gamma : S^1 \to \mathbb{R}^2, \; \frac{1}{2} \int \gamma \wedge \dot{\gamma} = c \right\}$$

Answer: Sphere (Proved by De Giorgi, '50s).

Brachistochrone Problem

Find a curve γ such that a mass m following along γ will take minimal time. Let $P = (0, 0)$ and $Q = (\overline{x}, \overline{y})$, with the curve represented as a graph $\gamma = \text{graph}(u)$.

Conservation of energy for derivation of the problem:

$$\frac{1}{2} m v^2 = g(-u(x)) \implies v = \sqrt{\frac{2g(-u)}{m}}$$

Note that $\dfrac{ds}{dt} = v \implies dt = \dfrac{ds}{v}$. With $ds = \sqrt{1 + (u')^2}\, dx$, we get the total time T by integrating:

$$T = \int_0^{\overline{x}} \frac{\sqrt{1 + (u')^2}}{\sqrt{2g(-u)/m}}\, dx$$

Problem Formulation (for $g/m = 1$):

$$\min \left\{ \int_0^{\overline{x}} \frac{\sqrt{1 + (u')^2}}{\sqrt{-u}}\, dx \;\middle|\; -u : u(0) = 0, u(\overline{x}) = \overline{y} \right\}$$

Geodesics Problem

Let $\gamma : [0,1] \to M$ be a curve on a Riemannian Manifold (M, g) such that $\gamma(0) = P$ and $\gamma(1) = Q$. The length functional is:

$$l(\gamma) := \int_0^1 \sqrt{g_{\gamma(t)}(\dot{\gamma}, \dot{\gamma})}\, dt$$

Where $g_x : T_xM \times T_xM \to \mathbb{R}$ is the scalar product, and $x \mapsto g_x$ is smooth.
Problem:

$$\min \left\{ \int_0^1 \sqrt{g_{\gamma(t)}(\dot{\gamma}, \dot{\gamma})}\, dt \;\middle|\; \gamma \right\}$$

This problem is mathematically analogous to the **Fermat Principle** in optics, where light minimizes travel time through a medium with refraction coefficient $n(x)$:

$$T = \int_0^1 n(\gamma(t)) \cdot |\dot{\gamma}(t)|\, dt$$

Exercise: Lifeguard Problem

Find the optimal point of entrance for a lifeguard to reach a target in minimal time across two mediums with speeds V_1 and V_2. **Solution:** Snell's Law applies:

$$\frac{\sin\alpha}{V_1} = \frac{\sin\beta}{V_2}$$

Optimal Tackling Problem

Determine the optimal direction $(x(t), y(t))$ for a sailboat to reach a finish line given a specific wind velocity field. Let $V_{wind} = \dfrac{1 - \cos(V_\alpha)}{2}$ and $V_R = \left(1 - \dfrac{y}{L}\right)^2$. With boundary conditions $y(0) = 0$ and $y(T) = 0$, the problem is maximizing the integral of horizontal velocity:

$$\int_0^1 V_{wind} + V_R\, dt = \int_0^1 \frac{1 - \cos(h \cdot \arctan(\dot{y}(t)))}{2} + \left(1 - \left(\frac{y(t)}{L}\right)^2\right) dt$$

which is equivalent to minimize

$$\int_0^1 \frac{\cos(h \cdot \arctan(\dot{y}(t))) - 1}{2} + \left(\frac{y(t)}{L}\right)^2 - 1\, dt$$

we find that an infimizing sequence exists (e.g., oscillating at 45° angles), but a smooth minimum is never reached due to unaccounted variables.

Plateau Problem (Soap Bubbles)

Given a boundary $\Gamma \subset \mathbb{R}^d$, look for a surface Σ such that $\partial\Sigma = \Gamma$ and the area of Σ is minimal.

Problem:

$$\min\{\text{Area}(\Sigma) : \partial\Sigma = \Gamma\}$$

Note that as one changes the meaning of surface area and boundary, the problem changes as well.

First Formalization (Graphs of a Map): We restrict to d-dimensional surfaces in $\mathbb{R}^N = \mathbb{R}^d \times \mathbb{R}^{N-d}$ that are graphs of $u : \Omega \subset \mathbb{R}^d \to \mathbb{R}^{N-d}$.

$$\min \left\{ \int_\Omega \sqrt{1 + \sum_{w=1}^{\min(d,N-d)} (M_w(Du))^2 \, dx} \;\middle|\; u = g \text{ on } \partial\Omega \right\}$$

Where $(M_w(Du))^2$ is the sum of squared w-minors of the Jacobian $Du \in \mathbb{R}^{N-d} \otimes \mathbb{R}^d$.

In particular, for $d = 2$ in $\mathbb{R}^4$: $\text{Area}(\text{graph } u) = \int_\Omega \sqrt{1 + |Du|^2 + (\det Du)^2} \, dx$ since we can compute for $u(x, y) = (u_1(x, y), u_2(x, y))$ that

$$Du = \begin{bmatrix} \frac{\partial u_1}{\partial x} & \frac{\partial u_1}{\partial y} \\ \frac{\partial u_2}{\partial x} & \frac{\partial u_2}{\partial y} \end{bmatrix}$$

and thus:

$$(M_1(Du))^2 = |Du|^2 \qquad (M_1(Du))^2 = (\det Du)^2$$

Second Formulation (Parametric Approach): Let $\Sigma = \varphi(D)$, where $\varphi : D \subset \mathbb{R}^2 \to \mathbb{R}^3$ (or $\mathbb{R}^d$, or (M, g)) and $\varphi(\partial D) = \Gamma$.

$$\min \left\{ \int_D |\partial_x \varphi \wedge \partial_y \varphi| \, dxdy \;\middle|\; \varphi : \partial D \to \Gamma \text{ is a parametrization} \right\}.$$

Third Formulation (Surfaces of Revolution):

$$\min \left\{ 2\pi \int_{-1}^1 u\sqrt{1 + (u')^2} \, dx \;\middle|\; u(-1) = u(1) = \alpha \right\}$$

Solution: Catenoid (rotation of the hyperbolic cosine).

Catenary and Elastic Membrane Problems

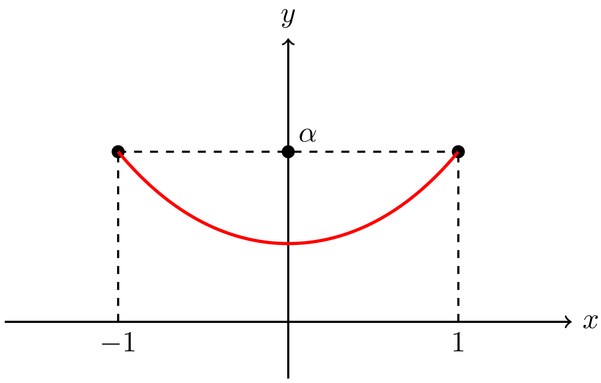
Catenary Problem: Minimize the potential energy $\int_{-1}^{1} y\,dl$ of a chain of fixed length

$$L = \int_{-1}^{1} \sqrt{1 + (u')^2}\,dx$$

with the formulation of such problem is

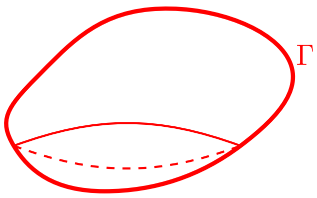
$$\min \left\{ \int_{-1}^{1} u\sqrt{1 + (u')^2}\,dx \;\middle|\; u(-1) = u(1) = \alpha \right\}$$

Solution: Hyperbolic cosine.



Elastic Membrane (Obstacle Problem): Expanding $\sqrt{1 + |\nabla u|^2} \sim 1 + \frac{1}{2}|\nabla u|^2$, the potential energy simplifies to the Dirichlet energy minus the load. If confined by a table at $z = 0$:

$$\min \left\{ \int_{\Omega} \left(\frac{|\nabla u|^2}{2} - gu \right) dx \;\middle|\; u = f \text{ on } \partial\Omega,\ u \geq 0 \text{ on } \Omega \right\}$$



2 The Euler Lagrange Equation

Before starting out, the list of references that is relevant to this course can be consulted in the Bibliography. Following the contents of this course, the most relevant references are [Dac07] and [Bra24]. The reader is advised to consult those two books since this note also follows from those sources, in addition to the lectures given.

Back to Calculus of Variations, we are in particular interested in problems of the form

$$\min \left\{ \int_\Omega F(x, u(x), Du(x))\, dx \;\middle|\; u \in \chi \right\}$$

for $F : \Omega \times \mathbb{R}^\ell \times (\mathbb{R}^\ell \otimes \mathbb{R}^d)$ where

$$\chi = \{u : \Omega \subseteq \mathbb{R}^d \to \mathbb{R}^\ell \text{ satisfying constraints}\}$$

Typically there are three types of constraints considered, they are given as follows:

1. **Boundary value constraint:** $u|_{\partial\Omega}$ is prescribed.
2. **Isoperimetric type constraint:** $\displaystyle\int_\Omega G(x, u, Du)$ is constant for some G.
3. **Holonomic constraint:** the image of u is presribed. An example of this case is the prescribed condition for geodesics problem.

Moreover, we may refer F as Lagrangian, integrand, or energy density, we will stay flexible with convention, but Lagrangian will be emphasized more throughout this notes.

§2.1 Deriving Euler-Lagrange Equations

For a finite dimensional function $f : \mathbb{R}^d \to \mathbb{R}$, if $\bar{x}$ is a minimal point, then

$$\frac{\partial f}{\partial r}(\bar{x}) = 0 \qquad \text{and} \qquad \frac{\partial^2 f}{\partial r^2} \geq 0$$

for any unit direction $r \in S^{d-1}$ assuming $\dfrac{\partial f}{\partial r}(\bar{x})$, $\dfrac{\partial^2 f}{\partial r^2}(\bar{x})$ exists.

Assume $\bar{u}$ is a minimizer for $\mathcal{F}(u) = \displaystyle\int_\Omega F(x, u, Du)\, dx$ and $\{v_\epsilon\}_{\epsilon \in (-\epsilon_0, \epsilon_0)}$ be a family of variations, i.e., a family of maps such that

$$\begin{cases} v_0(\bar{x}) = \bar{u}(x) & \forall x \in \Omega \\ v_\epsilon \in \chi & \forall |\epsilon| < \epsilon_0. \end{cases}$$

We define $g(\epsilon) = \mathcal{F}(v_\epsilon)$, then $g(\epsilon)$ has a minimum at 0. If g is differentiable, indeed $\left.\dfrac{d}{d\epsilon} g(\epsilon)\right|_{\epsilon=0} = 0$. Otherwise, we still have

$$\liminf_{\epsilon \to 0^+} \frac{g(\epsilon) - g(0)}{\epsilon} \geq 0 \text{ (slope is positive)}$$

In particular, we consider the case of

$$\chi = \{u : \Omega \subset \mathbb{R}^d \to \mathbb{R}^\ell : u|_{\partial\Omega} = g, u \in C^1\}$$

Fix $\varphi \in C_c^\infty(\Omega; \mathbb{R}^\ell)$, we consider the variation $v_\epsilon(x) = \bar{u}(x) + \epsilon\varphi(x) \in \chi$. Following our consideration

$$\frac{d}{d\epsilon} \int_\Omega F(x, u, Du)\, dx \bigg|_{\epsilon=0} = 0$$

Assuming interchangeability of derivative and integral and $F = F(x, z, p)$, we have

$$\int_\Omega \langle F_z(x, \bar{u}, D\bar{u}), \varphi\rangle + \langle F_p(x, \bar{u}, D\bar{u}), D\varphi\rangle\; dx \qquad \forall\varphi \in C_c^\infty(\Omega; \mathbb{R}^\ell). \tag{2.1}$$

Notes. The inner product $\langle F_p(x, \bar{u}, D\bar{u}), D\varphi\rangle$ is the Frobenius inner product, it is the same as when we consider tensors of $\mathbb{R}^\ell \otimes \mathbb{R}^d$ as vectors in $\mathbb{R}^{d \cdot l}$. In this note, I prefer this notion of inner product to not be mislead in further technical proofs.

However, there will be instance where I do not emphasize this notion for inner product. The reader is advised to stay vigilant and try to verify calculations and notions given in this note.

Before continuing, it is nice to review the result on integration by parts as follows.

Theorem 2.1.1 (Integration by Parts Formula)
Let $\Omega \subseteq \mathbb{R}^d$ be a bounded open set of class C^1, and C^1 functions $F : \Omega \to \mathbb{R}^d$ and $\varphi : \Omega \to \mathbb{R}$. Then

$$\int_\Omega \langle F, \nabla\varphi\rangle\; dx = -\int_\Omega \operatorname{div}(F)\varphi\, dx + \int_{\partial\Omega} \langle F, \nu\rangle\, \varphi\, d\sigma(x)$$

where ν is the normal outward pointing vector defined on $\partial\Omega$.

Proof. Apply Divergence Theorem on φF. ■

Assuming $\bar{u} \in C^2$, we can work with integration by parts (actually, we are working on a more general integration by parts) to have

$$\begin{aligned}
\int_\Omega \langle F_p(x, \bar{u}, D\bar{u}), D\varphi\rangle &= \sum_{i,j} \int_\Omega F_{p_i^j}(x, \bar{u}, D\bar{u})\partial_i\varphi^j \\
&= \sum_{i,j} \int_\Omega \partial_i \left(F_{p_i^j}(x, \bar{u}, D\bar{u})\varphi^j\right) - \partial_i \left(F_{p_i^j}(x, \bar{u}, D\bar{u})\right)\varphi^j \\
&= \int_\Omega \operatorname{div}\left(F_p(x, \bar{u}, D\bar{u})\varphi\right) - \langle \operatorname{div} F_p(x, \bar{u}, D\bar{u}), \varphi\rangle
\end{aligned}$$

Notes. The notion of divergence for $\operatorname{div}(F_p(x, \bar{u}, D\bar{u})\varphi)$ follows divergence on vector field which produces scalar value. However, the notion of divergence for $\operatorname{div} F_p(x, \bar{u}, D\bar{u})$ follows divergence of rank-2 tensor (in this case in $l \times d$ matrix) which is $\mathbb{R}^\ell$ valued where the α-th entry is

$$(\operatorname{div} F_p(x,\bar u, D\bar u))^\alpha = \sum_{i=1}^d \partial_i F_{p_i^\alpha}(x,\bar u, D\bar u)$$

We can apply divergence theorem on the term $\operatorname{div}(F_p(x,\bar u, D\bar u)\varphi)$ to have

$$\begin{aligned}\int_\Omega \langle F_p(x,\bar u, D\bar u), D\varphi\rangle &= \underbrace{\cancel{\int_{\partial\Omega} (F_p\varphi)\cdot \nu\, dS}}_{=0\ (\text{due to } \varphi|_{\partial\Omega}=0)} - \int_\Omega \langle \operatorname{div} F_p(x,\bar u, D\bar u), \varphi\rangle \\ &= \int_\Omega \langle -\operatorname{div} F_p(x,\bar u, D\bar u), \varphi\rangle\end{aligned}$$

Plugging this result back to (2.1), we have

$$\int_\Omega \langle F_z(x,\bar u, D\bar u) - \operatorname{div} F_p(x,\bar u, D\bar u), \varphi\rangle = 0 \quad \forall \varphi \in C_c^\infty(\Omega;\mathbb{R}^\ell) \tag{2.2}$$

We recall an important result to proceed as follows.

Lemma 2.1.2 (Du Bois-Reymond, Fundamental Lemma of Calculus of Variations)
Let $f \in L^1_{\text{loc}}(\Omega)$ such that $\int f\varphi\, dx = 0$ for any $\varphi \in C_c^\infty(\Omega)$. Then $f = 0$ a.e.

Proof. Let $A \subset \Omega$ be any open set. Consider an increasing sequence of mollifiers $\varphi_k \uparrow 1_A$. Hence, by Lebesgue Dominated Convergence Theorem, we have

$$\int_A f = \lim_{k\to\infty} \int_\Omega f\varphi_k = 0.$$

By regularity of Lebesgue measure, $\int_E f = 0$ for any measurable set E. Hence, $f = 0$ a.e. ■

Remark 2.1.3. We can improve this result to vectorial case of $\int \langle F, \varphi\rangle = 0$ by proving each entry of F is 0 a.e. considering appropriate φ.

Utilizing Du Bois-Reymond, equation (2.2) implies

$$\operatorname{div} F_p(x,\bar u, D\bar u) = F_z(x,\bar u, D\bar u) \tag{EL}$$

which we refer to as the Euler-Lagrange equation.

Still within the assumption of $\bar u \in C^2$, for all $\alpha = 1, \ldots, l$ we have

$$\begin{aligned}\operatorname{div} F_p(x,\bar u, D\bar u) = F_z(x,\bar u, D\bar u) &\iff \sum_{i=1}^d \partial_i (F_{p_i^\alpha})(x,\bar u, D\bar u) = F_{z^\alpha}(x,\bar u, D\bar u) \\ \iff \sum_i \frac{\partial}{\partial x^i} F_{p_i^\alpha} + \sum_{i,\beta} F_{p_i^\alpha z^\beta}(x,\bar u, D\bar u)\partial_i u^\beta &+ \sum_{i,j,\beta} F_{p_i^\alpha p_j^\beta}(x,\bar u, D\bar u)\partial_{ij} u^\beta = F_{z^\alpha}(x,\bar u, D\bar u)\end{aligned}$$

Thus, we get a system of second order PDEs

$$F_{pp} : D^2 u + F_{pz}\cdot Du + \operatorname{div}_x(F_p) = F_z. \tag{Strong EL}$$

As before, $\operatorname{div}(F_p)$ is defined for rank-2 tensor case.

Notes. The term $F_{pp} : D^2u$ denotes a double contraction between the rank-4 tensor F_{pp} and the rank-3 tensor D^2u. It is defined component-wise as:

$$(F_{pp} : D^2u)^\alpha = \sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{\beta=1}^{\ell} \frac{\partial^2 F}{\partial p_i^\alpha \partial p_j^\beta} \frac{\partial^2 u^\beta}{\partial x_i \partial x_j}$$

We summarize these result as follows.

Theorem 2.1.4 (System of Euler-Lagrange Equations)
Let $\Omega \subset \mathbb{R}^d$ be an open domain, and suppose $u \in C^2(\Omega, \mathbb{R}^\ell)$ is a critical point of the functional

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

where the Lagrangian $F \in C^2(\Omega \times \mathbb{R}^\ell \times \mathbb{R}^{l\times d})$ is sufficiently smooth. Then u satisfies the Euler-Lagrange system of equations:

$$\operatorname{div} F_p(x, u, Du) = F_z(x, u, Du). \tag{EL}$$

By applying the chain rule to expand the total spatial divergence, this is equivalent to the following system of second-order partial differential equations (the strong form):

$$F_{pp} : D^2u + F_{pz} \cdot Du + \operatorname{div}_x(F_p) = F_z, \tag{Strong EL}$$

where $\operatorname{div}_x(F_p)$ denotes the divergence of F_p taken strictly with respect to the explicit spatial variable x.

Remark 2.1.5. The PDE (Strong EL) is elliptic when F is convex in the p variable.

§2.2 Examples of Euler-Lagrange Equation

Poisson Problem

Consider $\mathcal{F}(u) = \int_\Omega \frac{|Du|^2}{2} - fu\, dx$ where $u : \Omega \subseteq \mathbb{R}^d \to \mathbb{R}$. In this case, $F(x, z, p) = |p|^2/2 - fz$. The Euler-Lagrange equation for this variational problem is

$$\operatorname{div}(F_p) = F_z \iff \operatorname{div}(Du) = -f \iff -\Delta u = f.$$

In the case of $f \equiv 0$. A harmonic function (solution to $\Delta u = 0$) can be found by finding minimizer of $\int_\Omega |Du|^2$ (Dirichlet principle).

Minimal Surface Problem

Consider $\mathcal{F}(u) = \int_\Omega \sqrt{1 + |Du|^2}$ where $u : \Omega \subseteq \mathbb{R}^d \to \mathbb{R}$. In this case, $F(x, z, p) = \sqrt{1 + |p|^2}$. The Euler-Lagrange equation for this variational problem is

$$\operatorname{div}(F_p) = F_z \iff \operatorname{div}\left(\frac{Du}{\sqrt{1 + |Du|^2}}\right) = 0.$$

For a hypersurface represented as a graph, the vector field $\nu = \dfrac{(-Du, 1)}{\sqrt{1+|Du|^2}}$ is the unit normal vector to the surface. In differential geometry, the divergence of the normal vector field gives the sum of the principal curvatures $(\lambda_1 + \cdots + \lambda_n)$. This sum is called the Mean Curvature (H). Therefore, the equation states that $H = 0$ everywhere on the surface.

Geodesic Problem

Consider $\mathcal{L}(u) = \int_0^1 \sqrt{g(\gamma(t), \gamma(t))}\, dt = \int_0^1 \sqrt{g_{\alpha\beta}(\gamma(t))\dot{\gamma}^\alpha(t)\dot{\gamma}^\beta(t)}\, dt$ (we use Einstein summation convention here) where $\gamma : [0,1] \subseteq \mathbb{R} \to (M, g)$ Riemannian manifold. In this case, $L(t, z, p) = \sqrt{g_{\alpha\beta}(z)p^\alpha p^\beta}$ and the Euler Lagrange equation for this variational problem is

$$\begin{aligned}\frac{d}{dt}(L_p) = F_z \iff & \frac{d}{dt}\left(\frac{1}{2L}\cdot\left(g_{\delta\beta}(\gamma)\dot{\gamma}^\beta + g_{\alpha\delta}(\gamma)\dot{\gamma}^\alpha\right)\right) = \frac{1}{2L}\cdot\partial_\delta g_{\alpha\beta}(\gamma)\dot{\gamma}^\alpha\dot{\gamma}^\beta \qquad \forall\delta \\ \iff & \frac{d}{dt}\left(\frac{1}{L}g_{\delta\alpha}(\gamma)\dot{\gamma}^\alpha\right) = \frac{1}{2L}\partial_\delta g_{\alpha\beta}(\gamma)\dot{\gamma}^\alpha\dot{\gamma}^\beta \quad (\because g \text{ symmetric})\end{aligned}$$

If the minimizer is parametrized by arclength, its speed is constant. Therefore, the denominator L is constant. Our equation is simplified into

$$\frac{d}{dt}(g_{\delta\alpha}(\gamma)\dot{\gamma}^\alpha) = \frac{1}{2}\partial_\delta g_{\alpha\beta}(\gamma)\dot{\gamma}^\alpha\dot{\gamma}^\beta \qquad \forall\delta.$$

We expand the left hand side of the equation to get

$$g_{\delta\alpha}(\gamma)\ddot{\gamma}^\alpha + \partial_\beta g_{\delta\alpha}(\gamma)\dot{\gamma}^\alpha\dot{\gamma}^\beta = \frac{1}{2}\partial_\delta g_{\alpha\beta}(\gamma)\dot{\gamma}^\alpha\dot{\gamma}^\beta.$$

Therefore, we isolate the velocity term:

$$\begin{aligned}g_{\delta\alpha}(\gamma)\ddot{\gamma}^\alpha &= \frac{1}{2}\partial_\delta g_{\alpha\beta}(\gamma)\dot{\gamma}^\alpha\dot{\gamma}^\beta - \partial_\beta g_{\delta\alpha}(\gamma)\dot{\gamma}^\alpha\dot{\gamma}^\beta \\ &= \frac{1}{2}\partial_\delta g_{\alpha\beta}(\gamma)\dot{\gamma}^\alpha\dot{\gamma}^\beta - \frac{1}{2}\left(\partial_\beta g_{\delta\alpha}(\gamma) + \partial_\alpha g_{\delta\beta}(\gamma)\right)\dot{\gamma}^\alpha\dot{\gamma}^\beta. \\ &= \frac{1}{2}(\partial_\delta g_{\alpha\beta}(\gamma) - \partial_\beta g_{\delta\alpha}(\gamma) - \partial_\alpha g_{\delta\beta}(\gamma))\dot{\gamma}^\alpha\dot{\gamma}^\beta.\end{aligned}$$

Indeed, we multiply by the inverse metric tensor $g^{\sigma\delta}$ to solve for $\ddot{\gamma}^\sigma$:

$$\ddot{\gamma}^\sigma = -\Gamma^\sigma_{\alpha\beta}\dot{\gamma}^\alpha\dot{\gamma}^\beta$$

where $\Gamma_{\alpha\beta}$ is the Christoffel symbol

$$\Gamma_{\alpha\beta} = \frac{1}{2}g^{\sigma\delta}\left(\partial_\delta g_{\alpha\beta} - \partial_\beta g_{\delta\alpha} - \partial_\alpha g_{\delta\beta}\right).$$

Hence, we obtained the classical geodesic equation

$$\ddot{\gamma}^\sigma + \Gamma^\sigma_{\alpha\beta}\dot{\gamma}^\alpha\dot{\gamma}^\beta = 0.$$

The Brachistochrone Problem

Consider $\mathcal{F}(u) = \int_0^1 \frac{\sqrt{1+(u')^2}}{\sqrt{-u}}\,dx$ where $u : [0,1] \subseteq \mathbb{R} \to \mathbb{R}$. In this case, $F(x,z,p) = \frac{\sqrt{1+p^2}}{\sqrt{-z}}$. The Euler-Lagrange equation for this variational problem is

$$(F_p)' = F_z \iff \left(\frac{1}{\sqrt{-u}} \cdot \frac{u'}{\sqrt{1+(u')^2}}\right) = \frac{1}{2} \cdot \frac{\sqrt{1+(u')^2}}{\sqrt{-u^3}}$$

This equation tells us that a linear function is not a minimizer. Actually, neither is an arc of circle a minimizer for the problem.

Indeed, most Euler-Lagrange equations are tedious, but in the case of $d = 1$ and $F(x,z,p) = F(z,p)$ (independent of x), we have an alternative approach to the Euler-Lagrange equation as follows.

Lemma 2.2.1 (Second Form of Euler-Lagrange, Du Bois-Reymond Equation)
Assume $u : I \subseteq \mathbb{R} \to \mathbb{R}^\ell$ and $F(x,z,p) = F(z,p)$ (independent of x) solves the Euler-Lagrange equation

$$(F_p(u,u'))' = F_z(u,u').$$

Then,

$$u' \cdot F_p(u,u') - F(u,u') = \text{constant}.$$

Where $\cdot$ is the vector dot product.

Rroof 1 (differentiating). We differentiate the target equation:

$$\frac{d}{dx}(u' \cdot F_p - F) = \left(u'' \cdot F_p + u' \cdot \frac{d}{dx}(F_p) - \frac{d}{dx}F(u,u')\right).$$

Now, because F explicitly depends only on u and u' (no x), we expand the total derivative of F using the multi-variable chain rule:

$$\frac{d}{dx}F(u,u') = F_z \cdot u' + F_p \cdot u''$$

Substitute this back into our equation:

$$\frac{d}{dx}(u' \cdot F_p - F) = u'' \cdot F_p + u' \cdot \frac{d}{dx}(F_p) - (F_z \cdot u' + F_p \cdot u'')$$

Fortunately, the $u'' \cdot F_p$ terms perfectly cancel out, leaving:

$$u' \cdot \frac{d}{dx}(F_p) - F_z \cdot u' = u' \cdot \left(\frac{d}{dx}(F_p) - F_z\right)$$

Because we assumed u solves the Euler-Lagrange equation, we know that $\frac{d}{dx}(F_p) - F_z = 0$. Therefore:

$$\frac{d}{dx}(u' \cdot F_p - F) = u' \cdot 0 = 0$$

Since the total derivative is exactly zero, the quantity must be a constant. ■

Proof 2 (Theorem 4.20 of [Dac07]). Let u be the solution of the variational problem with respect to the Lagrangian $F(u,u')$ and let $I=[a,b]$

We shall consider a technique known as *variations of the independent variables*; the classical derivation of Euler-Lagrange equation can be seen as a technique of variations of the dependent variables.

Let $\epsilon\in\mathbb{R}$, $\varphi\in C_c^\infty(a,b)$, $\lambda=(2\left\|\varphi'\right\|_\infty)^{-1}$, and consider

$$\xi(x,\epsilon)=x+\epsilon\lambda\varphi(x)=y.$$

Observe that for $|\epsilon|\le 1$, then $\xi(\cdot,\epsilon):I\to I$ is a diffeomorphism with $\xi(a,\epsilon)=a,\xi(b,\epsilon)=b$, and $\xi_x(x,e)>0$. Let $\eta(\cdot,\epsilon):I\to I$ be its inverse, i.e.

$$\xi(\eta(y,\epsilon),\epsilon)=y.$$

Differentiating with respect to y and ϵ respectively yields

$$\xi_x(\nu(y,\epsilon),\epsilon)\eta_y(y,\epsilon)=1 \text{ and } \xi_x(\eta(y,\epsilon),\epsilon)\eta_\epsilon(y,\epsilon)+\xi_\epsilon(\eta(y,\epsilon),\epsilon)=0.$$

With these equalities and noting the fact $|\epsilon\lambda\varphi'(y)|\le\dfrac{\epsilon}{2}<1$, we have

$$\begin{aligned}\eta_y(y,\epsilon)&=\frac{1}{\xi_x(\eta(y,\epsilon))}=\frac{1}{1+\epsilon\lambda\varphi'}=1-\epsilon\lambda\varphi'(y)+O(\epsilon^2)\\ \eta_\epsilon(y,\epsilon)&=-\frac{\xi_\epsilon(\eta(y,\epsilon),\epsilon)}{\xi_x(\eta(y,\epsilon),\epsilon)}=-\frac{\lambda\varphi}{1+\epsilon\lambda\varphi'}=-\lambda\varphi(y)+O(\epsilon)\end{aligned}$$

Consider $v_\epsilon(x)=u(\xi(x,\epsilon))$. Indeed, $\epsilon=0$ is a minimizer of $\epsilon\mapsto\int_I F(v_\epsilon,v_\epsilon')$ since $v_0=u$. Note that

$$\begin{aligned}\int_I F(v_\epsilon,v_\epsilon')\,dx&=\int_I F(u(\xi(x,\epsilon)),u'(\xi(x,\epsilon))\xi_x(x,\epsilon))\,dx\\ &=\int_I F\left(u(y),\frac{u'(y)}{\eta_y(y,\epsilon)}\right)\eta_y(y,\epsilon)\,dy.\end{aligned}$$

Hence,

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0}\int_I F(v_\epsilon,v_\epsilon')\,dx=\int_I\left.\frac{d}{d\epsilon}\right|_{\epsilon=0}F\left(u(y),\frac{u'(y)}{\eta_y(y,\epsilon)}\right)\eta_y(y,\epsilon)\,dy.=0$$

where

$$\begin{aligned}\left.\frac{d}{d\epsilon}\right|_{\epsilon=0}F\left(u(y),\frac{u'(y)}{\eta_y(y,\epsilon)}\right)\eta_y(y,\epsilon)&=\eta_{y\epsilon}(y,\epsilon)F\left(u(y),\frac{u'(y)}{\eta_y(y,\epsilon)}\right)-\frac{\eta_{y\epsilon}}{\eta_y}u'(y)\cdot F_p\left(u(y),\frac{u'(y)}{\eta_y(y,\epsilon)}\right)\Bigg|_{\epsilon=0}\\ &=\lambda(u'(y)\cdot F_p(u(y),u'(y))-F(u(y),u'(y)))\varphi'(y).\end{aligned}$$

Therefore, for any $\varphi\in C_c^\infty(I)$ we have

$$\int_I(u'\cdot F_p-F)\varphi'\,dx=0.$$

With integration by parts, the integral becomes

$$-\int_I\frac{d}{dx}\left(u'\cdot F_p-F\right)\varphi=0.$$

Therefore, $(u\cdot F_p(u,u')-F(u,u'))'=0$ in weak sense and thus $u\cdot F_p(u,u')-F(u,u')=$ constant. ■

Remark 2.2.2. Proof 2 is stronger than Proof 1 since Proof 2 works under weaker regularity hypotheses on the minimizer u.

A direct consequence of the second formulation is the solution for the Brachistochrone problem. Since $F(u,u') = \dfrac{\sqrt{1+(u')^2}}{\sqrt{-u}}$We have

$$\begin{aligned} u'F_p - F = C &\iff \frac{1}{\sqrt{-u}}\frac{(u')^2}{\sqrt{1+(u')^2}} - \frac{\sqrt{1+(u')^2}}{\sqrt{-u}} = C \\ &\implies \frac{-1}{\sqrt{1+(u')^2}} = C\sqrt{-u} \\ &\implies (-u)(1+(u')^2) = C_1. \end{aligned}$$

To solve this ODE, we look for a solution of the form $(x, u(x)) = (x(\tau), y(\tau))$. One can deduce the solution is

$$\begin{cases} x(\tau) = R(\tau - \sin(\tau)) \\ y(\tau) = -R\cos(\tau) \end{cases}$$

§2.3 Exercises

Exercise 1

Find the solution of the Euler-Lagrange equation for

$$\min\left\{\int_{-1}^{1} u\sqrt{1+(u')^2}\,dx : u(-1) = u(1) = \alpha\right\} \tag{2.3}$$

Solution.
The Lagrangian is $F(x,u,u') = u\sqrt{1+(u')^2}$. Since F has no explicit dependence on the spatial variable x, any extremizer must satisfy the Du Bois-Reymond equation:

$$u'F_{u'} - F = C$$

where C is a real constant. Computing the derivative with respect to u':

$$F_{u'} = \frac{uu'}{\sqrt{1+(u')^2}}$$

Substituting this into the initial equation yields:

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

By getting a common denominator, this strictly simplifies to:

$$\frac{u(u')^2 - u(1+(u')^2)}{\sqrt{1+(u')^2}} = C \implies \frac{-u}{\sqrt{1+(u')^2}} = C$$

Squaring both sides and solving for $(u')^2$ gives an autonomous separable ordinary differential equation:

$$\frac{u^2}{1+(u')^2} = C^2 \implies (u')^2 = \frac{u^2}{C^2} - 1 \implies u' = \pm\sqrt{\left(\frac{u}{C}\right)^2 - 1}$$

Separating variables and integrating gives:

$$\int \frac{du}{\sqrt{(u/C)^2 - 1}} = \pm \int C\,dx \implies C \operatorname{arcosh}\left(\frac{u}{C}\right) = \pm x + c_2$$

Thus, the general solution is the hyperbolic cosine (catenoid profile):

$$u(x) = C \cosh\left(\frac{x - x_0}{C}\right)$$

where C and x_0 are integration constants uniquely determined by the boundary conditions $u(-1) = u(1) = \alpha$.

Exercise 2

Write and find the solution to the Euler-Lagrange equation for the catenary problem:

$$\min\left\{\int_{-1}^{1} u\sqrt{1 + (u')^2}\,dx : u(-1) = u(1) = \alpha, \quad \int_{-1}^{1} \sqrt{1 + (u')^2}\,dx = L\right\} \tag{2.4}$$

Hint for 2):
Try to prove that a minimizer is a stationary point for

$$\int_{-1}^{1} u\sqrt{1 + (u')^2}\,dx - \lambda \int_{-1}^{1} \sqrt{1 + (u')^2}\,dx$$

for a suitable Lagrange multiplier λ.
Hint for hint: Consider two-parameter inner variations of the form

$$v_{t,s}(x) = \bar{u}(x) + t\varphi(x) + s\psi(x)$$

Let $f(t, s) = \mathcal{F}(v_{t,s})$ and $g(t, s) = \mathcal{G}(v_{t,s})$. Apply the standard finite-dimensional Lagrange multiplier theorem to find the minimum of $f(t, s)$ subject to the constraint $g(t, s) = L$.

Solution.
Following the hint, by the calculus of variations for isoperimetric constraints, the minimizer $\bar{u}$ must be a critical point of the augmented functional:

$$\mathcal{L}(u) = \int_{-1}^{1} (u - \lambda)\sqrt{1 + (u')^2}\,dx$$

where $\lambda \in \mathbb{R}$ is the Lagrange multiplier associated with the length constraint.

Define an auxiliary function $w(x) = u(x) - \lambda$. Notice that $w'(x) = u'(x)$. Under this substitution, the augmented functional becomes:

$$\tilde{\mathcal{L}}(w) = \int_{-1}^{1} w\sqrt{1 + (w')^2}\,dx$$

This is identical to the functional in Exercise 1. Therefore, $w(x)$ must satisfy the exact same Du Bois-Reymond equation, yielding the solution:

$$w(x) = C \cosh\left(\frac{x - x_0}{C}\right)$$

Substituting back $u(x) = w(x) + \lambda$, the general solution for the physical catenary problem is:

$$u(x) = \lambda + C \cosh\left(\frac{x - x_0}{C}\right)$$

The three unknowns (λ, C, x_0) are uniquely determined by the non-linear system formed by the two boundary conditions $u(-1) = \alpha$, $u(1) = \alpha$, and the arc length integral $L = \int_{-1}^{1} \sqrt{1 + (u')^2}\, dx = C\left[\sinh\left(\frac{1 - x_0}{C}\right) - \sinh\left(\frac{-1 - x_0}{C}\right)\right]$.

3 Some Properties of Minimizers

We start this lecture by putting a remark in the case of non-linear constrains on the image of $\bar{u}$, the variation of the form $u + \epsilon\varphi$ might be hard to deal with.

For example, the problem of minimizing

$$\inf\left\{\int_0^1 |u' - u| : u|_{\partial[0,1]} = g, u \geq 0\right\}$$

is difficult.

§3.1 Constant Integral Constraint

We shall consider the variation problem for $u : \Omega \subseteq \mathbb{R}^d \to \mathbb{R}$ with integral constraints stated as follows

$$\inf\left\{\int_\Omega F(x, u, Du)\,dx : u \text{ such that } \int_\Omega G(x, u, Du)\,dx = \text{constant}, \ u|_{\partial\Omega} = h\right\}. \tag{3.1}$$

Before continuing, given is a review of the implicit function theorem that we will use for the subsequent theorem.

Theorem 3.1.1 (Implicit Function Theorem)

Let $A \subset \mathbb{R}^n \times \mathbb{R}^m$ be an open set, and let $F : A \to \mathbb{R}^m$ be a continuously differentiable function of class C^k $(k \geq 1)$. We denote a point in A by (x, y), where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$.

Suppose that $(a, b) \in A$ is a point satisfying:

1. $F(a, b) = 0$,
2. The partial Jacobian matrix with respect to y, denoted $D_y F(a, b) \in \mathbb{R}^{m\times m}$, is invertible (i.e., $\det(D_y F(a, b)) \neq 0$).

Then, there exists an open neighborhood $U \subset \mathbb{R}^n$ of a, an open neighborhood $V \subset \mathbb{R}^m$ of b, and a unique function $g : U \to V$ such that:

$$F(x, g(x)) = 0 \quad \text{for all } x \in U.$$

Furthermore, the implicit function g is also of class C^k on U, and its Jacobian derivative at any $x \in U$ is given by:

$$Dg(x) = -\left[D_y F(x, g(x))\right]^{-1} D_x F(x, g(x)).$$

Theorem 3.1.2 (Lagrange Multiplier (Note: Regularity is not really concerned here))
If $\bar{u} : \Omega \subseteq \mathbb{R}^d \to \mathbb{R}$ is a minimizer of the variational problem (3.1) and there exists $\psi \in C_c^\infty(\Omega)$ such that

$$\int_\Omega G_z(x, \bar{u}, D\bar{u})\psi + \langle G_p(x, \bar{u}, D\bar{u}), D\psi \rangle \; dx \neq 0.$$

Then, there exists $\lambda \in \mathbb{R}$ such that $\bar{u}$ is a critical point for

$$\int_\Omega F(x, u, Du) + \lambda G(x, u, Du) \, dx$$

i.e., for any $\varphi \in C_c^\infty(\Omega)$

$$\int_\Omega \left[F_z(x, \bar{u}, D\bar{u}) + \lambda G_z(x, \bar{u}, D\bar{u})\right] \varphi + \left[F_p(x, \bar{u}, D\bar{u}) + \lambda G_p(x, \bar{u}, D\bar{u})\right] \cdot D\varphi \, dx = 0$$

Proof. Let φ be any smooth enough function and ψ as assumed above. Consider

$$f(t, s) = \int_\Omega F(x, \bar{u} + t\varphi + s\psi, D\bar{u} + tD\varphi + sD\psi) \, dx,$$
$$g(t, s) = \int_\Omega G(x, \bar{u} + t\varphi + s\psi, D\bar{u} + tD\varphi + sD\psi) \, dx.$$

Since $\bar{u}$ is a minimizer, $f(0,0) \leq f(t,s)$ for any (t,s) such that $g(t,s) = g(0,0) = C$.

Note that our assumption itself is a $\mathbb{R}^2$ problem

$$\begin{cases} g(0,0) = C \\ \partial_s g(0,0) = \displaystyle\int_\Omega G_z(x, \bar{u}, D\bar{u})\psi + \langle G_p(x, \bar{u}, D\bar{u}), D\psi \rangle \; dx \neq 0. \end{cases}$$

Hence, by implicit function theorem, there exists an open interval $I \ni 0$ and a function $s : I \to \mathbb{R}$ with $s(0) = 0$ such that

$$g(t, s(t)) = C \qquad \forall t \in I$$

and $s'(t) = -\dfrac{g_t(t, s(t))}{g_s(t, s(t))}$.

In particular, $\tilde{f}(t) := f(t, s(t))$ has a minimum in $t = 0$ and thus

$$0 = \tilde{f}'(0) = f_t(0,0) + f_s(0,0) \cdot s'(0) = f_t(0,0) - g_t(0,0) \cdot \frac{f_s(0,0)}{g_s(0,0)}.$$

Therefore, we have $\lambda = \lambda(\bar{u}, \psi) = -\dfrac{f_s(0,0)}{g_s(0,0)}$ as our desired constant. ∎

An application of the constant integral constraints following problem 3.1 is the first eigenvalue problem as follows.

Theorem 3.1.3 (The first eigenvalue of the Laplacian)
Let Ω be a bounded set, there exists $C = C(\Omega) > 0$ such that for any $\varphi \in C_c^1(\Omega)$ (or $W_0^{1,2}(\Omega)$), the following inequality holds

$$C \cdot \int_\Omega |\varphi|^2 \leq \int_\Omega |\nabla\varphi|^2.$$

The optimal constant C_{opt} is given by

$$C_{opt} = \lambda(\Omega) = \inf\left\{\int_\Omega |\nabla\varphi|^2 \middle| \varphi \in C_c^1(\Omega), \int_\Omega \varphi^2 = 1\right\}.$$

Proof. Our justification on finding the optimal C is from the sufficiency on proving the inequality in the case $\|\varphi\|_{L^2}^2 = \int_\Omega |\varphi|^2 = 1$ since for general case we can consider $\phi = \frac{\varphi}{\|\varphi\|_{L^2}}$ to assert our desired inequality.

Assume a minimizer $\bar{u}$ exists (the solution exists and it can be found by utilizing direct method (c.f. Lecture 4) with the aid of Rellich-Kondrachov (Theorem 6.1.9) and actually $\bar{u} \in W_0^{1,2}(\Omega))$ with $\int_\Omega \bar{u}^2 = 1$, and thus

$$\int_\Omega |\nabla\bar{u}|^2 = \inf\left\{\int_\Omega |\nabla\varphi|^2 \middle| \varphi \in C_c^1(\Omega), \int_\Omega \varphi^2 = 1\right\}.$$

In this case, the Lagrangian is $F(x,u,Du) = |Du|^2$ and the contraint is $G(x,u,Du) = u^2$ where $F_z = 0$, $F_p = 2p$, $G_z = 2z$, $G_p = 0$. We then have for $\psi \in C_c^\infty(\Omega)$

$$\int_\Omega G_z(x,\bar{u},D\bar{u})\psi + \langle G_p(x,\bar{u},D\bar{u}), D\psi\rangle \, dx = 2\int_\Omega \bar{u}\psi \, dx.$$

Since $\bar{u}$ cannot be a zero function, having the fact space of smooth compactly supported functions $C_c^\infty(\Omega)$ is dense in $L^2(\Omega)$, there must exist at least one smooth test function ψ that is not orthogonal to $\bar{u}$. Thus, we have our desired $\psi \in C_c^\infty(\Omega)$. By Theorem 3.1.2, there exists λ such that

$$\int_\Omega \langle D\bar{u}, D\varphi\rangle + \lambda\bar{u}\varphi = 0 \qquad \forall\varphi \in C_c^1(\Omega).$$

Fortunately, since $\bar{u} \in W_0^{1,2}(\Omega)$, there exists a sequence $\{\varphi_k\} \subset C_c^1(\Omega)$ such that $\varphi \xrightarrow{W^{1,2}} \bar{u}$. Thus, we have

$$\int_\Omega |D\bar{u}|^2 + \lambda(\bar{u})^2 \, dx = \int_\Omega |D\bar{u}|^2 + \lambda = 0.$$

Hence, we can choose $C_{opt} = \lambda(\Omega) = -\lambda$ and we have $\bar{u}$ as well a solution to the Poisson equation $-\Delta u = \lambda(\Omega)u$. Moreover, λ is an eigenvalue for the Laplacian operator (actually, λ is the smallest one). ■

Remark 3.1.4. By leaving the boundary values of u unconstrained in the minimization problem, the variational machinery naturally forces the normal derivative of the minimizer to be zero on the boundary. Hence, the Neumann boundary condition will be given $\frac{\partial\bar{u}}{\partial\nu} = \langle D\bar{u}, \nu\rangle = 0$ on $\partial\Omega$.

Another application of the Lagrange multiplier is one of the Poincaré inequalities as follows.

Theorem 3.1.5 (Poincaré-Wirtinger Inequality)
Let $\Omega \subset \mathbb{R}^d$ be a bounded, connected domain with Lipschitz boundary. Then there exists $C = C(\Omega) > 0$ such that for any $u \in W^{1,2}(\Omega)$ we have

$$\int_\Omega |u(x) - (u)_\Omega|^2 \, dx \le C \int_\Omega |\nabla u(x)|^2 \, dx,$$

where $(u)_\Omega = ⨍_\Omega u(y)\, dy$. The optimal constant is $C_{opt} = \dfrac{1}{\mu(\Omega)}$, where $\mu(\Omega)$ is the first non-trivial Neumann eigenvalue defined by:

$$\mu(\Omega) = \inf \left\{ \int_\Omega |\nabla u|^2 \, dx : u \in W^{1,2}(\Omega),\ \int_\Omega u^2 \, dx = 1,\ \int_\Omega u \, dx = 0 \right\}.$$

Proof. The proof of the inequality is omitted (the classical proof is by contradiction using the Rellich-Kondrachov compactness theorem (Theorem 6.1.9)).

Regarding the optimal constant, the direct method of the calculus of variations (c.f. Lecture 4) ensures the existence of a minimizer $\bar{u} \in W^{1,2}(\Omega)$. We define the Lagrangian $F(x, u, Du) = |Du|^2$ subject to the two integral constraints given by $G(x, u, Du) = u^2$ and $H(x, u, Du) = u$:

$$\int_\Omega G(x, u, Du)\, dx = 1 \quad \text{and} \quad \int_\Omega H(x, u, Du)\, dx = 0.$$

Because the minimization is taken over the unconstrained space $W^{1,2}(\Omega)$, the Lagrange multiplier theorem yields multipliers $\mu, \sigma \in \mathbb{R}$ such that the first variation vanishes for all test functions $\varphi \in W^{1,2}(\Omega)$ (and thus for all $\varphi \in C^1(\bar{\Omega})$):

$$\int_\Omega \big(F_p + \mu G_p + \sigma H_p\big) \cdot D\varphi + \big(F_z + \mu G_z + \sigma H_z\big)\varphi \, dx = 0 \quad \forall \varphi \in C^1(\bar{\Omega}).$$

Substituting the derivatives (and absorbing the factor of 2 into the multipliers), we obtain the weak equation:

$$\int_\Omega \langle D\bar{u}, D\varphi \rangle + \mu \bar{u}\varphi + \sigma\varphi \, dx = 0 \qquad \forall \varphi \in C^1(\bar{\Omega}).$$

Because $\varphi \equiv 1 \in C^1(\bar{\Omega})$ is a valid test function, we substitute it to find:

$$\int_\Omega \mu\bar{u} + \sigma \, dx = 0 \implies \mu \int_\Omega \bar{u}\, dx + \sigma|\Omega| = 0.$$

Since $\bar{u}$ satisfies the zero-mean constraint $\int_\Omega \bar{u} = 0$, we immediately conclude $\sigma = 0$. The weak equation thus simplifies to:

$$\int_\Omega \langle D\bar{u}, D\varphi \rangle + \mu\bar{u}\varphi \, dx = 0 \qquad \forall \varphi \in C^1(\bar{\Omega}). \tag{3.2}$$

Because $\bar{u} \in W^{1,2}(\Omega)$ itself is an admissible test function, we may set $\varphi = \bar{u}$ directly to obtain:

$$\int_\Omega |D\bar{u}|^2 \, dx + \mu \int_\Omega \bar{u}^2 \, dx = 0 \implies \mu(\Omega) + \mu = 0.$$

Therefore, $\mu = -\mu(\Omega)$.

To determine the strong form of the PDE and the boundary condition, we apply integration by parts to (3.2):

$$\int_\Omega (-\Delta \bar{u} + \mu \bar{u})\varphi \, dx + \int_{\partial\Omega} \varphi \frac{\partial \bar{u}}{\partial \nu} \, d\mathcal{H}^{d-1}(x) = 0 \qquad \forall \varphi \in C^1(\bar{\Omega}). \tag{3.3}$$

First, we restrict our test functions to $\varphi \in C_c^\infty(\Omega)$. The boundary integral vanishes, which implies that $\bar{u}$ satisfies the PDE $-\Delta \bar{u} = -\mu \bar{u} = \mu(\Omega)\bar{u}$ almost everywhere in Ω.

Second, knowing that the interior integral in (3.3) is identically zero, we return to general test functions $\varphi \in C^1(\bar{\Omega})$. The equation reduces entirely to:

$$\int_{\partial\Omega} \varphi \frac{\partial \bar{u}}{\partial \nu} \, d\mathcal{H}^{d-1}(x) = 0 \quad \forall \varphi \in C^1(\bar{\Omega}).$$

Because φ is arbitrary on the boundary $\partial\Omega$, we obtained the natural (Neumann) boundary condition $\dfrac{\partial \bar{u}}{\partial \nu} = 0$.

Thus, $\bar{u}$ is a solution to the Neumann eigenvalue problem:

$$\begin{cases} -\Delta u = \mu(\Omega) u & \text{in } \Omega \\ \dfrac{\partial u}{\partial \nu} = 0 & \text{on } \partial\Omega \end{cases}$$

■

§3.2 Exercises on Constant Integral Constraint (not yet attempted)

1) Consider the following problem when defined on Ω with a boundary $\partial\Omega$:

$$\inf \left\{ \int_\Omega \sqrt{1 + |\nabla u|^2} - \sigma \int_{\partial\Omega} u \, dS : \int_\Omega u \, dx = V \right\}$$

This models a water boundary (meniscus). Show that if u is a sufficiently regular minimizer, then it solves the constant mean curvature equation:

$$-\operatorname{div}\left(\frac{\nabla u}{\sqrt{1 + |\nabla u|^2}} \right) = \text{constant} \quad \text{in } \Omega$$

2) Let $\Omega_1, \Omega_2 \subseteq \mathbb{R}^2$ be smooth and disjoint domains. Define the admissible space of curves connecting their boundaries as:

$$\chi = \{\gamma : [0,1] \to \mathbb{R}^2 \setminus (\Omega_1 \cup \Omega_2) : \gamma(0) \in \partial\Omega_1, \gamma(1) \in \partial\Omega_2\}$$

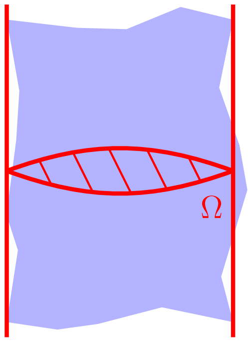
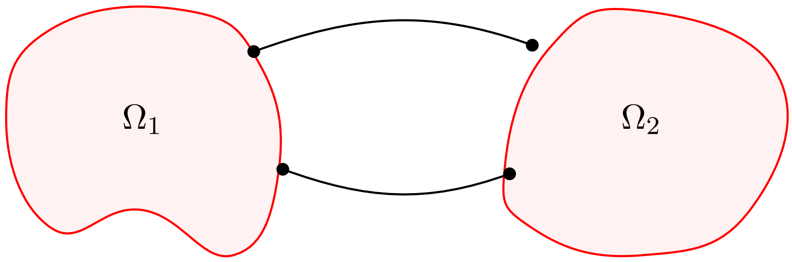


Figure 3.1: The Capillarity Problem (source: my drawing + Gemini)

If $\bar\gamma$ minimizes the length, then $\bar\gamma$ satisfies the transversality conditions:

$$\begin{cases} \ddot{\bar\gamma} = 0 \\ \dot{\bar\gamma}(0) \perp \partial\Omega_1 \\ \dot{\bar\gamma}(1) \perp \partial\Omega_2 \end{cases}$$



3) Suppose $\bar\gamma$ minimizes the weighted length:

$$\int_0^1 \eta(\gamma(t))|\dot\gamma(t)|\,dt \quad \text{subject to } \gamma(0) = P, \gamma(1) = Q$$

where the weight η is discontinuous across the axis $y = 0$:

$$\eta(x, y) = \begin{cases} \eta_-, & y \leq 0 \\ \eta_+, & y \geq 0 \end{cases}$$

Show that a minimizer should solve $\ddot{\bar\gamma} = 0$ away from the interface, and satisfy the transmission condition (Snell's Law) at $y = 0$.

§3.3 Examples for Cases Where Minimizers Suck

1) Lack of Minimizer: Consider the problem:

$$\inf \left\{ \int_0^1 \left(u^2 + ((u')^2 - 1)^2\right) dx : u(0) = u(1) = 0 \right\}$$

The infimizing sequence drives toward $\bar u = 0$, but if we plug that in, the energy $((u')^2 - 1)^2 \neq 0$, meaning the minimum is never attained in the standard space.

2) Minimizers exist but they are not C^1:

$$\min\left\{\int_{-1}^{1}((u')^2-1)^2\,dx : u(-1)=u(1)=0\right\}$$

The solution forms a zigzag pattern (sawtooth) with sharp corners where the derivative jumps between $+1$ and -1.

3) Minimizers might be C^1, but not C^2:

$$\min\left\{\int_{-1}^{1}u^2(u'-x)^2\,dx : u(-1)=0, u(1)=1\right\}$$

The minimizer is given by the piecewise function:

$$\bar{u}(x)=\begin{cases}0, & x\le 0\\ x^2/2, & x>0\end{cases}$$

4) Minimizers relax the constraint:

$$\min\left\{\int_{\Omega}\sqrt{1+|\nabla u|^2}\,dx : u=0 \text{ on } \partial\Omega,\ u(0)=1\right\}$$

This is known geometrically as the "Tende da circo" (circus tent) problem, where forcing a point constraint on a minimal surface causes the solution to break standard smoothness at the pinned point.

5) The Ball-Mizel Pathology (1985): There exists a 1-dimensional minimization problem for which the absolute minimizer is not a solution to the Euler-Lagrange equation.

§3.4 Modern vs. Classical Calculus of Variations Problems

The modern probabilistic and functional analytic approach to the Calculus of Variations centers around three core questions:

1. **Existence:** When do minimizers actually exist?
2. **Equations:** When do they solve Euler-Lagrange equations, and in which weak or strong sense?
3. **Regularity:** Are these minimizers "better" (more smooth, continuous, integrable) than a generic function in the admissible space?

4 Direct Methods in the Calculus of Variations

Within this lecture, we will discuss the abstract version of direct methods that is essential to the modern approach of Calculus of Variations

§4.1 Abstract Direct Methods

We shall start on building our main ingredients: lower semicontinuity and coercivity.

Definition 4.1.1 (Lower Semicontinuity)
Let X be a topological space. A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is said to be:

1. **lower semicontinuous (l.s.c.)** if for any $x \in X$ it holds that
$$f(x) \le \liminf_{y \to x} f(y)$$
2. **sequentially lower semicontinuous** if for any $x \in X$ and any sequence $\{x_n\} \subseteq X$ that converges to x it holds that
$$f(x) \le \liminf_{n \to \infty} f(x_n)$$

Remark 4.1.2. Note that the notion of lower semicontinuity is closed under suprema since for any family of lower semicontinuous functions $\{f_\alpha\}$, we have
$$f_\alpha(x) \le \liminf_{y \to x} f_\alpha(y) \le \liminf_{y \to x} \sup_\alpha f_\alpha(y)$$
and we can take the supremum over α to conclude $\sup_\alpha f_\alpha(x) \le \liminf_{y \to x} \sup_\alpha f_\alpha(y)$.

Definition 4.1.3 (Coercivity)
Let X be a topological space. A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is said to be:

1. **coercive** if for any $t \in \mathbb{R}$, the set $\{f \le t\}$ is precompact, i.e. $\overline{\{f \le t\}}$ is compact.
2. **sequentially coercive** if for any $t \in \mathbb{R}$, the set $\{f \le t\}$ is sequentially precompact, i.e. any sequence $\{x_n\} \subseteq \{f \le t\}$ admits a converging subsequence.

Remark 4.1.4.

- It is immediate by definition that l.s.c. implies sequential l.s.c. The converse is true if X is first countable, i.e. every it admits a countable base of neighborhoods for any $x \in X$.
- If X is metrizable, coercivity is equivalent with sequential coercivity.

Definition 4.1.5 (Epigraph)
Let X be a set. The **epigraph** of a function $f : X \to \mathbb{R} \cup \{+\infty\}$ is the set

$$\operatorname{Epi} f := \{(x, t) \in X \times \mathbb{R} \mid f(x) \le t\}$$

which is the set of points in $X \times \mathbb{R}$ that lies above the graph of f.

Proposition 4.1.6
Let X be a (first countable) topological space. A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is l.s.c if and only if $\operatorname{Epi} f$ is closed.

Proof. Let f be l.s.c. and $(x, y) \in \overline{\operatorname{Epi} f}$, there exists a sequence $\{(x_n, y_n)\} \subseteq \operatorname{Epi} f$ such that $x_n \to x$ and $y_n \to y$ whereas $f(x_n) \le y_n$ for all $n \in \mathbb{N}$. By lower semicontinuity, by passing $n \to \infty$ yields

$$f(x) \le \liminf_{n\to\infty} f(x_n) \le \lim y_n = y.$$

Hence, $(x, y) \in \operatorname{Epi} f$.

Conversely, let $\operatorname{Epi} f$ be closed and let $x \in X$ and $y < f(x)$ so that $(x, y) \notin \operatorname{Epi} f$. Hence, there exists a neighborhood of (x, y), say $U \times (y - \epsilon, y + \epsilon)$ such that all points in such neighborhood is not contained in $\operatorname{Epi} f$. Moreover, the structure of epigraph gives us $(U \times (-\infty, y + \epsilon)) \cap \operatorname{Epi} f = \varnothing$. Therefore, we have $f(u) \ge y + \epsilon$ for any $u \in U$. Let $\{x_n\} \subseteq X$ converges to x, For n sufficiently large, $x_n \in U$ and therefore,

$$y + \epsilon \le \liminf_{n\to\infty} f(x_n).$$

We can take $\epsilon \to 0^+$ and then $y \to f(x)^-$ to conclude lower semicontinuity of f. ■

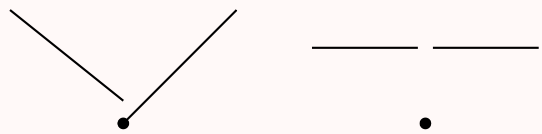
Corollary 4.1.7 (It was put as an exercise)
Let X be a (first countable) topological space and $f : X \to \mathbb{R} \cup \{+\infty\}$. The following statements are equivalent:

1. f is l.s.c.
2. $\{f \le t\}$ is closed for any $t \in \mathbb{R}$
3. $\{f > t\}$ is open for any $t \in \mathbb{R}$
4. $\operatorname{Epi} f$ is closed in $X \times \mathbb{R}$
5. $(X \times \mathbb{R}) \setminus \operatorname{Epi} f = \{(x, t) \in X \times \mathbb{R} : f(x) > t\}$ is open in $X \times \mathbb{R}$

Proof. Omitted. The most tedious equivalence (1. $\Leftrightarrow$ 4.) has been proved in Proposition 4.1.6. ■

Example 4.1.8

1. If A is an open set, then the indicator function $\mathbf{1}_A$ is lower semi-continuous (l.s.c.).
2. Graphical examples of l.s.c. functions at a point (where $f(x_0) \leq \liminf_{x \to x_0} f(x)$):



Exercise 4.1.9

Show that a function $f : \mathbb{R}^d \to [0, +\infty]$ is coercive if and only if $\lim_{|x| \to \infty} f(x) = \infty$.

Proof. ($\Rightarrow$) Suppose that f is coercive, for any $M \in \mathbb{R}$, the set $\{f \leq M\}$ is precompact. Therefore, there exists $R > 0$ such that $\{f \leq M\} \subseteq \{|x| \leq R\}$. Hence, for any $\{|x| > R\} \subseteq \{f > M\}$ and we just proved $\lim_{|x| \to \infty} f(x) = \infty$.

($\Leftarrow$) Suppose that $\lim_{|x| \to \infty} f(x) = \infty$. Therefore, for any $M \in \mathbb{R}$, there exists $R > 0$ such that

$$\{|x| > R\} \subseteq \{f > M\} \iff \{f \leq M\} \subseteq \{|x| \leq R\}.$$

Hence, $\{f \leq M\}$ is precompact. ■

Remark 4.1.10. There might be an issue with consistency on the notion of coercivity from my lecture. The above result also holds for any **Banach space** E if we define coercive as having $\{f \leq t\}$ **bounded for any** $t \in \mathbb{R}$. Hence, in the case of Banach space, we use this notion of coercivity instead of the one we defined.

With our ingredients set, we can naturally prove the following result which we also refer to as the (abstract) direct method.

Theorem 4.1.11 (Weierstrass Method)

Let X be a topological space, $f : X \to \mathbb{R} \cup \{+\infty\}$ such that

1. f is (sequentially) lowersemicontinuous
2. f is (sequentially) coercive, i.e., $\{f \leq t\}$ is (sequentially) compact for any $t \in \mathbb{R}$.

Then either $f \equiv +\infty$ or there exists $x_0 \in X$ such that $f(x_0) = \inf\{f(x) : x \in X\} < +\infty$.

Proof. Suppose there exists $x_0 \in X$ such that $f(x_0) < \infty$. We set $m = \inf_{x \in X} f(x) < \infty$.

We a minimizing sequence $\{x_n\}$ such that $f(x_n) \to m$. Without loss of generality, we assume $\{x_n\} \subseteq \{f \leq m+1\}$. By coercivity, $x_n \to \bar{x}$ up to a subsequence for some $\bar{x} \in X$. Therefore, by lower semicontinuity we conclude

$$m \leq f(\bar{x}) \leq \liminf_{n \to \infty} f(x_n) = m.$$

As a byproduct, we have also proved that $\inf_{x \in X} f(x) > -\infty$. ■

Remark 4.1.12. Given $f : X \to \mathbb{R} \cup \{+\infty\}$, finding a topology for X for which f is both coercive and l.s.c. might not be possible and these are conflicting requirements. Indeed, l.s.c. is easy when we have lots of open sets and compactness is easy with few open sets.

§4.2 Recap on Weak/Weak* Topologies on Banach Space

In this section, we consider E to be a Banach space and E^* is the (topological) dual of E. We denote the duality pairing

$$\langle x^*, x \rangle := x^*(x) \qquad \forall x^* \in E^*, \forall x \in E.$$

Definition 4.2.1

Let E be a Banach space.

1. We say that a sequence $\{x_n\} \subset E$ weakly converges to $x \in E$, denoted by $x_n \rightharpoonup x$, if
$$\langle x^*, x_n \rangle \to \langle x^*, x \rangle \qquad \forall x^* \in E^*.$$
2. We say that a sequence $\{x_n^*\} \subset E$ weakly* converges to $x^* \in E$, denoted by $x_n^* \rightharpoonup^* x^*$, if
$$\langle x_n^*, x \rangle \to \langle x^*, x \rangle \qquad \forall x \in E.$$

Proposition 4.2.2 (Banach-Steinhaus (Uniform Boundedness Theorem))

Let E be a Banach space.

1. If $x_n \rightharpoonup x$, then $\sup_n \|x_n\|_E < \infty$ and in particular $\|x\|_E \le \liminf_{n\to\infty} \|x_n\|_E$.
2. If $x_n^* \rightharpoonup^* x^*$, then $\sup_n \|x_n^*\|_{E'} < \infty$ and in particular $\|x^*\|_{E'} \le \liminf_{n\to\infty} \|x_n^*\|_{E'}$.

Proof. The proof of uniform boundedness is omitted. However, proving $\|x\|_E \le \liminf_{n\to\infty} \|x_n\|_E$ can be done by recalling the fact that

$$\|x\|_E = \sup_{\substack{x^* \in E^* \\ \|x^*\|_{E'} \le 1}} \langle x^*, x \rangle .$$

Then, in the weak topology, the map $x \mapsto \langle x^*, x \rangle$ is continuous for any $x^* \in E^*$ (and thus l.s.c.). Since the notion of l.s.c. is continuous under suprema and thus the map $x \mapsto \|x\|_E$ is l.s.c. in the weak topology. Therefore,

$$\|x\|_E \le \liminf_{n\to\infty} \|x_n\|_E .$$

Similar proof can be used to assert $\|x^*\|_{E'} \le \liminf_{n\to\infty} \|x_n^*\|_{E'}$. ■

The reader is advised to keep the next theorem in disposal since it will be of many use in subsequent lectures.

Theorem 4.2.3

Let E be a Banach space and let $f : E \to \mathbb{R}$ be a convex and l.s.c. (with respect to the norm/strong topology) function. Then, f is l.s.c. with respect to the weak topology.

Proof. To show that f is weakly l.s.c., it is necessary and sufficient to prove that for any arbitrary real number $t \in \mathbb{R}$, the sublevel set

$$K_t = \{x \in E \mid f(x) \leq t\}$$

is closed with respect to the weak topology.

From our hypotheses, we establish two critical properties of K_t:

1. **K_t is convex:** Let $x, y \in K_t$ and $\lambda \in [0, 1]$. By the convexity of f, we have:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \leq \lambda t + (1 - \lambda)t = t$$

   Thus, $\lambda x + (1 - \lambda)y \in K_t$, proving the set is convex.

2. **K_t is norm-closed:** By Corollary 4.1.7, in the norm (strong) topology, K_t is strongly closed.

A fundamental corollary of the Hahn-Banach Separation Theorem (often linked to Mazur's Lemma) dictates that for **any convex subset** of a locally convex topological vector space (such as a Banach space), **closure in the strong topology is equivalent to closure in the weak topology**. (one can consult most functional analysis text such as [Bre11] regarding this theorem or follow the lecture of functional analysis in the department!)

Because the set K_t is both convex and strongly closed, it is automatically **weakly closed** and thus f is weakly l.s.c. ■

Theorem 4.2.4 (Banach-Alaoglu)

Let E be a Banach space, then the set $\{\|x\|_{E'} \leq R\}$ is weak* compact for any $R \geq 0$.

Proof. Omitted. One may consider most functional analysis textbook such as [Bre11] for the proof. ■

Corollary 4.2.5

Let E be a Banach space and $f : E^* \to \mathbb{R} \cup \{+\infty\}$. Then tbe following statements are equivalent.

1. f is weak* coercive
2. $\lim_{\|x^*\|_{E'} \to \infty} f(x^*) = +\infty$
3. There exists a monotone function $\varphi : [0, +\infty] \to [0, +\infty]$ where $\varphi \uparrow \infty$ such that $f(x) \geq \varphi(\|x^*\|_{E'})$.

Proof. The proof of 3. $\implies$ 2. $\implies$ 1. follows by the aid of Exercise 4.1.9 trivially.

1. $\implies$ 2. Assume otherwise, there exists a sequence $\{x_n^*\} \subseteq$ such that $\|x_n^*\|_{E'} \to \infty$, but $f(x_n^*) \leq M$ for some M. Hence, by weak* coercivity, up to subsequence $x_n^* \rightharpoonup^* x^*$ for some $x^* \in E^*$. However, Theorem 4.2.2 implies boundedness of $\{x_n\}$ up to subsequence which contradicts our assumption.

2. $\implies$ 3. One can define for any $s \geq 0$ $\varphi(s) := \inf\{f(x^*) : \|x^*\|_{E'} \geq s\}$ which is monotonely increasing and converges to $+\infty$ as $s \to +\infty$. ■

From all of our consideration, we can assert the direct methods for reflexive Banach space as follows.

Theorem 4.2.6 (Direct Methods, E reflexive Banach space)
Let E be a reflexive Banach space, i.e. $E \cong E^{**}$ and $f : E \to \mathbb{R} \cup \{+\infty\}$, $f \not\equiv +\infty$ such that:

1. f is convex
2. f is strongly l.s.c.
3. f is coercive, i.e., $\lim_{\|x\|_E \to \infty} f(x) = +\infty$

Then, f admits a minimum.

Proof. We shall apply the abstract direct method (Theorem 4.1.11) to the space $X = E$ endowed with the **weak topology**. To do so, we must verify that f is both weakly lower semicontinuous and weakly coercive.

It is immediate from hypotheses 1. and 2., we have f weakly l.s.c. by Theorem 4.2.3. Hence, we only need to prove weak coercivity.

We must show that for any $t \in \mathbb{R}$, the sublevel set $K_t = \{f \leq t\}$ is weakly compact. By hypothesis 3., As established in Exercise 4.1.9, this is equivalent to say every sublevel set K_t is bounded in the norm topology, meaning there exists some $R > 0$ such that $K_t \subseteq \{\|x\|_E \leq R\}$. Since E is a reflexive Banach space ($E \cong E^{**}$), the weak topology on E coincides exactly with the weak* topology on E^{**}. By the Banach-Alaoglu Theorem, the norm-closed ball $\{\|x\|_E \leq R\}$ is weakly compact. Furthermore, from Step 1, f is weakly l.s.c., which guarantees that its sublevel set K_t is weakly closed (by Corollary 4.1.7). Because K_t is a weakly closed subset of a weakly compact set (the closed bounded ball), K_t itself must be weakly compact. This confirms that f is weakly coercive. ■

Remark 4.2.7. Note that the theorem still holds even when we apply to a function $f : C \to \mathbb{R} \cup \{+\infty\}$ where $C \subseteq E$ is a convex set by considering $\tilde{f} : E \to \mathbb{R} \cup \{+\infty\}$ as

$$\tilde{f}(x) := \begin{cases} f(x) & , \ x \in C \\ +\infty & , \ x \notin C. \end{cases}$$

The function $\tilde{f}$ still maintains the same property we needed (convex, strongly l.s.c., strongly coercive).

Remark 4.2.8. Hypotheses 1. and 2. of Theorem 4.2.6 play well with weak topology by Theorem 4.2.3 and hypothesis 3. plays well with weak* topology by Corollary 4.2.5. This is the reason we require E to be reflexive to make the two topologies coincide.

5 Recap on L^p Space

> **Notes.** Actually, some of the contents written here were actually given during lecture 4 in the class. However, I decided to put it here for coherency.

§5.1 Preliminary Properties of L^p Space

Let $\Omega \subseteq \mathbb{R}^d$, $p \in [1, \infty]$. Define

$$L^p(\Omega; \mathbb{R}^\ell) = \begin{cases} \left\{ u : \Omega \to \mathbb{R}^\ell : \int_\Omega |u|^p < \infty \right\} & \text{if } p < \infty \\ \left\{ u : \Omega \to \mathbb{R}^\ell : \operatorname{ess\,sup}_\Omega |u| < \infty \right\} & \text{if } p = \infty \end{cases}$$

L^p is a Banach space with norm

$$\|u\|_p = \|u\|_{L^p} = \left(\int_\Omega |u|^p \right)^{1/p}, \quad p < \infty$$
$$\|u\|_\infty = \|u\|_{L^\infty} = \operatorname{ess\,sup}_\Omega |u| := \inf\{m \geq 0 : \mathcal{L}(\{|u| > m\}) = 0\}, \quad p = \infty$$

For $1 \leq p < \infty$, $(L^p(\Omega; \mathbb{R}^\ell))^* = L^{p'}(\Omega; \mathbb{R}^\ell)$ where $\dfrac{1}{p} + \dfrac{1}{p'} = 1$.
So, in particular

$$(L^1)^* = L^\infty \quad \text{and} \quad L^p \text{ is reflexive for } 1 < p < \infty$$

Therefore, we can define weak (and weak*) convergence for any sequence $\{u_n\} \subseteq L^p(\Omega; \mathbb{R}^\ell)$:

$$u_n \rightharpoonup u \text{ in } L^p \ (p < \infty) \quad \text{if} \quad \int \langle u_n, v \rangle \to \int \langle u, v \rangle \quad \forall v \in L^{p'}$$
$$u_n \overset{*}{\rightharpoonup} u \text{ in } L^\infty \quad \text{if} \quad \int \langle u_n, v \rangle \to \int \langle u, v \rangle \quad \forall v \in L^1$$

Theorem 5.1.1

Let $\{u_n\} \subseteq L^p(\Omega; \mathbb{R}^\ell)$, we have $u_n \rightharpoonup u$ in $L^p, 1 \leq p < \infty$ if and only if

1. $\sup_n \|u_n\|_p < \infty$

and one of the following two conditions hold:

a. $\displaystyle\int \langle u_n, \varphi \rangle \to \int \langle u, \varphi \rangle \quad \forall \varphi \in C_c^\infty(\Omega; \mathbb{R}^\ell)$

b. $\displaystyle\int_E u_n \to \int_E u \quad \forall E$ Borel (or measurable)

Proof of 1. + a. $\implies$ weak convergence. Let $v \in L^{p'}$. By density, $\exists \{\varphi_k\} \subseteq C_c^\infty(\Omega, \mathbb{R}^\ell)$ s.t. $\varphi_k \to v$ in $L^{p'}$.
Therefore, for any k, $n \in \mathbb{N}$:

$$\left| \int \langle u_n, v \rangle - \int \langle u, v \rangle \right| \leq \left| \int \langle u_n, v - \varphi_k \rangle - \int \langle u, v - \varphi_k \rangle \right| + \left| \int \langle u_n - u, \varphi_\eta \rangle \right|$$
$$\overset{\text{Hölder}}{\leq} \left(\sup_n \|u_n\|_p + \|u\|_p \right) \cdot \|v - \varphi_k\|_{L^{p'}} + \left| \int \langle u_n - u, \varphi_k \rangle \right|.$$

Thus, passing $n \to \infty$ and then $k \to \infty$ yields

$$\limsup_{n\to\infty} \left| \int \langle u_n, v \rangle - \int \langle u, v \rangle \right| = 0.$$

Proof of 1. + 2. $\implies$ weak convergence is applied to convergence against simple functions, and we recall the fact that the set of simple functions is dense. ■

Remark 5.1.2. Similar conclusion applies to $p = \infty$ by substituting weak convergence with weak* convergence.

§5.2 Weak Convergence versus Strong Convergence

Theorem 5.2.1 (Vitali)
Let $\{u_k\} \subseteq L^p(\mathbb{R}^d; \mathbb{R}^\ell)$ be a sequence of L^p functions such that

1. $u_k \to u$ in measure $(\forall \varepsilon > 0, |\{|u_k - u| > \varepsilon\}| \to 0)$
   (this implies a.e. pointwise convergence up to subsequence)
2. $\displaystyle \lim_{R\to\infty} \left(\sup_k \int_{\mathbb{R}^d \setminus B_R} |u_k|^p \right) = 0$
3. $\displaystyle \lim_{M\to\infty} \left(\sup_k \int_{\{|u_k| > M\}} |u_k|^p \right) = 0$

Then $u_k \to u$ strongly in L^p.

Proof. Exercise. ■

Remark 5.2.2. Note that if $u_n \to u$ strongly in L^p, then 1., 2., and 3. are satisfied.

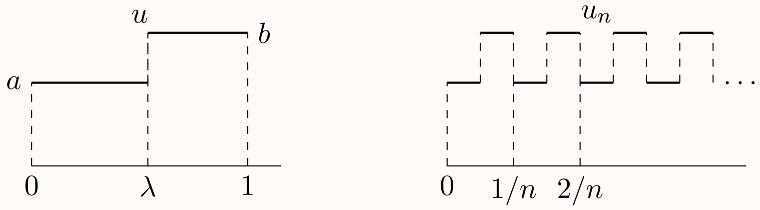
Example 5.2.3 (Examples when only weak convergence hold)
Let $p \in (1, +\infty)$

1. $u_k = \dfrac{1}{k^{1/p}} \mathbf{1}_{[0,1/k]} \overset{L^p}{\rightharpoonup} 0$ but not strongly.
2. $u_k = \mathbf{1}_{[k,k+1]} \rightharpoonup 0$ in L^p but not strongly.
3. $u_k = \varphi(x - ke_i)$, $\quad k \in \mathbb{N}$, $\quad \varphi \in C_c^\infty(B)$, we have $u_k \overset{L^p}{\rightharpoonup} u \equiv 0$ but not strongly
4. $u_k = k^{d/p}\varphi(kx)$, we have $u_k \overset{L^p}{\rightharpoonup} u \equiv 0 \quad (1 < p < \infty)$ but not strongly

Example 5.2.4

Let $p \in [1, +\infty]$ and consider

$$u(x) = \begin{cases} a, & x \in [0, \lambda] \\ b, & x \in (\lambda, 1] \end{cases}$$



and extend u by 1-periodicity, $u_n(x) = u(nx)$. $\{u_n\} \subseteq L^\infty([0,1])$ (or $L^p([0,1])$). One can check for any interval $I \subseteq [0,1]$,

$$\int_I u_n \to |I| \cdot (\lambda a + (1-\lambda)b).$$

This can be done by considering $I = [c, d]$ and notice we will have $N = \left\lfloor \dfrac{d-c}{n} \right\rfloor$ full periods and a leftover for which the "leftover" pieces at the ends of the interval have a width of at most $\dfrac{2}{n}$.

Since the class of simple functions is dense and the Lebesgue measure is Borel regular, we can extend our obtained result to hold for any E Borel and condition 1. + b. holds for Theorem 5.1.1 to fully assert

$$\implies u_n \xrightarrow[\text{or } L^\infty]{L^p} u \equiv \lambda a + (1-\lambda)b$$

Exercise (Requires Fourier Analysis)

Let $\{f_k\}$ be a uniformly bounded sequence in $L^2(\mathbb{R}^d)$, and let $f \in L^2(\mathbb{R}^d)$. Denote by $\hat{f}$ the Fourier transform of f in $L^2(\mathbb{R}^d)$.

Prove the following two statements:

Part I: If the sequence $\{f_k\}$ converges weakly to f in $L^2(\mathbb{R}^d)$ (denoted $f_k \rightharpoonup f$), then the sequence of Fourier transforms converges pointwise:

$$\lim_{k\to\infty} \hat{f}_k(\xi) = \hat{f}(\xi) \quad \text{for almost every } \xi \in \mathbb{R}^d.$$

Part II: The sequence $\{f_k\}$ converges strongly to f in $L^2(\mathbb{R}^d)$ if and only if the following three conditions are simultaneously satisfied:

(i) **Pointwise Convergence in Frequency:** $\lim\limits_{k\to\infty} \hat{f}_k(\xi) = \hat{f}(\xi)$ for almost every $\xi \in \mathbb{R}^d$.

(ii) **No Escape of Mass to Spatial Infinity (Tightness in Space):**

For every $\varepsilon > 0$, there exists a radius $R > 0$ such that

$$\sup_{k\in\mathbb{N}} \int_{\mathbb{R}^d\setminus B(0,R)} |f_k(x)|^2\,dx < \varepsilon.$$

(iii) **No Escape of Mass to Frequency Infinity (Tightness in Frequency):** For every $\varepsilon > 0$, there exists a radius $R > 0$ such that

$$\sup_{k\in\mathbb{N}} \int_{\mathbb{R}^d\setminus B(0,R)} |\hat{f}_k(\xi)|^2\,d\xi < \varepsilon.$$

Proof. Omitted, it requires Fourier analysis. ■

Actually, Example 5.2.4 is encoded in the Riemann-Lebesgue Lemma which stated as follows.

Lemma 5.2.5 (Riemann-Lebesgue / Periodic Homogenization)
Let $Q = [0,1)^d$ be the unit cube in $\mathbb{R}^d$, and let $p \in [1,\infty]$. Suppose $V \in L^p_{loc}(\mathbb{R}^d;\mathbb{R}^m)$ is a $\mathbb{Z}^d$-periodic function. Define the sequence of highly oscillatory functions:

$$u_k(x) = V(kx) \quad \text{for } k \in \mathbb{N}.$$

Then, for any bounded measurable domain $\Omega \subset \mathbb{R}^d$, the sequence $\{u_k\}$ converges to the constant average of V:

$$\bar{u}(x) \equiv \frac{1}{|Q|}\int_Q V(y)\,dy \quad \text{(note: } |Q| = 1)$$

in the following sense:

1. If $1 \le p < \infty$, then $u_k \rightharpoonup \bar{u}$ weakly in $L^p(\Omega;\mathbb{R}^m)$.
2. If $p = \infty$, then $u_k \overset{*}{\rightharpoonup} \bar{u}$ weakly* in $L^\infty(\Omega;\mathbb{R}^m)$.

Proof. We shall consider the proof for $p \in (1,+\infty)$, $d = 1$, and $\Omega = [0,1]$ for simplicity. Here, we want to apply condition 1. + a. from Theorem 5.1.1.

Step 1: Proving condition 1.
For any $k \in \mathbb{N}$, we do uniform partition of $[0,1]$ into k intervals to have

$$\int_0^1 |u_k(x)|^p\,dx = \sum_{i=1}^k \int_{\frac{i-1}{k}}^{\frac{i}{k}} |V(kx)|^p\,dx = \sum_{i=1}^k \frac{1}{k}\int_0^1 |V(y)|^p\,dy = \int_0^1 |V(y)|^p\,dy < +\infty.$$

Hence, $\sup_k \int_0^1 |u_k|^p\,dx < +\infty$.

Step 2: Proving condition a.
Let $\varphi \in C_c^\infty([0,1])$, we shall prove

$$\int_0^1 \langle u_k, \varphi\rangle\,dx \to \left\langle \int_0^1 V\,dx, \int_0^1 \varphi\,dx \right\rangle.$$

For any $k \in \mathbb{N}$, we have

$$\int_0^1 \langle u_k, \varphi \rangle \, dx = \sum_{i=1}^k \int_{\frac{i-1}{k}}^{\frac{i}{k}} \langle V(kx), \varphi(x) \rangle \, dx$$
$$= \underbrace{\sum_{i=1}^k \left\langle \int_{\frac{i-1}{k}}^{\frac{i}{k}} V(kx)\, dx, \varphi\left(\frac{i}{k}\right) \right\rangle}_{I_1} + \underbrace{\sum_{i=1}^k \int_{\frac{i-1}{k}}^{\frac{i}{k}} \left\langle V(kx), \varphi(x) - \varphi\left(\frac{i}{k}\right) \right\rangle \, dx}_{I_2}$$

where

$$I_1 = \sum_{i=1}^k \frac{1}{k} \left\langle \int_0^1 V(y)\, dy, \varphi\left(\frac{i}{k}\right) \right\rangle$$
$$= \left\langle \int_0^1 V(y)\, dy, \sum_{i=1}^k \frac{1}{k} \varphi\left(\frac{i}{k}\right) \right\rangle \to \left\langle \int_0^1 V\, dx, \int_0^1 \varphi\, dx \right\rangle.$$

and by Cauchy-Schwarz on inner product and mean value inequality:

$$|I_2| \leq \sum_{i=1}^k \int_{\frac{i-1}{k}}^{\frac{i}{k}} |V(kx)| \cdot \|\varphi'\|_\infty \frac{1}{k}\, dx = \sum_{i=1}^k \frac{\|\varphi'\|_\infty}{k^2} \int_0^1 |V(y)|\, dy \to 0.$$

Hence,

$$\int_0^1 \langle u_k, \varphi \rangle \, dx \to \left\langle \int_0^1 V\, dx, \int_0^1 \varphi\, dx \right\rangle.$$

■

Theorem 5.2.6

Let $\Omega \subset \mathbb{R}^d$ be a bounded measurable set, and let $p \in (1, \infty)$. Suppose $f : \mathbb{R}^\ell \to \mathbb{R} \cup \{+\infty\}$ is a Borel measurable function that is bounded from below.

Define the integral functional $\mathcal{F} : L^p(\Omega; \mathbb{R}^\ell) \to \mathbb{R} \cup \{+\infty\}$ by

$$\mathcal{F}(u) = \int_\Omega f(u(x))\, dx.$$

Then, the following statements are equivalent:

1. $\mathcal{F}$ is l.s.c. with respect to the weak topology on $L^p(\Omega; \mathbb{R}^\ell)$.
2. f is convex and lower semi-continuous on $\mathbb{R}^\ell$.

Proof. **(2. $\implies$ 1.):** Fortunately, L^p is a Banach space. We can utilize Theorem 4.2.3 to immediately prove this direction. Since f is convex, $\mathcal{F}$ is trivially convex. Furthermore, because f is l.s.c. and bounded from below (and Ω is bounded), we can apply Fatou's Lemma to deduce that the map

$$L^p \ni u \mapsto \int_\Omega f(u(x))\, dx$$

is l.s.c. in the strong topology of L^p. Since $\mathcal{F}$ is convex and strongly l.s.c., it is l.s.c. with respect to the weak topology.

(1. $\implies$ 2.): To prove f is l.s.c., let $y \in \mathbb{R}^\ell$ and $\{y_k\} \subseteq \mathbb{R}^\ell$ such that $y_k \to y$. By taking $u_k \equiv y_k \in L^p(\Omega;\mathbb{R}^\ell)$ and $u \equiv y \in L^p(\Omega;\mathbb{R}^\ell)$, we have $u_k \to u$ strongly (and thus weakly), and by the fact $\mathcal{F}$ is l.s.c.:

$$|\Omega|\, f(y) = \mathcal{F}(u) \le \liminf_{k\to\infty} \mathcal{F}(u_k) = |\Omega| \liminf_{k\to\infty} f(y_k).$$

Now, we only need to prove f is convex. Consider the unit cube $Q = [0,1)^d$. Let $y_1, y_2 \in \mathbb{R}^\ell$ and take any $E \subseteq Q$ such that $|E| = \lambda \in (0,1)$ (one option is $E = \sqrt[d]{\lambda} Q$) and consider $\tilde{v} = y_1 \mathbf{1}_E + y_2 \mathbf{1}_{Q\setminus E}$.

Let v be the $\mathbb{Z}^d$ periodic extension of $\tilde{v}$ and define

$$u_k(x) := v(kx)|_\Omega.$$

Hence, by the Riemann-Lebesgue Lemma (Lemma 5.2.5) we have

$$u_k \rightharpoonup u \equiv \int_Q v(z)\, dz = \lambda y_1 + (1-\lambda) y_2.$$

Notice that the composition $f(v(x))$ is also $\mathbb{Z}^d$-periodic. Again, we apply Riemann-Lebesgue to $f(v(kx))$ to obtain

$$f(v(k\cdot)) \rightharpoonup \int_Q f(v(z))\, dz = \lambda f(y_1) + (1-\lambda) f(y_2).$$

By choosing the test function $\mathbf{1}_\Omega \in L^{p'}(\Omega)$, weak convergence implies

$$\mathcal{F}(u_k) = \int_\Omega f(u_k(x))\, dx \to |\Omega| \left(\lambda f(y_1) + (1-\lambda) f(y_2)\right).$$

Hence, by the weak lower semicontinuity of $\mathcal{F}$ we have

$$|\Omega|\, f(\lambda y_1 + (1-\lambda) y_2) = \mathcal{F}(u) \le \liminf_{k\to\infty} \mathcal{F}(u_k) = |\Omega| \left(\lambda f(y_1) + (1-\lambda) f(y_2)\right)$$

Dividing by $|\Omega|$, we conclude $f(\lambda y_1 + (1-\lambda) y_2) \le \lambda f(y_1) + (1-\lambda) f(y_2)$. ■

Remark 5.2.7 (The Bolza Example and Convexity). The Bolza problem perfectly illustrates why convexity with respect to the derivative (momentum) variable is strictly required for weak lower semicontinuity. Consider the minimization problem:

$$\inf \left\{ \int_0^1 \left(((u')^2 - 1)^2 + u^2 \right) dx : u(0) = 0, u(1) = 0 \right\}$$

The Lagrangian $F(x,u,u') = ((u')^2-1)^2 + u^2$ is **not convex** with respect to u' because of the "double-well potential" $((u')^2-1)^2$, which strongly prefers slopes of exactly $+1$ or -1. Simultaneously, the u^2 term prefers the function itself to be exactly 0. To minimize both terms, an infimizing sequence $\{u_k\}$ must oscillate wildly between slopes of $+1$ and -1 (creating a microscopic zigzag/sawtooth pattern) to keep the derivative energy near zero, while compressing the amplitude of these zigzags so that the function stays arbitrarily close to the x-axis ($u_k \approx 0$).

As $k \to \infty$, this sequence converges uniformly (and thus weakly in $W^{1,2}$) to the flat line:

$$\bar{u}(x) \equiv 0$$

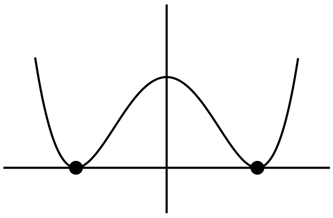


Figure 5.1: The graph of the double-well potential.

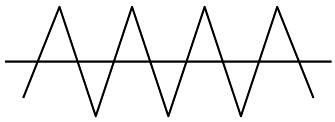


Figure 5.2: The graph of the minimizer.

However, if we evaluate the energy of this limit function:

$$\mathcal{F}(\bar{u}) = \int_0^1 \left((0-1)^2 + 0^2\right) dx = \int_0^1 1\, dx = 1$$

While the limit of the sequence's energies is $\lim_{k\to\infty} \mathcal{F}(u_k) = 0$, the energy of the limit function is 1. Because

$$\mathcal{F}(\bar{u}) > \liminf_{k\to\infty} \mathcal{F}(u_k),$$

the functional fails to be weakly lower semi-continuous. The minimum is never attained in $W^{1,2}$ strictly because the double-well potential breaks convexity.

6 Recap on Sobolev Space

The needed proofs for the mentioned results in Section 6.1 can be consulted from [Eva10] and [Bre11]. However, those two references only considered the scalar-valued result. Nonetheless, the generalization to vector-valued functions follows canonically.

§6.1 Preliminary Properties of Sobolev Space

Let $\Omega \subseteq \mathbb{R}^d$ be an open set (often assumed to have a Lipschitz boundary for embedding and trace theorems, but not required for basic definitions).

Definition 6.1.1 (Sobolev Space $W^{1,p}$)
For $1 \leq p \leq \infty$, the Sobolev space $W^{1,p}(\Omega; \mathbb{R}^\ell)$ is defined as:

$$\left\{ u \in L^p(\Omega; \mathbb{R}^\ell) : \begin{array}{l} \exists g \in L^p(\Omega; \mathbb{R}^{\ell \times d}) \text{ such that} \\ \displaystyle\int_\Omega u_i(x) \partial_j \varphi(x)\, dx = - \int_\Omega g_{ij}(x) \varphi(x)\, dx \quad \forall \varphi \in C_c^\infty(\Omega; \mathbb{R}) \end{array} \right\}.$$

The function g is unique (up to a set of measure zero), and we denote the weak derivative as $Du = g$.

We define the Sobolev norm to the Sobolev space $W^{1,p}$ as

$$\|u\|_{W^{1,p}(\Omega;\mathbb{R}^\ell)} = \|u\|_p + \|Du\|_p \simeq \left(\int_\Omega |u|^p\, dx + \int_\Omega |Du|^p\, dx \right)^{1/p}.$$

Remark 6.1.2.

1. $(W^{1,p}(\Omega; \mathbb{R}^\ell), \|\cdot\|_{W^{1,p}(\Omega;\mathbb{R}^\ell)})$ is a Banach space.
2. $W^{1,2}(\Omega; \mathbb{R}^\ell)$ is a Hilbert space with

   $$\langle u, w \rangle_{W^{1,2}(\Omega;\mathbb{R}^\ell)} = \int_\Omega \langle u, v \rangle + \langle Du, Dv \rangle \, dx$$

   where $\langle Du, Dv \rangle$ is the Frobenius inner product.
3. Alternatively, one can define the space $H^{1,p}(\Omega; \mathbb{R}^\ell)$ as the completion of smooth functions $C^\infty(\Omega; \mathbb{R}^\ell)$ with respect to the Sobolev norm $\|\cdot\|_{W^{1,p}(\Omega;\mathbb{R}^\ell))}$ and obviously $H^{1,p}(\Omega; \mathbb{R}^\ell) \subseteq W^{1,p}(\Omega; \mathbb{R}^\ell)$

Theorem 6.1.3 (Meyers-Serrin, 1964)
Let $\Omega \subseteq \mathbb{R}^d$ be **any open set**. *(Crucially, no regularity on the boundary $\partial\Omega$ is required).* Then smooth functions are dense in $W^{1,p}(\Omega)$, meaning:

$$H^{1,p}(\Omega; \mathbb{R}^\ell) = W^{1,p}(\Omega; \mathbb{R}^\ell) \quad \text{for } 1 \leq p < \infty.$$

Remark 6.1.4. For Meyers-Serrin Theorem, it is important we define $H^{1,p}(\Omega;\mathbb{R}^\ell)$ on completion over $C^\infty(\Omega;\mathbb{R}^\ell)$ and not other spaces. For instance, approximation by $C^1(\overline{\Omega};\mathbb{R}^\ell)$ generally fails if Ω has an irregular or disconnected boundary.

Example (The Split Disc): Let $D \subset \mathbb{R}^2$ be the open unit disc and $L = \{(x,0) : -1 < x < 1\}$ be its horizontal diameter. Define $\Omega = D \setminus L$. The function

$$u(x,y) = \begin{cases} 1 & \text{if } y > 0 \\ 0 & \text{if } y < 0 \end{cases}$$

satisfies $Du = 0$ inside Ω, so $u \in W^{1,p}(\Omega)$. However, u cannot be approximated by any sequence $v_k \in C^\infty(\bar{\Omega})$, because v_k must be uniformly continuous across the slit L, which would force the gradient norm $\|Dv_k\|_{L^p} \to \infty$ to accommodate the jump discontinuity.

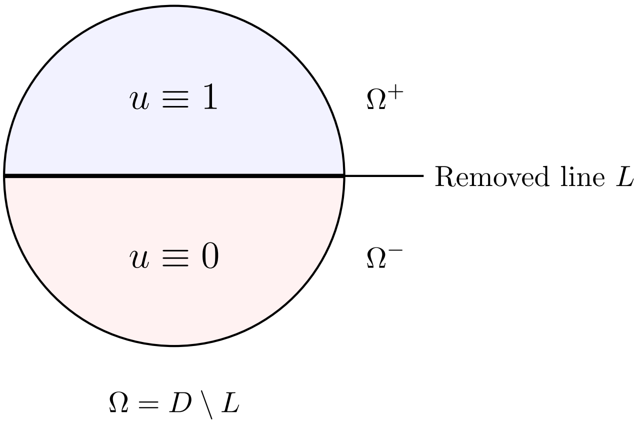


Figure 6.1: The split disc domain. A jump discontinuity across L prevents approximation by $C^\infty(\bar{\Omega})$ functions.

Theorem 6.1.5 (Properties on Lipschitz Domains)

If $\Omega \subset \mathbb{R}^d$ is a bounded open set with a Lipschitz boundary, then:

1. **Global Extension:** There exists a continuous, linear extension operator $\text{Ext} : W^{1,p}(\Omega;\mathbb{R}^\ell) \to W^{1,p}(\mathbb{R}^d;\mathbb{R}^\ell)$ such that $\text{Ext}(u)|_\Omega = u$ a.e.
2. **Density:** $C^1(\bar{\Omega};\mathbb{R}^\ell)$ is dense in $W^{1,p}(\Omega;\mathbb{R}^\ell)$.
3. **Trace Operator:** There exists a unique continuous linear operator $\text{tr} : W^{1,p}(\Omega;\mathbb{R}^\ell) \to L^p(\partial\Omega;\mathbb{R}^\ell)$ extending the classical restriction $\text{tr}(u) = u|_{\partial\Omega}$ for $u \in C^0(\bar{\Omega}) \cap W^{1,p}(\Omega)$.
4. **Gauss-Green Formula:** For all $u \in W^{1,p}(\Omega;\mathbb{R}^\ell)$ and $\varphi \in C^1(\bar{\Omega})$:

$$\int_\Omega \varphi Du \, dx = -\int_\Omega u \otimes D\varphi \, dx + \int_{\partial\Omega} \varphi \, \text{tr}(u) \otimes \nu \, d\mathcal{H}^{d-1}$$

where ν is the outward unit normal.

Definition 6.1.6 (Zero-Trace Sobolev Space)
We define $W_0^{1,p}(\Omega;\mathbb{R}^\ell)$ as the closure of compactly supported smooth functions under the Sobolev norm:

$$W_0^{1,p}(\Omega;\mathbb{R}^\ell) = \overline{C_c^\infty(\Omega;\mathbb{R}^\ell)}^{\|\cdot\|_{W^{1,p}}}$$

The space $W_0^{1,p}(\Omega)$ is defined as the closure of $C_c^\infty(\Omega)$ in $W^{1,p}(\Omega)$. For Lipschitz domains, this coincides exactly with the kernel of the trace operator:

$$W_0^{1,p}(\Omega) = \left\{u \in W^{1,p}(\Omega) : \operatorname{tr}(u) = 0 \text{ on } \partial\Omega\right\}$$

For variational problems with a non-zero boundary condition $g \in W^{1,p}(\Omega)$, the set of admissible functions forms the affine space:

$$g + W_0^{1,p}(\Omega) = \left\{u \in W^{1,p}(\Omega) : \operatorname{tr}(u) = \operatorname{tr}(g) \text{ on } \partial\Omega\right\}.$$

Theorem 6.1.7 (The Fundamental Sobolev Embedding)
For the whole space $\mathbb{R}^d$, the fundamental Sobolev embeddings are given by:

$$W^{1,p}(\mathbb{R}^d;\mathbb{R}^\ell) \hookrightarrow \begin{cases} L^{p^*}(\mathbb{R}^d;\mathbb{R}^\ell), & \text{if } 1 \le p < d \\ L^q_{\text{loc}}(\mathbb{R}^d;\mathbb{R}^\ell)\ \forall q \in [1,\infty) \text{ and } L^q\ \forall q \in [p,\infty), & \text{if } p = d \\ C^{0,1-d/p}(\mathbb{R}^d;\mathbb{R}^\ell), & \text{if } p > d \end{cases}$$

where $\dfrac{1}{p^*} = \dfrac{1}{p} - \dfrac{1}{d}$.

In particular:

1. $W_0^{1,p}(\Omega;\mathbb{R}^\ell)$ embeds in the same spaces for **any** open set Ω.
2. $W^{1,p}(\Omega;\mathbb{R}^\ell)$ embeds in the same spaces for any Ω with a **Lipschitz boundary**.

Remark 6.1.8 (Regarding 1.). We can always think of $W_0^{1,p}(\Omega;\mathbb{R}^\ell)$ as a subset of $W^{1,p}(\mathbb{R}^d;\mathbb{R}^\ell)$ via the "zero" extension outside of Ω.

In particular, we have the general local embedding:

$$W^{1,p}(\mathbb{R}^d;\mathbb{R}^\ell) \hookrightarrow L^q_{\text{loc}}(\mathbb{R}^d;\mathbb{R}^\ell) \quad \forall q < p^* \quad (\text{where } p^* = \infty \text{ if } p = d) \tag{6.1}$$

Theorem 6.1.9 (Rellich-Kondrachov)
The embedding (6.1) is compact. That is, if $\{u_k\}$ is a bounded sequence in $W^{1,p}(\mathbb{R}^d;\mathbb{R}^\ell)$, then there exists a subsequence $\{u_{k'}\}$ and a function $u \in W^{1,p}(\mathbb{R}^d;\mathbb{R}^\ell)$ such that for any ball $B_R \subseteq \mathbb{R}^d$:

$$\|u_{k'} - u\|_{L^q(B_R;\mathbb{R}^\ell)} \to 0 \quad \text{as } k' \to \infty$$

Furthermore, if $p \ge d$, the same is true for $q = \infty$ (implying uniform convergence on compact sets).

Remark 6.1.10. Note in particular that if Ω is **bounded**, we have the compact embeddings:

$$W_0^{1,p}(\Omega) \underset{\text{compact}}{\hookrightarrow} \begin{cases} L^q(\Omega) & \forall 1 \le q < p^*, \quad p < d \\ L^q(\Omega) & \forall q \in [1,\infty), \quad p = d \\ L^\infty(\Omega) & p > d \end{cases}$$

The exact same compact embeddings hold for the standard space $W^{1,p}(\Omega)$ provided the bounded domain Ω is Lipschitz.

Theorem 6.1.11 (Fundamental Sobolev Inequalities)
For 1. and 2., let $u \in W^{1,p}(\mathbb{R}^d;\mathbb{R}^\ell)$.

1. **Sobolev Inequality:**

$$\|u\|_{L^{p^*}(\mathbb{R}^d;\mathbb{R}^\ell)} \le S\|Du\|_{L^p(\mathbb{R}^d)(\mathbb{R}^d;\mathbb{R}^\ell)} \quad \text{for } 1 < p < d, \frac{1}{p^*} = \frac{1}{p} - \frac{1}{d}$$

2. **Morrey Inequality:**

$$[u]_{C^{0,1-d/p}(\mathbb{R}^d;\mathbb{R}^\ell)} := \sup_{\substack{x,y\in\mathbb{R}^d \\ x\neq y}} \frac{|f(x)-f(y)|}{|x-y|^{1-d/p}} \le S\|Du\|_{L^p(\mathbb{R}^d;\mathbb{R}^\ell)} \quad \text{for } p > d$$

3. **Poincaré Inequalities:** Let $\Omega \subset \mathbb{R}^d$ be a bounded domain.
   (I) For functions with zero trace, $\forall u \in W_0^{1,p}(\Omega;\mathbb{R}^\ell)$:

$$\|u\|_{L^p(\Omega;\mathbb{R}^\ell)} \le C|\Omega|^{1/d}\|Du\|_{L^p(\Omega;\mathbb{R}^\ell)}$$

   (II) If Ω is Lipschitz and connected, then $\forall u \in W^{1,p}(\Omega)$:

$$\left\| u - \fint_\Omega u \right\|_{L^p(\Omega;\mathbb{R}^\ell)} \le C(\Omega)\cdot \|Du\|_{L^p(\Omega;\mathbb{R}^\ell)}$$

   (Note: Dimensionally, the constant $C(\Omega) \sim$ length).
   (III) In particular, if Ω is a ball $B_R(x_0)$, the inequality scales precisely with the radius:

$$\left\| u - \fint_{B_R} u \right\|_{L^p(\Omega;\mathbb{R}^\ell)} \le C(d)\cdot R \cdot \|Du\|_{L^p(\Omega;\mathbb{R}^\ell)} \tag{6.2}$$

Exercise 6.1.12
Use (6.2) to check that if $u \in W^{1,p}(B_1)$ and $u \equiv 0$ on a set $E \subseteq B_1$ with $|E| \ge \mu\,|B_1|$ for some $\mu \in (0,1)$ then

$$\|u\|_{L^p(B_1)} \le \frac{C(d)}{\mu}\,\|Du\|_{L^p(B_1)}\,.$$

Solution. Let $\bar{u} = \fint_{B_1} u\,dx$. Since $u \equiv 0$ almost everywhere on E, we have $|u - \bar{u}| = |\bar{u}|$ on E. Integrating this over E provides a bound on the average:

$$|\bar{u}|^p|E| = \int_E |u - \bar{u}|^p\,dx \leq \int_{B_1} |u - \bar{u}|^p\,dx = \|u - \bar{u}\|^p_{L^p(B_1)}$$

Using the hypothesis $|E| \geq \mu|B_1|$, we can evaluate the L^p norm of the constant function $\bar{u}$ over the entire ball B_1:

$$\|\bar{u}\|^p_{L^p(B_1)} = |\bar{u}|^p|B_1| \leq \frac{|B_1|}{|E|}\|u - \bar{u}\|^p_{L^p(B_1)} \leq \frac{1}{\mu}\|u - \bar{u}\|^p_{L^p(B_1)}$$

Taking the p-th root yields $\|\bar{u}\|_{L^p(B_1)} \leq \mu^{-1/p}\|u - \bar{u}\|_{L^p(B_1)}$.

By Minkowski's inequality and the Poincaré inequality (6.2) applied to B_1 ($R = 1$), we bound u:

$$\begin{aligned}\|u\|_{L^p(B_1)} \leq \|u - \bar{u}\|_{L^p(B_1)} + \|\bar{u}\|_{L^p(B_1)} &\leq \left(1 + \mu^{-1/p}\right)\|u - \bar{u}\|_{L^p(B_1)} \\ &\leq \left(1 + \mu^{-1/p}\right) C(d)\|Du\|_{L^p(B_1)}\end{aligned}$$

Finally, since $\mu \in (0, 1)$ and $p \geq 1$, we can simplify the prefactor: $1 + \mu^{-1/p} \leq 1 + \mu^{-1} \leq 2\mu^{-1}$. Absorbing the constant 2 into $C(d)$ yields the desired result:

$$\|u\|_{L^p(B_1)} \leq \frac{\tilde{C}(d)}{\mu}\|Du\|_{L^p(B_1)}$$

■

Remark 6.1.13 (Weak Convergence). For $1 < p < \infty$, the Sobolev space $W^{1,p}(\Omega; \mathbb{R}^\ell)$ (and its zero-trace counterpart $W^{1,p}_0(\Omega; \mathbb{R}^\ell)$) is a reflexive Banach space. The weak topology is characterized as follows:

$$u_k \rightharpoonup u \text{ in } W^{1,p} \iff \begin{cases} \int_\Omega \langle u_k, v\rangle\,dx \to \int_\Omega \langle u_k, v\rangle\,dx & \forall v \in L^{p'}(\Omega; \mathbb{R}^\ell) \\ \int_\Omega \langle Du_k, F\rangle\,dx \to \int_\Omega \langle Du, F\rangle\,dx & \forall F \in L^{p'}(\Omega; \mathbb{R}^\ell \otimes \mathbb{R}^d) \end{cases}$$

Remark 6.1.14 (Compactness). By the Rellich-Kondrachov compactness theorem (and the uniform boundedness principle), if $u_k \rightharpoonup u$ weakly in $W^{1,p}$, then $u_k \to u$ strongly in L^q_{loc}. If Ω is bounded and Lipschitz, this strong convergence holds globally in $L^q(\Omega)$ for $1 \leq q < p^*$.

§6.2 Application of the Sobolev Space to the Abstract Direct Methods

Theorem 6.2.1 (Existence of Minimizers)
Let $\Omega \subset \mathbb{R}^d$ be a bounded open set with Lipschitz boundary. Let $f : \Omega \times \mathbb{R}^\ell \times \mathbb{R}^{\ell \times d} \to \mathbb{R}$ be a Lagrangian $f(x, z, \xi)$ satisfying:

1. $x \mapsto f(x, z, \xi)$ is Lebesgue measurable.
2. $(z, \xi) \mapsto f(x, z, \xi)$ is continuous and convex for a.e. $x \in \Omega$.
3. **Coercivity:** $|f(x, z, \xi)| \geq C|\xi|^p - M|z|^q - h(x)$ for some $1 \leq q < p$, C, $M > 0$, and $h \in L^1(\Omega)$.

Then, the integral functional $\mathcal{F}(u) = \displaystyle\int_\Omega f(x, u(x), Du(x))\, dx$ is well-defined. Furthermore, for any boundary data $g \in W^{1,p}(\Omega; \mathbb{R}^\ell)$, there exists a minimum for the problem:

$$\inf_{u \in g + W_0^{1,p}(\Omega; \mathbb{R}^\ell)} \mathcal{F}(u)$$

Remark 6.2.2. We refer condition 1. and 2. (without convexity) as the Carathéodory conditions.

Remark 6.2.3. This coercivity coniditon (3.) is a generalization of the coercivity mentioned in the lectures. This coercivity condition follows from Theorem 3.30 of [Dac07].

Proposition 6.2.4 (Euler-Lagrange Equations and Further Properties)
Suppose the hypotheses of the previous theorem hold, and let $\bar{u} \in g + W_0^{1,p}(\Omega; \mathbb{R}^\ell)$ be a minimizer of $\mathcal{F}$. Depending on further growth conditions on f, the following properties hold:

(a) **Absence of Lavrentiev Phenomenon:** If we additionally assume the upper bound $|f(x, z, \xi)| \le \alpha(x) + C(|\xi|^p + |z|^p)$ where $\alpha \in L^1(\Omega)$, then the infimum over smooth functions matches the Sobolev minimum:

$$\inf_{\substack{u \in C^1(\bar{\Omega};\mathbb{R}^\ell) \\ u=g \text{ on } \partial\Omega}} \mathcal{F}(u) = \min_{\substack{u \in W^{1,p}(\Omega;\mathbb{R}^\ell) \\ u=g \text{ on } \partial\Omega}} \mathcal{F}(u)$$

(b) **Euler-Lagrange (E-L) Equations:** If

$$|D_z f| + |D_\xi f| \le \beta(x) + C\left(|\xi|^{p-1} + |z|^{p-1}\right)$$

where $\beta \in L^{p'}(\Omega)$ and $C \ge 0$, then any minimizer $\bar{u}$ satisfies the E-L equations in weak form:

$$\int_\Omega \langle D_z f(x, \bar{u}, D\bar{u}), \varphi \rangle + \langle D_\xi f(x, \bar{u}, D\bar{u}), D\varphi \rangle \, dx = 0 \quad \forall \varphi \in W_0^{1,p}(\Omega; \mathbb{R}^\ell)$$

(c) **Sufficiency:** By convexity of f, any solution to the weak (E-L) equation is a minimizer.

(d) **Uniqueness:** If the function f is either *strictly* convex in z or in ξ, then the minimizer is unique.

Remark 6.2.5. This condition posed in (a) and (b) is a generalization of the conditions mentioned in the lectures. This coercivity condition follows from Theorem 3.30 of [Dac07].

Remark 6.2.6 (Context and Pathologies).

(i) Assumption (a) ensures that relaxing the problem from C^1 to $W^{1,p}$ does not strictly decrease the infimum. When this fails, it is called the **Lavrentiev Phenomenon**.

Manià's Example (1934): Consider $\mathcal{F}(u) = \displaystyle\int_0^1 (u^3 - x)^2 (u')^6 \, dx$. This Lagrangian fails the coercivity condition (3). The true minimizer is $u(x) = x^{1/3} \in W^{1,1}$, yielding $\mathcal{F}(u) = 0$. However, $x^{1/3}$ has an infinite derivative at $x = 0$, so it is not in $W^{1,\infty}$ (or C^1). Any sequence of smooth functions attempting to approximate this curve forces the energy to blow up, leading to:

$$0 = \min_{\substack{u \in W^{1,1} \\ u(0)=0 \\ u(1)=1}} \mathcal{F}(u) < \inf_{\substack{u \in C^1 \\ u(0)=0 \\ u(1)=1}} \mathcal{F}(u)$$

(ii) To prove statement (a), we only rely on the upper bound, not on convexity.

(iii) Statement (b) simply provides the growth conditions under which we can rigorously differentiate the functional to derive the E-L equations.

(iv) If the upper bound in (b) is relaxed to $|D_z f| + |D_\xi f| \le C(|\xi|^p + |z|^p + 1)$, the (E-L) equations still hold, but the test functions must be restricted to $\varphi \in C_c^1(\Omega; \mathbb{R}^\ell)$ (since the gradients of f may only be in L^1).

(v) Statement (b) does not require convexity to hold; it is a purely differential result.

(vi) In the lower bound (condition 3), it is crucial that the exponent $q < p$. For instance, consider $p = 2, q = 2$ with $C > 1$:

$$\inf_{u \in W_0^{1,2}(0,\pi)} \int_0^\pi \left((u')^2 - Cu^2\right) dx = -\infty$$

This can be seen by plugging in $u(x) = M \sin(x)$ and letting $M \to \infty$.

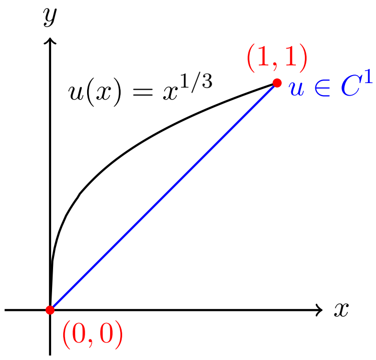


Figure 6.2: Manià's Example illustrating the Lavrentiev Phenomenon.

7 Existence of Solution to Variational Problem: Proof of Theorem 6.2.1 and Proposition 6.2.4

§7.1 Proof of Existence

The structure of the proof is divided into three steps: proving $\mathcal{F}$ is well-defined; $\mathcal{F}$ is coercive; $\mathcal{F}$ is lower semisemicontinuous. With these three steps done, we can invoke direct methods (Theorem 4.1.11) to prove Theorem 6.2.1

Proving $\mathcal{F}$ is well-defined. It is sufficient to prove $x \mapsto f(x, u(x), Du(x))$ is a measurable map.

Notice that the map $x \mapsto f(x, u(x), Du(x))$ is the composition of two maps: $x \mapsto (u(x), Du(x))$ which is measurable immediately and $(z(x), \xi(x)) \mapsto f(x, u(x), \xi(x))$ where $z(x), \xi(x)$ measurable.

To prove measurability of $(z(x), \xi(x)) \mapsto f(x, u(x), \xi(x))$, we start from the case $(z(x), \xi(x))$ is a simple function, i.e.

$$(z(x), \xi(x)) = \sum_{j=1}^{n} (z_j, \xi_j) \mathbf{1}_{E_j}.$$

We have

$$f(x, z(x), \xi(x)) = \sum_{j=1}^{n} f(x, z_j, \xi_j) \mathbf{1}_{E_j}.$$

From assumption 1., $x \mapsto f(x, z, \xi)$ is measurable. Therefore, $f(x, z(x), \xi(x))$ is measurable.

To prove the general case, let $(z(x), \xi(x))$ be measurable. There exists a sequence of simple functions $s_n(x)$ converging pointwise to $(z(x), \xi(x))$ almost everywhere. Because $(z, \xi) \mapsto f(x, z, \xi)$ is continuous for a.e. x, we have $f(x, s_n(x)) \to f(x, z(x), \xi(x))$ pointwise almost everywhere. Since pointwise limits of measurable functions are measurable, the composition is measurable. ■

Remark 7.1.1. Proving $\mathcal{F}$ is well-defined is an important matter, there is a case $f(x, z(x), \xi(x))$ is not measurable. Let $E \subseteq [0, 1]$ be a non-measurable set. We consider $\tilde{E} = \{(x, y) \in E \times E : x = y\} \subseteq [0, 1]^2$. Hence, the function $f(x, z, \xi) = 1 - \mathbf{1}_{\tilde{E}}(x, z)$ is not measurable even though it is still lower semicontinuous and thus $\mathcal{F}$ is not well-defined.

§7.1.1 Proving Coercivity of $\mathcal{F}$

To prove coercivity, we shall prove the following result

Theorem 7.1.2

Under the standing assumptions of Theorem 6.2.1, there exists $C_a = C(\Omega) > 0$ and $C_b = C(\Omega, g) \in \mathbb{R}$ such that for any $u \in g + W_0^{1,p}(\Omega; \mathbb{R}^l)$

$$\mathcal{F}(u) \geq C_a \|u\|_{W^{1,p}}^p + C_b.$$

Before starting the proof, we recall some elementary inequalities that will serve the technicalities of the proof.

Lemma 7.1.3

Let $p > 1$ and let p' be its Hölder conjugate such that $\frac{1}{p} + \frac{1}{p'} = 1$. Then the following inequalities hold:

1. **Convexity of the norm:** For all $a, b \in \mathbb{R}^n$,

$$|a + b|^p \leq 2^{p-1}(|a|^p + |b|^p).$$

   This inequality still holds for $p = 1$ since it is the triangle inequality.

2. **Weighted norm inequality:** For any $\varepsilon > 0$, there exists a constant $C_\varepsilon > 0$ such that for all $a, b \in \mathbb{R}^n$,

$$|a + b|^p \leq (1 + \varepsilon)|a|^p + C_\varepsilon |b|^p$$

   where the explicit constant is given by $C_\varepsilon = \left(1 - (1 + \varepsilon)^{-\frac{1}{p-1}}\right)^{1-p}$.

3. **Young's Inequality with ε:** For any $\varepsilon > 0$, there exists a constant $C_\varepsilon > 0$ such that for all $a, b \in \mathbb{R}$,

$$ab \leq \varepsilon |a|^p + C_\varepsilon |b|^{p'}$$

   where the explicit constant is given by $C_\varepsilon = \frac{1}{p'}(\varepsilon p)^{-\frac{p'}{p}} = \frac{p-1}{p}(\varepsilon p)^{-\frac{1}{p-1}}$.

Proof of Theorem 7.1.2. By the coercivity condition of f, we have

$$f(x, u(x), Du(x)) \geq C|Du(x)|^p - M|u(x)|^q + h(x).$$

Since $q < p$ and $h \in L^1$, we have $f(x, u(x), Du(x)) \in L^1(\Omega)$ and thus $\mathcal{F}(u) > -\infty$ for any $u \in g + W_0^{1,p}(\Omega; \mathbb{R}^l)$.

Notice that by convexity of the norm, we have

$$|Du|^p \geq 2^{1-p}|D(u - g)|^p - |Dg|^p$$

and

$$|u|^q \leq 2^{q-1}(|u - q|^q + |g|^q).$$

Therefore,

$$\mathcal{F}(u) = \int_\Omega f(x, u, Du) \geq C \int_\Omega \left(2^{1-p}|D(u - g)|^p - |Dg|^p\right) - M \int_\Omega \left(2^{q-1}(|u - g|^q + |g|^q)\right) + \int_\Omega h$$

Let

$$-C_0(g,\Omega) := \int_\Omega h - C\,|Dg|^p - M2^{q-1}|g|^q. \tag{7.1}$$

Since $u - g \in W_0^{1,p}(\Omega;\mathbb{R}^l)$, we can use Poincaré Inequality I (Theorem 6.1.11) to have

$$\int_\Omega |D(u-g)|^p \geq C_1(\Omega)\int_\Omega |u-g|^p\,. \tag{7.2}$$

On the other hand, by Hölder and Young's inequality to $p/q > 1$, we have

$$\int_\Omega |u-g|^q \leq C_2(\Omega)\cdot\left(\int_\Omega |u-g|^p\right)^{q/p} \leq C_2(\Omega)\left[\epsilon\int_\Omega |u-g|^p + C_\epsilon \cdot 1^{(q/p)'}\right]. \tag{7.3}$$

Considering the results of (7.1), (7.2), (7.3), we have the inequality (and noting $|D(u-g)|^p = \frac{1}{2}|D(u-g)|^p + \frac{1}{2}|D(u-g)|^p$):

$$\begin{aligned}\mathcal{F}(u) &\geq C2^{-p}\cdot\left(\int_\Omega |D(u-g)|^p + C_1|u-g|^p\right) - M2^{q-1}C_2\left[\epsilon\int_\Omega |u-g|^p + C_\epsilon\right] - C_0\\ &= C2^{-p}\cdot\left(\int_\Omega |D(u-g)|^p\right) + (C2^{-p}C_1 - (M2^{q-1}C_2)\epsilon)\cdot\left(\int_\Omega |u-g|^p\right) - M2^{q-1}C_2C_\epsilon C_0\end{aligned}$$

We choose $\epsilon > 0$ such that $C2^{-p}C_1 - (M2^{q-1}C_2)\epsilon \geq C2^{-p}C_1/2$ to get

$$\mathcal{F}(u) \geq \min\left\{C2^{-p}, C2^{-p}C_1/2\right\}\|u-g\|^p_{W^{1,p}} - M2^{q-1}C_2C_\epsilon C_0.$$

By convexity of the norm,

$$\|u-g\|^p_{W^{1,p}} \geq 2^{1-p}\|u\|^p_{W^{1,p}} - \|g\|^p_{W^{1,p}}.$$

Substituting this back yields

$$\mathcal{F}(u) \geq C_a\|u\|^p_{W^{1,p}} + C_b,$$

concluding the proof. ∎

With Theorem 7.1.2, proving coercivity is easy. For any $t \in \mathbb{R}$ and any $u \in \{\mathcal{F} \leq t\}$, we have

$$\|u\|^p_{W^{1,p}} \leq \frac{t - C_b}{C_a}.$$

Hence, the set $\{\mathcal{F} \leq t\}$ is bounded and it is precompact (in the weak topology of $W^{1,p}(\Omega;\mathbb{R}^\ell)$)by Rellich-Kondrachov (Theorem 6.1.9).

§7.1.2 Proving Lower Semicontinuity of $\mathcal{F}$

It is sufficient to prove $\mathcal{F}$ is convex and strongly l.s.c., i.e. l.s.c. with respect to the norm topology of $W^{1,p}(\Omega;\mathbb{R}^\ell)$ and thus we have lower semicontinuity in the weak topology from Theorem 4.2.3.

Indeed, by assumption 2., since $(z,\xi) \mapsto f(x,z,\xi)$ is convex for a.e. $x \in \Omega$, the convexity of $\mathcal{F}$ is trivial.

To prove strong l.s.c., let $u \in g + W_0^{1,p}(\Omega; \mathbb{R}^\ell)$ and $\{u_k\} \subseteq g + W_0^{1,p}(\Omega; \mathbb{R}^\ell)$ such that $u_k \to u$. Let $L = \liminf_{k\to\infty} \mathcal{F}(u_k)$. By the definition of the limit inferior, there exists a subsequence $\{u_{k_j}\}$ such that:

$$\lim_{j\to\infty} \mathcal{F}(u_{k_j}) = L$$

Because the subsequence $\{u_{k_j}\}$ still converges strongly to u in $W^{1,p}$, we can invoke the L^p convergence theorem to extract a further subsequence $\{u_{k_{j_m}}\}$ such that $u_{k_{j_m}} \to u$ and $Du_{k_{j_m}} \to Du$ pointwise almost everywhere, and there exists an L^1 dominating function $H(x)$ such that $|u_{k_{j_m}}(x)| \le H(x)$ a.e.

By the coercivity condition of f applied to this sub-subsequence, we have

$$f(x, u_{k_{j_m}}, Du_{k_{j_m}}) \ge -M|u_{k_{j_m}}(x)|^q + h(x) \ge -M|H(x)|^q + h(x).$$

Set $\Phi(x) = -M|H(x)|^q + h(x) \in L^1(\Omega)$. Since $f(x, u_{k_{j_m}}, Du_{k_{j_m}}) \ge \Phi(x)$, we can invoke Fatou's Lemma to obtain:

$$\mathcal{F}(u) = \int_\Omega \lim_{m\to\infty} f(x, u_{k_{j_m}}, Du_{k_{j_m}})\, dx \le \liminf_{m\to\infty} \int_\Omega f(x, u_{k_{j_m}}, Du_{k_{j_m}})\, dx = \liminf_{m\to\infty} \mathcal{F}(u_{k_{j_m}}).$$

However, because the parent sequence $\mathcal{F}(u_{k_j})$ converges to L, any subsequence of it must also converge to L. Therefore,

$$\liminf_{m\to\infty} \mathcal{F}(u_{k_{j_m}}) = \lim_{j\to\infty} \mathcal{F}(u_{k_j}) = L = \liminf_{k\to\infty} \mathcal{F}(u_k).$$

Chaining these inequalities together yields $\mathcal{F}(u) \le \liminf_{k\to\infty} \mathcal{F}(u_k)$, which concludes the proof.

§7.2 Proof of Proposition 6.2.4

§7.2.1 Continuity of $\mathcal{F}$

Proof of Theorem 6.2.1 (a). We assume the upper growth condition $|f(x, z, \xi)| \le C(|\xi|^p + |z|^p) + \alpha(x)$ holds with $\alpha \in L^1(\Omega)$. With this assumption, we can prove $\mathcal{F}$ is continuous in the strong topology of $W^{1,p}(\Omega; \mathbb{R}^\ell)$.

Let $u \in g + W_0^{1,p}(\Omega; \mathbb{R}^\ell)$ and $\{u_k\} \subseteq g + W_0^{1,p}(\Omega; \mathbb{R}^\ell)$ such that $u_k \to u$ strongly in $W^{1,p}$. By passing to a subsequence (without relabeling), we can assume $u_k \to u$ and $Du_k \to Du$ pointwise almost everywhere. By the continuity of f in its second and third arguments, we have:

$$f(x, u_k(x), Du_k(x)) \to f(x, u(x), Du(x)) \quad \text{a.e. } x \in \Omega$$

By the upper growth condition, the sequence is bounded pointwise by:

$$|f(x, u_k, Du_k)| \le C(|u_k(x)|^p + |Du_k(x)|^p) + \alpha(x) := H_k(x)$$

Because $u_k \to u$ strongly in $W^{1,p}$, the dominating sequence $H_k(x)$ converges strongly in $L^1(\Omega)$ to the limit function $H(x) := C(|u(x)|^p + |Du(x)|^p) + \alpha(x)$.

Since $f_k \to f$ a.e. and is dominated by an L^1-convergent sequence $H_k \to H$, we can invoke the Generalized Lebesgue Dominated Convergence Theorem (Pratt's Lemma) to conclude:

$$\mathcal{F}(u_k) = \int_\Omega f(x, u_k, Du_k)\, dx \to \int_\Omega f(x, u, Du)\, dx = \mathcal{F}(u)$$

(Since every subsequence contains a sub-subsequence converging to $\mathcal{F}(u)$, the entire original sequence converges). Thus, $\mathcal{F}$ is strongly continuous.

Now, we address the absence of the Lavrentiev phenomenon. Let $\bar{u} \in g + W_0^{1,p}(\Omega; \mathbb{R}^\ell)$ be the global minimizer. We can write $\bar{u} = g + w$, where $w \in W_0^{1,p}$. By the definition of the zero-trace Sobolev space, there exists a sequence $\{\psi_k\} \subset C_c^\infty(\Omega; \mathbb{R}^\ell)$ such that $\psi_k \to w$ strongly in $W^{1,p}$.

Define $v_k = g + \psi_k$. Clearly, $v_k \to \bar{u}$ strongly in $W^{1,p}$. Assuming the boundary data g is smooth enough to admit a $C^1(\bar{\Omega})$ extension, the sequence v_k belongs to $C^1(\bar{\Omega}; \mathbb{R}^\ell)$ and strictly satisfies $v_k = g$ on $\partial\Omega$.

By the strong continuity of $\mathcal{F}$ just established, $\lim_{k\to\infty} \mathcal{F}(v_k) = \mathcal{F}(\bar{u})$. Therefore,

$$\inf_{\substack{v \in C^1(\bar{\Omega}; \mathbb{R}^\ell) \\ v = g \text{ on } \partial\Omega}} \mathcal{F}(v) \leq \lim_{k\to\infty} \mathcal{F}(v_k) = \mathcal{F}(\bar{u}) = \min_{\substack{u \in W^{1,p}(\Omega; \mathbb{R}^\ell) \\ u = g \text{ on } \partial\Omega}} \mathcal{F}(u)$$

Since $C^1(\bar{\Omega}) \subset W^{1,p}(\Omega)$, the reverse inequality holds trivially. We conclude that the smooth infimum equals the Sobolev minimum. ■

§7.2.2 Derivation of the Euler-Lagrange Equations

First, we start by noting that the our assumption also implies the same growth condition as part a. through the following result.

Theorem 7.2.1

Suppose f satisfies the growth condition:

$$|D_z f(x, z, \xi)| + |D_\xi f(x, z, \xi)| \leq \beta(x) + C_1(|\xi|^{p-1} + |z|^{p-1})$$

with $\beta \in L^{p'}(\Omega)$. If $f(x, 0, 0) \in L^1(\Omega)$, then there exists $\alpha \in L^1(\Omega)$ and $\tilde{C} \geq 0$ such that:

$$|f(x, z, \xi)| \leq \alpha(x) + \tilde{C}(|\xi|^p + |z|^p)$$

Proof. By the Fundamental Theorem of Calculus along the segment from $(0, 0)$ to (z, ξ):

$$|f(x, z, \xi)| \leq |f(x, 0, 0)| + \int_0^1 \Big(|D_z f(x, tz, t\xi)||z| + |D_\xi f(x, tz, t\xi)||\xi|\Big)\, dt$$

Factoring $(|z| + |\xi|)$ and applying the assumed growth condition at the scaled points $(tz, t\xi)$:

$$|f(x, z, \xi)| \leq |f(x, 0, 0)| + \int_0^1 \Big[\beta(x) + C_1 t^{p-1}(|z|^{p-1} + |\xi|^{p-1})\Big](|z| + |\xi|)\, dt$$

Evaluating the integral with respect to t yields:

$$|f(x, z, \xi)| \leq |f(x, 0, 0)| + \beta(x)(|z| + |\xi|) + \frac{C_1}{p}(|z|^{p-1} + |\xi|^{p-1})(|z| + |\xi|)$$

Apply Young's Inequality ($ab \leq \frac{a^{p'}}{p'} + \frac{b^p}{p}$) to the $\beta(x)$ terms:

$$\beta(x)(|z| + |\xi|) \leq \frac{2}{p'}\beta(x)^{p'} + \frac{1}{p}(|z|^p + |\xi|^p)$$

Distribute the algebraic term and apply Young's Inequality (with exponents $\frac{p}{p-1}$ and p) to the cross terms:

$$\begin{aligned}(|z|^{p-1}+|\xi|^{p-1})(|z|+|\xi|) &= |z|^p+|\xi|^p+|z|^{p-1}|\xi|+|\xi|^{p-1}|z| \\ &\le |z|^p+|\xi|^p+\left(\frac{p-1}{p}|z|^p+\frac{1}{p}|\xi|^p\right)+\left(\frac{p-1}{p}|\xi|^p+\frac{1}{p}|z|^p\right) \\ &= 2(|z|^p+|\xi|^p)\end{aligned}$$

Substitute these bounds back into the main inequality:

$$|f(x,z,\xi)| \le \underbrace{\left(|f(x,0,0)|+\frac{2}{p'}\beta(x)^{p'}\right)}_{:=\alpha(x)}+\underbrace{\left(\frac{1+2C_1}{p}\right)}_{:=\tilde{C}}(|z|^p+|\xi|^p)$$

Since $f(x,0,0)\in L^1(\Omega)$ by assumption, and $\beta\in L^{p'}(\Omega)$ implies $\beta^{p'}\in L^1(\Omega)$, it follows that $\alpha\in L^1(\Omega)$. ■

Proof of Theorem 6.2.1 (b). Let $\bar{u}\in g+W_0^{1,p}(\Omega;\mathbb{R}^\ell)$ be a minimizer of $\mathcal{F}$, and let $\varphi\in C_c^\infty(\Omega;\mathbb{R}^\ell)$ be a smooth test function with compact support.

Under the growth condition on f, we know $\mathcal{F}(v)$ is finite for any $v\in W^{1,p}$. Thus, the function $\varepsilon\mapsto\mathcal{F}(\bar{u}+\varepsilon\varphi)$ is well-defined and finite for $\varepsilon\in\mathbb{R}$. We seek to compute the Gâteaux derivative at $\varepsilon=0$:

$$\frac{d}{d\varepsilon}\mathcal{F}(\bar{u}+\varepsilon\varphi)\Big|_{\varepsilon=0}=\lim_{\varepsilon\to0}\int_\Omega\underbrace{\frac{f(x,\bar{u}+\varepsilon\varphi,D\bar{u}+\varepsilon D\varphi)-f(x,\bar{u},D\bar{u})}{\varepsilon}}_{:=g_\varepsilon(x)}\,dx$$

Pointwise, as $\varepsilon\to0$, the difference quotient $g_\varepsilon(x)$ converges for almost every $x\in\Omega$ to the directional derivative:

$$g_\varepsilon(x)\xrightarrow{\varepsilon\to0}\langle D_z f(x,\bar{u},D\bar{u}),\varphi\rangle+\langle D_\xi f(x,\bar{u},D\bar{u}),D\varphi\rangle\quad\text{a.e. } x\in\Omega$$

To pass the limit inside the integral via the Dominated Convergence Theorem (DCT), we must find an $L^1(\Omega)$ bound for $g_\varepsilon(x)$. By the Fundamental Theorem of Calculus, for any differentiable function h, we have $\frac{1}{\varepsilon}(h(\varepsilon)-h(0))=\int_0^1 h'(\varepsilon s)\,ds$. Applying this to our Lagrangian yields:

$$g_\varepsilon(x)=\int_0^1\Big[\langle D_z f(x,\bar{u}+\varepsilon s\varphi,D\bar{u}+\varepsilon sD\varphi),\varphi\rangle+\langle D_\xi f(x,\bar{u}+\varepsilon s\varphi,D\bar{u}+\varepsilon sD\varphi),D\varphi\rangle\Big]ds$$

Taking the absolute value and applying the growth condition $|D_z f|+|D_\xi f|\le\beta(x)+C(|z|^{p-1}+|\xi|^{p-1})$, we can bound the integrand. For $|\varepsilon|\le1$, we have:

$$\begin{aligned}|g_\varepsilon(x)| &\le \int_0^1\Big(|D_z f|+|D_\xi f|\Big)(|\varphi|+|D\varphi|)\,ds \\ &\le \int_0^1\Big[\beta(x)+C\Big(|\bar{u}+\varepsilon s\varphi|^{p-1}+|D\bar{u}+\varepsilon sD\varphi|^{p-1}\Big)\Big](|\varphi|+|D\varphi|)\,ds \\ &\le \Big[\beta(x)+\tilde{C}\Big(|\bar{u}|^{p-1}+|\varphi|^{p-1}+|D\bar{u}|^{p-1}+|D\varphi|^{p-1}\Big)\Big](|\varphi|+|D\varphi|)\end{aligned}$$

where $\tilde{C}$ is a constant depending on C and p. Because $\bar{u}, \varphi \in W^{1,p}(\Omega)$, the terms with power $p-1$ belong to $L^{p/(p-1)}(\Omega) = L^{p'}(\Omega)$, and the test function terms belong to $L^p(\Omega)$. By Hölder's inequality, their product is in $L^1(\Omega)$. Furthermore, $\beta \in L^{p'}(\Omega)$, so $\beta(x)(|\varphi| + |D\varphi|) \in L^1(\Omega)$.

Thus, $|g_\varepsilon(x)|$ is uniformly bounded by an $L^1(\Omega)$ function. By the Lebesgue Dominated Convergence Theorem, we can pass the limit inside the integral:

$$0 = \frac{d}{d\varepsilon}\mathcal{F}(\bar{u}+\varepsilon\varphi)\Big|_{\varepsilon=0} = \int_\Omega \Big(\langle D_z f(x,\bar{u},D\bar{u}),\varphi\rangle + \langle D_\xi f(x,\bar{u},D\bar{u}),D\varphi\rangle \Big)\,dx \quad \forall\varphi \in C_c^\infty(\Omega;\mathbb{R}^\ell)$$

Finally, because the linear mapping $\varphi \mapsto \int_\Omega (\langle D_z f,\varphi\rangle + \langle D_\xi f, D\varphi\rangle)\,dx$ is continuous with respect to the $W^{1,p}$ norm (due to the exact same Hölder bounds used above), and $C_c^\infty(\Omega)$ is dense in $W_0^{1,p}(\Omega)$, the equality extends to all test functions. Therefore:

$$\int_\Omega \Big(\langle D_z f(x,\bar{u},D\bar{u}),\varphi\rangle + \langle D_\xi f(x,\bar{u},D\bar{u}),D\varphi\rangle \Big)\,dx = 0 \quad \forall\varphi \in W_0^{1,p}(\Omega;\mathbb{R}^\ell).$$

■

§7.2.3 Sufficiency of the Euler-Lagrange Equations

Proof of Theorem 6.2.1 (c). We wish to show that if f is convex in (z,ξ) and $\bar{u}$ satisfies the Euler-Lagrange equations, then $\bar{u}$ is a minimizer.

Recall the first-order condition for convexity: if a function $F : V \to \mathbb{R}$ on a vector space is differentiable and convex, it satisfies $F(a) \geq F(b) + \langle \nabla F(b), a-b\rangle$ for all $a, b \in V$.

Applying this pointwise to the integrand $f(x,\cdot,\cdot)$ for almost every $x \in \Omega$, and taking any admissible competitor $v \in g + W_0^{1,p}(\Omega;\mathbb{R}^\ell)$, we have:

$$\begin{aligned} f(x,v(x),Dv(x)) \geq\ & f(x,\bar{u}(x),D\bar{u}(x)) \\ & + \langle D_z f(x,\bar{u}(x),D\bar{u}(x)), (v(x)-\bar{u}(x))\rangle \\ & + \langle D_\xi f(x,\bar{u}(x),D\bar{u}(x)), (Dv(x)-D\bar{u}(x))\rangle \end{aligned}$$

Integrating this inequality over the domain Ω yields:

$$\begin{aligned} \int_\Omega f(x,v,Dv)\,dx \geq\ & \int_\Omega f(x,\bar{u},D\bar{u})\,dx \\ & + \int_\Omega \Big(\langle D_z f(x,\bar{u},D\bar{u}), (v-\bar{u})\rangle + \langle D_\xi f(x,\bar{u},D\bar{u}), D(v-\bar{u})\rangle \Big)\,dx \end{aligned}$$

Notice that because both v and $\bar{u}$ belong to the affine space $g + W_0^{1,p}(\Omega;\mathbb{R}^\ell)$, their difference $\varphi := v - \bar{u}$ belongs to the zero-trace space $W_0^{1,p}(\Omega;\mathbb{R}^\ell)$.

Since $\bar{u}$ is assumed to be a solution to the weak Euler-Lagrange equations, the integral of the derivative terms evaluated at the test function $(v-\bar{u})$ is exactly zero:

$$\int_\Omega \Big(\langle D_z f(x,\bar{u},D\bar{u}), (v-\bar{u})\rangle + \langle D_\xi f(x,\bar{u},D\bar{u}), D(v-\bar{u})\rangle \Big)\,dx = 0$$

Consequently, the inequality simplifies to:

$$\mathcal{F}(v) \geq \mathcal{F}(\bar{u}) \quad \forall v \in g + W_0^{1,p}(\Omega;\mathbb{R}^\ell)$$

Thus, $\bar{u}$ is a global minimizer of the functional. ■

§7.2.4 Uniqueness Under Strict Convexity

Proof of Theorem 6.2.1 (d). Assume for contradiction that there exist two distinct minimizers $u \neq v$ in $g + W_0^{1,p}(\Omega; \mathbb{R}^\ell)$. Let $m = \mathcal{F}(u) = \mathcal{F}(v) = \min \mathcal{F}$.

Since the admissible space is affine, the arithmetic mean $w = \frac{1}{2}(u+v)$ is also admissible. By the convexity of $\mathcal{F}$:

$$m \leq \mathcal{F}(w) \leq \frac{1}{2}\mathcal{F}(u) + \frac{1}{2}\mathcal{F}(v) = m$$

This forces $\mathcal{F}(w) = m$. Rewriting this as a single integral yields:

$$\int_\Omega \left[\frac{1}{2} f(x, u, Du) + \frac{1}{2} f(x, v, Dv) - f(x, w, Dw) \right] dx = 0$$

Because f is convex, the integrand is non-negative, so it must be zero almost everywhere:

$$f(x, w, Dw) = \frac{1}{2} f(x, u, Du) + \frac{1}{2} f(x, v, Dv) \quad \text{a.e. in } \Omega$$

We now obtain a contradiction depending on the strict convexity assumption:

- **If f is strictly convex in z:** Since $u \neq v$ in L^p, there exists a set $E \subset \Omega$ of positive measure where $u(x) \neq v(x)$. On E, strict convexity enforces $f(x, w, Dw) < \frac{1}{2} f(x, u, Du) + \frac{1}{2} f(x, v, Dv)$, contradicting the a.e. equality.
- **If f is strictly convex in ξ:** Since $(u - v) \in W_0^{1,p}(\Omega; \mathbb{R}^\ell)$ and $u \neq v$, Poincaré's inequality ensures $\|Du - Dv\|_{L^p} > 0$. Thus, $Du \neq Dv$ on a set of positive measure, again triggering a strict pointwise inequality that contradicts the a.e. equality.

Therefore, $u = v$ almost everywhere, and the minimizer is unique. ■

8 Example of Applications of Results from Theorem 6.2.1 and Section 6.2.4

The Direct Method and the Existence Theorem developed previously serve as the foundation for solving a vast array of variational problems. Below is a catalog of classical applications and boundary value problems derived from these principles.

1. The Dirichlet Energy (Laplace Equation)

The simplest application is the minimization of the Dirichlet energy:

$$\min_{u \in g + W_0^{1,2}(\Omega)} \int_\Omega \frac{1}{2}|Du|^2 \, dx$$

The functional is coercive and strictly convex. The unique minimizer $\bar{u}$ exists and weakly satisfies the Laplace equation:

$$\begin{cases} -\Delta \bar{u} = 0 & \text{in } \Omega \\ \bar{u} = g & \text{on } \partial\Omega \end{cases}$$

2. The Poisson Equation and Generalizations

For a given source term $f \in L^2(\Omega)$, consider the problem:

$$\min_{u \in g + W_0^{1,2}(\Omega)} \int_\Omega \left(\frac{1}{2}|Du|^2 - fu \right) dx$$

The associated Euler-Lagrange (E-L) equation is the Poisson equation $-\Delta u = f$. Similarly, if we consider a vector field $F \in L^2(\Omega; \mathbb{R}^d)$, minimizing the functional $\int_\Omega \left(\frac{1}{2}|Du|^2 - \langle F, Du \rangle \right) dx$ yields the weak formulation of $-\Delta u = \text{div} F$.

3. The Brachistochrone Problem

Consider the classical problem of finding the curve of fastest descent:

$$\min_{u(0)=0} \int_0^1 \frac{\sqrt{1 + (u')^2}}{\sqrt{-u}} \, dx$$

At first glance, this Lagrangian is highly non-convex. However, by introducing the change of variables $v = \sqrt{-u}$, we obtain $v' = \dfrac{-u'}{2\sqrt{-u}}$. The functional transforms into:

$$\int_0^1 \sqrt{\frac{1}{v^2} + 4(v')^2} \, dx$$

The new integrand $(z, \xi) \mapsto \sqrt{\dfrac{1}{z^2} + 4\xi^2}$ is strictly convex in (z, ξ). Because convexity is established, any solution to the E-L equations for this transformed problem is guaranteed to be a global minimizer (yielding arcs of cycloids) (refer to (2.2)).

4. The p-Laplacian

Generalizing the Dirichlet energy to L^p spaces gives the p-Dirichlet integral:

$$\min_{u \in g + W_0^{1,p}(\Omega)} \int_\Omega \left(\frac{1}{p} |Du|^p - fu \right) dx$$

The E-L equation is the quasilinear p-Laplace equation:

$$\begin{cases} -\text{div}(|Du|^{p-2} Du) = f & \text{in } \Omega \\ u = g & \text{on } \partial\Omega \end{cases}$$

5. Linear Elliptic Equations with Variable Coefficients

Let $A(x)$ be a measurable $d \times d$ matrix field. We minimize:

$$\min_{u \in g + W_0^{1,2}(\Omega)} \int_\Omega \left(\frac{1}{2} \langle A(x)Du, Du \rangle - fu \right) dx$$

The corresponding E-L equation is $-\text{div}(A(x)Du) = f$. For the functional to be coercive, we require $A(x)$ to satisfy a uniform ellipticity condition: $\langle A(x)\xi, \xi \rangle \geq \lambda |\xi|^2$ for some $\lambda > 0$ and almost every $x \in \Omega$.

6. Linearized Elasticity

In the theory of linearized elasticity, the energy depends on the symmetric part of the gradient, representing the deformation tensor:

$$\min_{u \in g + W_0^{1,2}(\Omega; \mathbb{R}^d)} \int_\Omega |Du + (Du)^T|^2 \, dx$$

The mapping $\xi \mapsto |\xi + \xi^T|^2$ is convex. However, a pointwise coercivity bound like $\langle A\xi, \xi \rangle \geq \lambda |\xi|^2$ is impossible here because the symmetric gradient operator has a non-trivial kernel (infinitesimal rigid body rotations). Coercivity over $W_0^{1,2}$ is instead recovered globally via **Korn's Inequality**:

$$\int_\Omega |Du + (Du)^T|^2 \, dx \geq C(\Omega) \int_\Omega |Du|^2 \, dx$$

7. The Obstacle Problem

We restrict the admissible space to functions lying above a certain obstacle (here, $u \geq 0$):

$$\min \left\{ \int_\Omega \left(\frac{1}{2} |Du|^2 - fu \right) dx : u \in g + W_0^{1,2}(\Omega), \, u \geq 0 \text{ a.e.} \right\}$$

The admissible set $\mathcal{C}$ is a closed and convex subset of $W^{1,2}(\Omega)$. Because the functional is coercive and weakly lower semi-continuous, we only need to verify that $\mathcal{C}$ is weakly closed. If a minimizing sequence $u_k \rightharpoonup \bar{u}$ weakly, Mazur's Lemma guarantees a sequence of convex combinations converging strongly (and thus pointwise a.e.), which preserves the condition $u \geq 0$. The direct method applies perfectly.

Remark 8.0.1. For this problem, $\bar{u}$ is unique and satisfies a variational inequality rather than an equation: $\int_\Omega \langle D\bar{u}, D\varphi \rangle - f\varphi \geq 0$ for all admissible test functions $\varphi \geq 0$.

8. Natural (Neumann) Boundary Conditions

Consider the minimization over the entire Sobolev space without trace restrictions:

$$\min_{u \in W^{1,2}(\Omega)} \int_\Omega \left(\frac{1}{2}|Du|^2 - fu \right) dx$$

Formally, a minimizer satisfies $-\Delta u = f$ in Ω and $\partial_\nu u = 0$ on $\partial\Omega$. A necessary compatibility condition for existence is $\displaystyle\int_\Omega f\,dx = 0$ (otherwise, sending $u \equiv c \to \pm\infty$ forces the energy to $-\infty$). Under this condition, the energy is invariant under constant shifts $u \mapsto u + c$. By restricting the minimization to the closed subspace of functions with zero mean, $\tilde{W} = \{u \in W^{1,2}(\Omega) : \displaystyle\int_\Omega u = 0\}$, the Poincaré-Wirtinger inequality (Theorem 3.1.5) restores coercivity.

9. The Rayleigh Quotient

The optimal constant for the Poincaré inequality can be found by minimizing the Rayleigh quotient:

$$\inf_{u \in W_0^{1,2}(\Omega)\setminus\{0\}} \frac{\int_\Omega |Du|^2\,dx}{\int_\Omega u^2\,dx}$$

Because the quotient is invariant under homothety ($u \mapsto \lambda u$), this problem is equivalent to minimizing the gradient energy subject to an L^2 spherical constraint:

$$\min \left\{ \int_\Omega |Du|^2\,dx : u \in W_0^{1,2}(\Omega),\ \int_\Omega u^2\,dx = 1 \right\}$$

10. Pathologies and Limitations of the Classical Assumptions

The standard assumptions ($p > 1$ and convexity in ξ) are sometimes too restrictive for physical and geometric realities:

- **Minimal Surfaces:** The functional $\displaystyle\int_\Omega \sqrt{1 + |Du|^2}\,dx$ has linear growth ($p = 1$). Bounded energy sequences belong to $W^{1,1}$, which is not reflexive. Weak limits may exhibit jump discontinuities, requiring the problem to be relaxed into the space of Bounded Variation (BV).
- **Lack of Convexity in z:** Phase transition models often utilize double-well potentials, such as the Allen-Cahn energy $\displaystyle\int_\Omega \big(|Du|^2 + W(u)\big)\,dx$ where $W(u) = (1 - u^2)^2$.
- **Lack of Convexity in ξ:** In nonlinear elasticity (e.g., Neo-Hookean materials), the energy depends on the deformation volume, yielding functionals like $\displaystyle\int_\Omega \big(|Du|^2 + g(\det Du)\big)\,dx$. The map $A \mapsto \det A$ is strictly non-convex, requiring weaker topological conditions such as *polyconvexity* to establish lower semicontinuity.

9 Beyond Convexity - 1

§9.1 Is Convexity Necessary in the Gradient Variable?

Recall that the Direct Method (Theorem 4.1.11) relies on two pillars:

1. **Coercivity** (Growth conditions and lower bounds) to guarantee a bounded minimizing sequence.
2. **Weak Lower Semicontinuity (L.S.C.)** to ensure the limit of the sequence minimizes the functional.

Previously in, we established that if $f(x, z, \xi)$ is jointly convex in (z, ξ), then $\mathcal{F}(u) = \int_\Omega f(x, u, Du)\, dx$ is convex. By abstract nonsense (Theorem 4.2.3), strong l.s.c. combined with convexity implies weak l.s.c. for $\mathcal{F}$

However, requiring convexity in the state variable z is extremely restrictive (it excludes phase transition models, double-well potentials, etc. as shown in the previous lecture). We need to relax this.

Remark 9.1.1 (The Role of Oscillations). If $u_k \rightharpoonup u$ weakly in $W^{1,p}(\Omega)$, the Rellich-Kondrachov compactness theorem ensures that $u_k \to u$ strongly in $L^p(\Omega)$ (for bounded Ω). Consequently, up to a subsequence, $u_k \to u$ pointwise almost everywhere.

Because the sequence $\{u_k\}$ converges pointwise, it does not suffer from high-frequency oscillations. The weak convergence $Du_k \rightharpoonup Du$, however, means the gradient *can* oscillate heavily. Therefore, we strictly need convexity in the gradient variable ξ to "absorb" these oscillations, but we do not need convexity in the state variable z.

Theorem 9.1.2 (Morrey-Tonelli)
Let $\Omega \subset \mathbb{R}^d$ be a domain of finite measure. Let $f : \Omega \times \mathbb{R}^\ell \times \mathbb{R}^{\ell \times d} \to \mathbb{R}$ be a Lagrangian satisfying the Carathéodory conditions:

1. $x \mapsto f(x, z, \xi)$ is Lebesgue measurable for all (z, ξ).
2. $(z, \xi) \mapsto f(x, z, \xi)$ is continuous for almost every $x \in \Omega$.

Assume further that:

(3) **Lower Bound:** f is non-negative (or bounded below by an $L^1(\Omega)$ function).

(4) **Partial Convexity:** For almost every $x \in \Omega$ and all $z \in \mathbb{R}^\ell$, the map $\xi \mapsto f(x, z, \xi)$ is convex.

Then, the integral functional $\mathcal{F}(u) = \int_\Omega f(x, u(x), Du(x))\, dx$ is well-defined and weakly lower semicontinuous with respect to the weak topology of $W^{1,p}(\Omega)$ for any $p \geq 1$.

Proof. To avoid heavy measure-theoretic technicalities, we prove this under the additional simplifying assumptions that the partial derivative $D_\xi f(x,u,\xi)$ exists and the map $(u,\xi) \mapsto D_\xi f(x,u,\xi)$ is continuous.

We also assume $f \geq 0$, if f is bounded below by $h \in L^1(\Omega)$, we consider $f - h \geq 0$.

Let $u_k \rightharpoonup u$ weakly in $W^{1,p}(\Omega)$. We want to show $\mathcal{F}(u) \leq \liminf_{k\to\infty} \mathcal{F}(u_k)$.

Step 1: Subsequence Extraction.
We pass to a subsequence (without relabeling) that achieves the limit inferior:

$$\lim_{k\to\infty} \mathcal{F}(u_k) = \liminf_{k\to\infty} \mathcal{F}(u_k)$$

Because $u_k \rightharpoonup u$ in $W^{1,p}(\Omega)$, Rellich-Kondrachov implies $u_k \to u$ strongly in $L^p(\Omega)$. Thus, we can extract a further sub-subsequence such that $u_k(x) \to u(x)$ pointwise a.e. in Ω.

Step 2: Pointwise Limits.
Because $u_k(x) \to u(x)$ a.e., the continuity of f and $D_\xi f$ with respect to the state variable implies:

$$f(x, u_k(x), Du(x)) \xrightarrow{k\to\infty} f(x, u(x), Du(x)) \quad \text{a.e.} \tag{9.1}$$

$$D_\xi f(x, u_k(x), Du(x)) \xrightarrow{k\to\infty} D_\xi f(x, u(x), Du(x)) \quad \text{a.e.} \tag{9.2}$$

Step 3: The Subgradient Inequality.
Because $\xi \mapsto f(x,z,\xi)$ is convex and differentiable, it lies above its tangent plane. For any k, setting the base point as Du and the evaluation point as Du_k, we obtain pointwise a.e.:

$$f(x, u_k, Du_k) \geq f(x, u_k, Du) + \langle D_\xi f(x, u_k, Du), Du_k - Du\rangle$$

Step 4: Passing to the Limit.
Fix an arbitrary $\varepsilon > 0$. We apply two measure-theoretic approximations to obtain a well-behaved compact domain:

1. By **Lusin's Theorem** applied to the measurable functions u and ∇u, there exists a compact set $K_1 \subset \Omega$ such that $|\Omega \setminus K_1| < \varepsilon/2$ and in K_1, the map $x \mapsto D_\xi f(x, u(x), Du(x))$ is uniformly continuous.
2. By **Egorov's Theorem**, there exists a compact set $K_2 \subset \Omega$ such that $|\Omega \setminus K_2| < \varepsilon/2$ and $u_k \to u$, and the convergences of (9.1) and (9.2) are uniform.

We define $K_\varepsilon := K_1 \cap K_2$ where $|\Omega \setminus K_\varepsilon| < \varepsilon$, and in K_ε both results of Lusin and Egorov hold simultaneously. Since we assume $f \geq 0$, we have $\int_\Omega f(x, u_k, Du_k)\,dx \geq \int_{K_\varepsilon} f(x, u_k, Du_k)\,dx$. We apply the subgradient inequality with Du as the base point:

$$\begin{aligned}
\int_\Omega f(x,u_k,Du_k)\,dx &\geq \int_{K_\varepsilon} f(x,u_k,Du)\,dx + \int_{K_\varepsilon} \langle D_\xi f(x,u_k,Du), (Du_k - Du)\rangle\,dx \\
&= \underbrace{\int_{K_\varepsilon} f(x,u_k,Du)\,dx}_{I} + \underbrace{\int_\Omega \mathbf{1}_{K_\varepsilon} \langle D_\xi f(x,u,Du), (Du_k - Du)\rangle\,dx}_{II} \\
&\quad + \underbrace{\int_{K_\varepsilon} \langle D_\xi f(x,u_k,Du) - D_\xi f(x,u,Du), (Du_k - Du)\rangle\,dx}_{III}
\end{aligned}$$

Note that we assumed $u_k \rightharpoonup u$ in $W^{1,p}(\Omega;\mathbb{R}^\ell)$. Thus, $\|u_k\|_{W^{1,p}}$ is bounded. Hence, we can safely pass by limit to each term of the integration to have:

$$I \xrightarrow[\text{uniform convergence}]{k\to\infty} \int_{K_\varepsilon} f(x,u,Du)\,dx \xrightarrow[\text{Monotone Conv. since } f\geq 0]{\varepsilon\to 0^+} \int_\Omega f(x,u,Du)\,dx = \mathcal{F}(u).$$

$II \to 0$ since $(u,\xi)\mapsto D_\xi f(x,u,\xi)$ is continuous and thus $\mathbf{1}_{K_\varepsilon} D_\xi f(x,u,\xi) \in L^\infty(\Omega) \subseteq L^{p'}(\Omega)$.

$$|III| \leq \underbrace{(\|Du_k - Du\|_{L^p})}_{\text{bounded}} \cdot \left(|K_\varepsilon|^{1/p'} \cdot \|D_\xi f(\cdot,u_k,Du) - D_\xi f(\cdot,u,Du)\|_{L^\infty(K_\varepsilon)}\right) \xrightarrow{k\to\infty} 0.$$

Therefore, we have just showed $\liminf_{k\to\infty} \mathcal{F}(u_k) \geq \mathcal{F}(u)$. ■

Now, we shall try to find the right notion of "convexity" on necessity of weak lower semicontinuity to be "convexing". Throughout this lecture, we assume

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

and we want to find the necessary condition for lower semicontinuity of it.

We introduce the following notion of convexities.

Definition 9.1.3

Let $f: \mathbb{R}^\ell \otimes \mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$. The function f is said to be:

1. polyconvex if there exists $F: \mathbb{R}^{\tau(d,\ell)} \to \mathbb{R}\cup\{+\infty\}$ convex, such that
$$f(\xi) = F(T(\xi))$$
where $T: \mathbb{R}^\ell \otimes \mathbb{R}^d \to \mathbb{R}^{\tau(\ell,d)}$ is such that
$$T(\xi) := (\xi, \operatorname{adj}_2, \xi, \dots, \operatorname{adj}_{\ell\wedge d}, \xi)$$
where $\operatorname{adj}_m \xi$ stands for the matrix of all $m\times m$ minors of the matrix $\xi \in \mathbb{R}^\ell \otimes \mathbb{R}^d$, $2 \leq m \leq \ell\wedge d = \min\{\ell, d\}$, and
$$\tau(\ell,d) := \sum_{s=1}^{\ell\wedge d} \sigma(m), \quad \text{where} \quad \sigma(m) := \binom{\ell}{m}\binom{d}{m}.$$
2. quasiconvex if the inequality
$$f(\xi) \leq \fint_D f(\xi + D\varphi)\,dx$$
holds for every bounded set $D \subset \mathbb{R}^d$, for every $\xi \in \mathbb{R}^\ell \otimes \mathbb{R}^d$, and for every $\varphi \in W_0^{1,\infty}(D;\mathbb{R}^\ell)$.
3. rank one convex if
$$f(\lambda\xi + (1-\lambda)\eta) \leq \lambda f(\xi) + (1-\lambda) f(\eta)$$
for every $\lambda \in [0,1]$, ξ, $\eta \in \mathbb{R}^\ell \otimes \mathbb{R}^d$ with $\operatorname{rank}(\xi - \nu) \leq 1$.

Indeed, in the subsequent part of this notes, we shall prove the following relation

$$f \text{ convex} \implies f \text{ polyconvex} \implies f \text{ quasiconvex} \implies f \text{ rank-one convex.}$$

The reason of our consideration is due to the following result regarding necessity of a l.s.c. functional which will be proven later in this note.

Theorem (Morrey)
Let $\Omega \subset \mathbb{R}^d$ be a bounded open set and $f : \mathbb{R}^\ell \otimes \mathbb{R}^d \to \mathbb{R}$ be a Carathéodory function. Assume that $\mathcal{F}(u) = \int_\Omega f(Du)\,dx$ is l.s.c. with respect to the weak* topology of $W^{1,\infty}(\Omega;\mathbb{R}^\ell)$, i.e.

$$\mathcal{F}(u) \leq \liminf_{k\to\infty} \mathcal{F}(u_k)$$

for any $u \in W^{1,\infty}(\Omega;\mathbb{R}^\ell)$ and any $\{u_k\} \subseteq W^{1,\infty}(\Omega;\mathbb{R}^\ell)$ such that $u_k \xrightarrow{L^\infty} u$ and $Du_k \rightharpoonup^* Du$. Then, f is quasiconvex.

§9.2 Observations on Quasiconvexity

Indeed, we define quasiconvexity with arbitrary bounded set $D \subset \mathbb{R}^d$. However, it is sufficient to proof it for one set, such as the unit cube $Q = [0,1)^d$.

Proposition 9.2.1
Let $f : \mathbb{R}^\ell \otimes \mathbb{R}^d \to \mathbb{R}$ be Borel measurable and locally bounded. Let $D \subset \mathbb{R}^n$ be a bounded open set and let the inequality

$$f(\xi)\cdot|D| \leq \int_D f(\xi + D\varphi(x))\,dx \tag{9.3}$$

holds for every $\xi \in \mathbb{R}^\ell \otimes \mathbb{R}^d$ and for every $\varphi \in W_0^{1,\infty}(D;\mathbb{R}^\ell)$. Then the inequality

$$f(\xi)\cdot|E| \leq \int_E f(\xi + D\varphi(x))\,dx \tag{9.4}$$

holds for every bounded set $E \subset \mathbb{R}^d$, for every $\xi \in \mathbb{R}^\ell \otimes \mathbb{R}^d$, and for every $\varphi \in W_0^{1,\infty}(E;\mathbb{R}^\ell)$.

Proof. Let $E \subset \mathbb{R}^d$ and $\psi \in W_0^{1,\infty}(E;\mathbb{R}^l)$. Cover E by a sufficiently large open cube $Q_a := (-a,a)^d$. Extend ψ to V as

$$V(x) := \begin{cases} \psi(x) & \text{if } x \in E \\ 0 & \text{if } x \in Q_a \setminus E \end{cases}$$

and now $V \in W_0^{1,\infty}(Q_a;\mathbb{R}^l)$.

We shall translate and rescale Q_a such that it is contained in D. Let $x_0 \in D$. Since D is open, we can shrink Q_a by $\frac{1}{\mu}$ for μ large enough such that

$$x_0 + \frac{1}{\mu}Q_a \subset D.$$

We can define

$$\varphi(x) := \begin{cases} \frac{1}{\mu} V(\mu(x - x_0)) & \text{if } x \in x_0 + \frac{1}{\nu} Q_a \\ 0 & \text{otherwise} \end{cases}$$

where

$$D\varphi(x) = DV(\mu(x - x_0)).$$

Therefore, $\varphi \in W_0^{1,\infty}(D; \mathbb{R}^l)$ and so

$$\begin{aligned} \int_D f(\xi + D\varphi(x))\, dx &= \int_{D \setminus \left(x_0 + \frac{1}{\mu} Q_a\right)} f(\xi)\, dx + \int_{x_0 + \frac{1}{\mu} Q_a} f(\xi + D\varphi(x))\, dx \\ &= f(\xi) \left[|D| - \frac{1}{\mu^d} |Q_a| \right] + \frac{1}{\mu^d} \int_{Q_a} f(\xi + DV(x))\, dx \\ &= f(\xi) \left[|D| - \frac{1}{\mu^d} |Q_a| \right] + \frac{f(\xi)}{\mu^d} |Q_a \setminus E| + \frac{1}{\mu^d} \int_E D\psi(x)\, dx \\ &= f(\xi) \left[|D| - \frac{1}{\mu^d} |E| \right] + \frac{1}{\mu^d} \int_E D\psi(x)\, dx. \end{aligned}$$

We can then use (9.3) and simplifying the terms yields (9.4) as desired. ■

Remark 9.2.2. With this proposition, we can show the implication

$$f \text{ convex} \implies f \text{ quasiconvex}$$

immediately. Considering the unit cube Q that has regular boundary regular, we can use Jensen's inequality by convexity of f and the Divergence theorem afterwards to have

$$\fint_Q f(A + D\varphi)\, dx \geq f\left(\fint_Q A + D\varphi\, dx \right) = f\left(A + \fint_{\partial Q} \varphi \otimes \nu\, d\mathcal{H}^{d-1}(x) \right) = f(A).$$

If we have proved

$$f \text{ quasiconvex} \implies f \text{ rank-one convex}$$

if $d = 1$, or $\ell = 1$, we have the equivalence between convexity and rank-one convexity. Therefore, we also have equivalence between convexity and quasiconvexity (in addition, rank-one convexity as well).

Unfortunately, if ℓ, $d \geq 2$, there are quasiconvex functions which are not convex which will be shown after the following preliminary proposition.

Definition 9.2.3

Given a matrix $A \in \mathbb{R}^\ell \otimes \mathbb{R}^d$ and m such that $1 \leq m \leq \tau(\ell, d)$, we denote $\mathcal{M}_m(A)$ as the vector of all $m \times m$ minors where $\mathcal{M}_m(A) \in \mathbb{R}^{\binom{\ell}{m}\binom{d}{m}}$. We also denote

$$\mathcal{M}(A) = (\mathcal{M}_1(A), \ldots, \mathcal{M}_{\ell \wedge d}(A)) \in \mathbb{R}^{\tau(\ell, d)}.$$

Remark 9.2.4. We still put this definition even after defining the function $T : \mathbb{R}^\ell \oplus \mathbb{R}^d \to \mathbb{R}^{\tau(\ell,d)}$ to stay faithful to the lectures.

Proposition 9.2.5

Let $\Omega \subset \mathbb{R}^d$ be an open set with C^1 boundary and let $\psi_1,\, \psi_2 \in C^1(\overline{\Omega}; \mathbb{R}^\ell)$ such that ψ_1 and ψ_2 agrees in $\partial\Omega$. Then, for any $m \times m$ minor $\mathcal{M}_m$ where $1 \leq m \leq \tau(\ell, d)$, we have

$$\int_\Omega \mathcal{M}_m(D\psi_1)\, dx = \int_\Omega \mathcal{M}_m(D\psi_2)\, dx$$

Proof. Without loss of generality, it suffices to consider the principal case where the dimensions of the domain and codomain are equal to m (i.e., $\Omega \subset \mathbb{R}^m$ and $\ell = d = m$), and the minor $\mathcal{M}_m(A)$ is the determinant $\det(A)$. The general theorem follows by applying this exact argument to arbitrary $m \times m$ submatrices of the full Jacobian.

The key to establishing the equality of these integrals is recognizing that the determinant of a Jacobian is a null Lagrangian. In the language of differential geometry, it can be expressed as an exact form.

1. The Determinant as a Pullback

Let $\omega_0 = dy^1 \wedge \cdots \wedge dy^m$ be the standard volume form on the target space $\mathbb{R}^m$. The pullback of ω_0 by a sufficiently smooth function ψ is given by:

$$\psi^*\omega_0 = d\psi^1 \wedge \cdots \wedge d\psi^m$$

By expanding each differential as $d\psi^i = \sum_{k=1}^{m} \frac{\partial \psi^i}{\partial x^k} dx^k$ and applying the alternating property of the wedge product ($dx^i \wedge dx^j = -dx^j \wedge dx^i$), this pullback naturally isolates the determinant of the Jacobian matrix:

$$\psi^*\omega_0 = \det(D\psi)\, dx^1 \wedge \cdots \wedge dx^m$$

2. Exactness and the Piola Identity

Assume temporarily that $\psi \in C^2(\overline{\Omega}; \mathbb{R}^m)$. The volume form ω_0 is exact on $\mathbb{R}^m$. Because the exterior derivative commutes with pullbacks, $\psi^*\omega_0$ must also be exact. We can explicitly construct the $(m-1)$-form α_ψ such that $d\alpha_\psi = \psi^*\omega_0$:

$$\begin{aligned}
\det(D\psi)\, dx^1 \wedge \cdots \wedge dx^m &= d\psi^1 \wedge \cdots \wedge d\psi^m \\
&= d\big(\psi^1 d\psi^2 \wedge \cdots \wedge d\psi^m\big) \\
&= d\big((-1)^{j+1}\psi^j d\psi^1 \wedge \cdots \wedge d\psi^{j-1} \wedge d\psi^{j+1} \wedge \cdots \wedge d\psi^m\big)
\end{aligned}$$

By averaging these symmetrized forms over all m indices, we obtain the differential representation of the Piola identity:

$$\det(D\psi)\, dx^1 \wedge \cdots \wedge dx^m = d\left(\frac{1}{m}\sum_{j=1}^{m}(-1)^{j+1}\psi^j d\psi^1 \wedge \cdots \wedge \widehat{d\psi^j} \wedge \cdots \wedge d\psi^m\right) = d\alpha_\psi$$

(In classical vector calculus, the exactness $d^2 = 0$ corresponds to the row-wise divergence-free property of the cofactor matrix, $\operatorname{div}(\operatorname{cof}(D\psi)) = 0$, leading to $\det(D\psi) = \operatorname{div}(\operatorname{cof}(D\psi)\psi)$).

3. Applying Stokes' Theorem for Boundary Agreement

Let $\psi_1, \psi_2 \in C^2(\overline{\Omega}; \mathbb{R}^m)$ be two functions that agree on the boundary, such that $\psi_1|_{\partial\Omega} =$

$\psi_2|_{\partial\Omega}$. Because they coincide everywhere on $\partial\Omega$, their tangential derivatives along the boundary manifold are perfectly identical. Consequently, the pullbacks of their $(m-1)$-forms evaluated strictly on the boundary are the same: $\alpha_{\psi_1}|_{\partial\Omega} = \alpha_{\psi_2}|_{\partial\Omega}$.

By the generalized Stokes' Theorem for manifolds with boundary:

$$\int_\Omega \det(D\psi_1)\,dx = \int_\Omega d\alpha_{\psi_1} = \int_{\partial\Omega} \alpha_{\psi_1}$$

Because the boundary forms match perfectly, we deduce:

$$\int_{\partial\Omega} \alpha_{\psi_1} = \int_{\partial\Omega} \alpha_{\psi_2} = \int_\Omega d\alpha_{\psi_2} = \int_\Omega \det(D\psi_2)\,dx$$

4. Extension to C^1 via Density

The proposition assumes $\psi_1, \psi_2 \in C^1$, but the construction of the exact form and the nilpotency ($d^2 = 0$) require second weak derivatives. However, $C^2(\overline{\Omega})$ is dense in $C^1(\overline{\Omega})$. We can approximate ψ_1 and ψ_2 by sequences of C^2 functions that maintain the boundary agreement. Because the integral of the determinant is continuous with respect to the C^1 topology, the equality extends to $C^1(\overline{\Omega};\mathbb{R}^m)$ via a standard density argument, completing the proof. ■

Proposition 9.2.6

Let $g : \mathbb{R} \to \mathbb{R}$ is convex, then the map $G : \mathbb{R}^2 \otimes \mathbb{R}^2 \to \mathbb{R}$ given by $G(\xi) = g(\det \xi)$ is quasiconvex, but in general not convex.

Proof. If $g(1) > g(0)$, one can take $A = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$ where $G(A) = G(B) = 0$, but $G\left(\frac{1}{2}A + \frac{1}{2}B\right) = G(\mathrm{Id}) = g(1)$. Hence, G is not convex.

To prove quasiconvexity, we consider the unit cube Q. Let $\xi \in \mathbb{R}^2 \otimes \mathbb{R}^2$, for any $\varphi \in W_0^{1,\infty}(Q;\mathbb{R}^2)$, we use Proposition 9.2.5 (after passing by density of $C_C^\infty \subset C^1 \cap W_0^{1,\infty}$) to have

$$\fint_Q \det(A + D\varphi)\,dx = \fint \det A\,dx = \det\left(\fint A\,dx\right) = \det\left(\fint A + D\varphi\,dx\right).$$

Therefore, we can use Jensen's inequality to conclude

$$\fint g(\det(A + D\varphi))\,dx \geq g\left(\fint \det(A + D\varphi)\,dx\right) = g(\det A).$$

■

We can actually generalize this argument to $\mathcal{A}$ (or $T(A)$) instead to have the following result.

Corollary 9.2.7

If $f : \mathbb{R}^\ell \otimes \mathbb{R}^d \to \mathbb{R}$ is polyconvex, f is quasiconvex.

10 Beyond Convexity - 2

The notes on this lecture is shorter than the others because half of the lectures was reviewing contents that I decided to put on Lecture 9.

§10.1 Observations on Rank-One Convexity and Quasiconvexity

Recalling the definition of rank-one convexity, we can also write rank-one convexity as the condition of the mapping

$$t \mapsto g(A + t(p \otimes q))$$

is convex for any $A \in \mathbb{R}^\ell \otimes \mathbb{R}^d$, $p \in \mathbb{R}^\ell$, and $q \in \mathbb{R}^d$.

Remark 10.1.1. We summarized our current result (and goal) as follows

$$\begin{array}{ccccccc} f \text{ convex} & \Longrightarrow & f \text{ polyconvex} & \Longrightarrow & f \text{ quasiconvex} & \Longrightarrow & f \text{ rank-one convex} \\ & \not\Leftarrow & & \not\Leftarrow & & \not\Leftarrow & (\text{for } l \geq 3 \text{ by Šverák}) \end{array}$$

Following Remark 9.2.2, if $l = 1$, we have all notions being equivalent. However, the case $l = 2$ it is still an open problem (Morrey conjecture).

We only need to proof quasiconvexity implies rank-one convexity to settle our chain of implications.

Proposition 10.1.2

Let $f : \mathbb{R}^\ell \otimes \mathbb{R}^d \to \mathbb{R}$. If f is quasiconvex, then f is rank-one convex.

We will have two proofs for this proposition.

Proof 1 (requires regularity on f). We assume $f \in C^2$. Note that for any bounded open set $\Omega \subset \mathbb{R}^d$ and $\varphi \in C_c^\infty(\Omega; \mathbb{R}^\ell)$, the notion of convexity implies the function

$$h(\epsilon) := \int_\Omega f(A + \epsilon D\varphi)$$

attains its minimum at $\epsilon = 0$. Therefore, we have

$$h''(0) = \int_\Omega \frac{\partial^2 f}{\partial \xi_i^\alpha \partial \xi_j^\beta}(A) \partial_i \varphi^\alpha \partial_j \varphi^\beta \, dx \geq 0. \tag{10.1}$$

(we are using Einstein summation notion)

Let $a \in \mathbb{R}^d$ and $b \in \mathbb{R}^\ell$, we take $\varphi \in C_c^\infty(\Omega)$ and define $\varphi_1(x) := \psi(x) b \cos(\tau a \cdot x)$, $\varphi_2(x) := \psi(x) b \sin(\tau a \cdot x)$ for $\tau \in \mathbb{R}$ where $\varphi_1, \varphi_2 \in C_c^\infty(\Omega; \mathbb{R}^\ell)$. We plug φ_1 and φ_2 to (10.1) to have (after simplifying terms)

$$\int_\Omega \frac{\partial^2 f}{\partial \xi_i^\alpha \partial \xi_j^\beta}(A) \left[\partial_i \psi(x) \partial_j \psi(x) b^\alpha b^\beta \cos^2(\tau a \cdot x) + \psi^2 \tau^2 a_i a_j b^\alpha b^\beta \sin^2(\tau a \cdot x) \right] dx \geq 0$$

$$\int_\Omega \frac{\partial^2 f}{\partial \xi_i^\alpha \partial \xi_j^\beta}(A) \left[\partial_i \psi(x) \partial_j \psi(x) b^\alpha b^\beta \sin^2(\tau a \cdot x) + \psi^2 \tau^2 a_i a_j b^\alpha b^\beta \cos^2(\tau a \cdot x) \right] dx \geq 0$$

summing the two terms, we have

$$\int_\Omega \frac{\partial^2 f}{\partial \xi_i^\alpha \partial \xi_j^\beta}(A)\left[\partial_i\psi(x)\partial_j\psi(x) + \psi^2\tau^2 a_i a_j\right] b^\alpha b^\beta \, dx \geq 0$$

We then divide by τ^2 on both side and pas $\tau \to +\infty$ to conclude

$$\int_\Omega \left[\frac{\partial^2 f}{\partial \xi_i^\alpha \partial \xi_j^\beta}(A) a_i a_j b^\alpha b^\beta\right] \psi^2 \, dx \geq 0$$

Since $\psi \in C_c^\infty(\Omega)$ is taken arbitrarily, we have $\frac{\partial^2 f}{\partial \xi_i^\alpha \partial \xi_j^\beta}(A) a_i a_j b^\alpha b^\beta \geq 0$. Therefore, the map

$$t \mapsto f\left(A + t(a \otimes b)\right)$$

is convex. ■

Proof 2 (free of regularities). Let $A_1, A_2 \in \mathbb{R}^\ell \otimes \mathbb{R}^d$ be two matrices such that $\operatorname{rank}(A_1 - A_2) = 1$. By the rank-one property, there exist non-zero vectors $a \in \mathbb{R}^\ell$ and $b \in \mathbb{R}^d$ such that

$$A_1 - A_2 = a \otimes b.$$

Let $\lambda \in (0,1)$ and define the convex combination matrix $A \in \mathbb{R}^\ell \otimes \mathbb{R}^d$ as

$$A = \lambda A_1 + (1-\lambda) A_2.$$

Our objective is to demonstrate the rank-one convexity inequality: $f(A) \leq \lambda f(A_1) + (1-\lambda) f(A_2)$. Using the definition of A, we can algebraically rewrite A_1 and A_2 in terms of A and the rank-one direction $a \otimes b$:

$$\begin{aligned} A_1 &= A + (1-\lambda) a \otimes b, \\ A_2 &= A - \lambda a \otimes b. \end{aligned}$$

To utilize the quasiconvexity of f, we construct a highly oscillating sequence of test functions (a laminate microstructure). We begin by defining a 1-periodic, piecewise affine function $h : \mathbb{R} \to \mathbb{R}$. We prescribe its derivative as:

$$h'(t) = \begin{cases} 1-\lambda & \text{if } t \in [0,\lambda), \\ -\lambda & \text{if } t \in [\lambda, 1), \end{cases}$$

and extend it periodically to all of $\mathbb{R}$. By integrating $h'(t)$, we obtain $h(t)$. Notice that over one period, $\int_0^1 h'(t)\, dt = \lambda(1-\lambda) + (1-\lambda)(-\lambda) = 0$, ensuring that $h(0) = h(1) = 0$ and that h is continuous and globally bounded.

Next, we introduce a scaling parameter $k \in \mathbb{N}$ to increase the oscillation frequency. Define $h_k(t) = \frac{1}{k} h(kt)$. As $k \to \infty$, the amplitude shrinks such that $h_k \to 0$ uniformly on $\mathbb{R}$, while its derivative $h_k'(t) = h'(kt)$ maintains the same oscillating values. We then define the vector-valued function $v_k : \mathbb{R}^d \to \mathbb{R}^\ell$ as:

$$v_k(x) = h_k(x \cdot b) a.$$

By the chain rule, the gradient of this function is $\nabla v_k(x) = h'(k(x \cdot b)) a \otimes b$.

Because v_k does not necessarily vanish on the boundary $\partial\Omega$, it is not an admissible test function for the quasiconvexity condition. To localize it, let $\eta \in C_c^\infty(\Omega)$ be a standard cut-off function satisfying $0 \le \eta \le 1$. We define our admissible test function $\varphi_k \in W_0^{1,\infty}(\Omega; \mathbb{R}^\ell)$ as:

$$\varphi_k(x) = \eta(x) v_k(x).$$

Applying the product rule, the gradient is given by:

$$\nabla\varphi_k(x) = \eta(x)\nabla v_k(x) + v_k(x) \otimes \nabla\eta(x).$$

We evaluate the quasiconvexity inequality at A using our valid test function φ_k:

$$f(A) \le \fint_\Omega f\big(A + \eta(x)\nabla v_k(x) + v_k(x) \otimes \nabla\eta(x)\big)\, dx. \tag{10.2}$$

We now rigorously pass to the limit as $k \to \infty$. Notice that $\nabla v_k(x)$ only takes values from the set $\{(1-\lambda)a \otimes b, -\lambda a \otimes b\}$. Because $0 \le \eta(x) \le 1$ and $\nabla\eta(x)$ is bounded, there exists a compact set $K \subset \mathbb{R}^\ell \otimes \mathbb{R}^d$ that contains the matrix $A + \eta(x)\nabla v_k(x)$ and the perturbed matrix $A + \eta(x)\nabla v_k(x) + v_k(x) \otimes \nabla\eta(x)$ for all $x \in \Omega$ and all $k \in \mathbb{N}$.

Assuming f is Carathéodory, we have f is continuous and thus it is uniformly continuous on the compact set K. As $k \to \infty$, $v_k \to 0$ uniformly, which implies $v_k(x) \otimes \nabla\eta(x) \to 0$ uniformly on Ω. By the uniform continuity of f on K, this guarantees that:

$$\lim_{k\to\infty} \Big| f\big(A + \eta(x)\nabla v_k(x) + v_k(x) \otimes \nabla\eta(x)\big) - f\big(A + \eta(x)\nabla v_k(x)\big)\Big| = 0$$

uniformly for all $x \in \Omega$.

The remaining sequence in the integrand is $f\big(A + \eta(x)\nabla v_k(x)\big)$. The argument of f oscillates rapidly between two discrete states:

- $A + \eta(x)(1-\lambda)a \otimes b$ (occurring on strips with volume fraction λ)
- $A - \eta(x)\lambda a \otimes b$ (occurring on strips with volume fraction $1-\lambda$)

By the Riemann-Lebesgue Lemma (Lemma 5.2.5) for highly oscillating periodic functions, the weak* limit of this composition converges to its expected value (the average over one period). Taking the limit of (10.2) as $k \to \infty$ yields:

$$f(A) \le \fint_\Omega \Big[\lambda f\big(A + \eta(x)(1-\lambda)a \otimes b\big) + (1-\lambda) f\big(A - \eta(x)\lambda a \otimes b\big)\Big]\, dx.$$

Finally, we must remove the influence of the cut-off function. Choose a sequence of cut-off functions $\eta_j \in C_c^\infty(\Omega)$ such that $\eta_j \uparrow \chi_\Omega$ almost everywhere in Ω. As $\eta_j(x) \to 1$ a.e., the arguments of f inside the integral converge point-wise back to our original matrices A_1 and A_2:

$$A + (1-\lambda)a \otimes b = A_1,$$
$$A - \lambda a \otimes b = A_2.$$

Applying the Dominated Convergence Theorem, we can pass the limit $\eta_j \to 1$ inside the integral to obtain:

$$f(A) \le \fint_\Omega \Big[\lambda f(A_1) + (1-\lambda) f(A_2)\Big]\, dx = \lambda f(A_1) + (1-\lambda) f(A_2).$$

This completes the direct proof that a continuous quasiconvex function is necessarily rank-one convex. ∎

11 Beyond Convexity - 3

In the modern calculus of variations, particularly in the study of microstructure and materials science, we often encounter the problem of differential inclusions. Instead of solving a standard partial differential equation, we seek a function whose gradient takes values in a prescribed set of matrices.

Remark 11.0.1 (Exact and Approximate Inclusions). This leads to two foundational problems:

- **Exact Inclusions:** Given a set of "allowed states" $K \subseteq \mathbb{R}^l \otimes \mathbb{R}^d$, we seek to understand the structure of functions $\Psi \in W^{1,\infty}(\Omega; \mathbb{R}^\ell)$ satisfying boundary conditions such that the weak derivative satisfies $D\Psi(x) \in K$ for almost every $x \in \Omega$. Equivalently, one can look for a matrix field $V \in L^\infty(\Omega; \mathbb{R}^\ell \otimes \mathbb{R}^d)$ such that $V(x) \in K$ a.e. and $\operatorname{curl} V = 0$ in the sense of distributions.
- **Approximate Inclusions:** Given the same set K, we study minimizing sequences or sequences of functions $\{\Psi_k\}$ such that the L^∞-distance to the set K vanishes, i.e., $\operatorname{dist}(D\Psi_k, K) \to 0$ in measure or almost everywhere.

The simplest, yet highly non-trivial, instance of these problems occurs when the target set K consists of exactly two matrices, $K = \{A, B\}$.

Let $\Omega \subset \mathbb{R}^d$ be an open, bounded, and connected domain. Let $\operatorname{Lip}(\Omega; \mathbb{R}^l)$ denote the space of Lipschitz continuous functions, which is isomorphic to the Sobolev space $W^{1,\infty}(\Omega; \mathbb{R}^l)$.

Theorem 11.0.2 (Ball-James)

Let $\Psi \in \operatorname{Lip}(\Omega; \mathbb{R}^l)$ be such that its weak derivative satisfies $D\Psi(x) \in \{A, B\}$ for almost every $x \in \Omega$, where $A, B \in \mathbb{R}^{l \times d}$.

a) If $\operatorname{rank}(A - B) \geq 2$, then the gradient must be strictly constant almost everywhere: either $D\Psi \equiv A$ a.e. or $D\Psi \equiv B$ a.e.

b) If $\{\Psi_k\}_{k=1}^\infty$ is a uniformly Lipschitz sequence of functions (i.e., $\sup_k \|D\Psi_k\|_{L^\infty} < \infty$) such that $\operatorname{dist}(D\Psi_k, \{A, B\}) \to 0$ almost everywhere, then, up to a subsequence, either

$$\int_\Omega |D\Psi_k - A| \, dx \to 0 \quad \text{or} \quad \int_\Omega |D\Psi_k - B| \, dx \to 0.$$

Proof of part a) in 2D. To illustrate the rigidity imposed by the rank-two condition, we provide the proof for Part (a) in the specific case where $l = d = 2$. Let $\Omega \subset \mathbb{R}^2$ be a connected domain. We assume without loss of generality that $D\Psi(x) \in \{A, B\}$ for almost every $x \in \Omega$.

Define the set $E = \{x \in \Omega : D\Psi(x) = B\}$. The complement up to a null set is $\Omega \setminus E = \{x \in \Omega : D\Psi(x) = A\}$. We can express the gradient globally using the indicator function $\mathbf{1}_E$ of the set E:

$$D\Psi = A + (B - A)\mathbf{1}_E. \tag{11.1}$$

Since $\Psi \in W^{1,\infty}(\Omega;\mathbb{R}^2)$, its distributional curl must vanish. Letting $C = B - A$, we have:

$$0 = \operatorname{curl}(D\Psi) = \operatorname{curl}\left(A + C\mathbf{1}_E\right) = \operatorname{curl}(C\mathbf{1}_E). \tag{11.2}$$

Here, the curl of a matrix field is applied row by row. Writing $C = \begin{pmatrix} C_1^1 & C_1^2 \\ C_2^1 & C_2^2 \end{pmatrix}$, equation (11.2) yields a system of equations in the sense of distributions. For a vector field $V = (V^1, V^2)$, $\operatorname{curl} V = \partial_1 V^2 - \partial_2 V^1 = 0$. For the row vectors of $C\mathbf{1}_E$, we factor out the constant matrix C:

$$\operatorname{curl}(C\mathbf{1}_E) = \begin{pmatrix} \partial_1(C_1^2\mathbf{1}_E) - \partial_2(C_1^1\mathbf{1}_E) \\ \partial_1(C_2^2\mathbf{1}_E) - \partial_2(C_2^1\mathbf{1}_E) \end{pmatrix} = \begin{pmatrix} C_1^1 & C_1^2 \\ C_2^1 & C_2^2 \end{pmatrix} \begin{pmatrix} -\partial_2\mathbf{1}_E \\ \partial_1\mathbf{1}_E \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{11.3}$$

This is equivalent to $C\nabla^\perp\mathbf{1}_E = 0$, where $\nabla^\perp = (-\partial_2, \partial_1)^\top$.

We are given that $\operatorname{rank}(A - B) \geq 2$. Since $C = B - A$ is a 2×2 matrix, having rank at least 2 implies that C is strictly of rank 2, and therefore invertible ($\det C \neq 0$). Thus, the kernel of C contains only the zero vector: $\ker C = \{0\}$.

This forces the distributional derivatives of the indicator function to vanish identically:

$$\begin{pmatrix} -\partial_2\mathbf{1}_E \\ \partial_1\mathbf{1}_E \end{pmatrix} \in \ker C \implies \nabla\mathbf{1}_E = 0 \text{ in } \mathscr{D}'(\Omega). \tag{11.4}$$

A fundamental theorem in the calculus of distributions states that a function whose distributional gradient vanishes on a connected domain must be a constant almost everywhere. Consequently, $\mathbf{1}_E$ must be constant. Since it is an indicator function, its only possible values are 0 or 1.

Therefore, we conclude:

$$\mathbf{1}_E = \begin{cases} 1 & \text{a.e. on } \Omega \\ 0 & \text{a.e. on } \Omega \end{cases} \tag{11.5}$$

This means either E has full measure ($D\Psi \equiv B$ a.e.) or E has measure zero ($D\Psi \equiv A$ a.e.), establishing the rigid phase separation. ■

§11.1 Lower Semicontinuity for Quasiconvex Functions at $p = +\infty$

With our intuitions built up, we can now prove lowersemicontinuity for quasiconvex functions in the case of $p = +\infty$. If we are given growth condition on f, we can actually have the same result for $1 < p < +\infty$ which will be given later.

Theorem 11.1.1 (Morrey)
Let $\Omega \subset \mathbb{R}^d$ be a bounded open set and $f : \mathbb{R}^\ell \otimes \mathbb{R}^d \to \mathbb{R}$ be a Carathéodory function. Assume that $\mathcal{F}(u) = \displaystyle\int_\Omega f(Du)\,dx$ is l.s.c. with respect to the weak* topology of $W^{1,\infty}(\Omega;\mathbb{R}^\ell)$, i.e.

$$\mathcal{F}(u) \leq \liminf_{k\to\infty} \mathcal{F}(u_k)$$

for any $u \in W^{1,\infty}(\Omega;\mathbb{R}^\ell)$ and any $\{u_k\} \subseteq W^{1,\infty}(\Omega;\mathbb{R}^\ell)$ such that $u_k \xrightarrow{L^\infty} u$ and $Du_k \rightharpoonup^* Du$. Then, f is quasiconvex.

Proof. The definition of quasiconvexity relies only on the properties of f over a unit cell, so without loss of generality, we may assume $\Omega = Q$, where $Q = [0,1)^d$ is the unit cube in $\mathbb{R}^d$. Notice that $|Q| = 1$.

Fix an arbitrary matrix $A \in \mathbb{R}^l \otimes \mathbb{R}^d$ and an arbitrary test function $\varphi \in W_0^{1,\infty}(Q;\mathbb{R}^l)$. To show quasiconvexity, our goal is to prove that $f(A) \leq \fint_Q f(A + D\varphi(x))\,dx$.

First, we extend φ to all of $\mathbb{R}^d$ by $\mathbb{Z}^d$-periodicity. Since $\varphi \in W_0^{1,\infty}(Q)$, this periodic extension remains in $W^{1,\infty}_{\text{loc}}(\mathbb{R}^d;\mathbb{R}^l)$. For each $k \in \mathbb{N}$, we define the highly oscillating sequence of functions:

$$\varphi_k(x) = \frac{1}{k}\varphi(kx).$$

By the chain rule, the weak gradient is $D\varphi_k(x) = D\varphi(kx)$.

Now we analyze the convergence of the sequence $u_k(x) := Ax + \varphi_k(x)$ as $k \to \infty$:

1. **Uniform Convergence:** Since φ is bounded (by M), $|\varphi_k(x)| \leq \dfrac{M}{k}$. Thus, $\varphi_k \to 0$ uniformly on Q, implying $u_k \xrightarrow{L^\infty} Ax$.
2. **Weak-* Convergence of Gradients:** The gradient is $Du_k(x) = A + D\varphi(kx)$. Because $D\varphi$ is Q-periodic and L^∞-bounded, the Riemann-Lebesgue Lemma states that $D\varphi(kx)$ converges weakly-* in $L^\infty(Q)$ to its mean value over one period. Since $\varphi = 0$ on ∂Q, by the Divergence Theorem, its mean value is zero:
$$\fint_Q D\varphi(x)\,dx = \fint_Q D\varphi(x)\,dx = 0.$$
Therefore, $Du_k \xrightharpoonup{L^\infty *} A + 0 = A$.

Together, these two conditions mean $u_k \xrightharpoonup{W^{1,\infty} *} Ax$. We can now apply the weak-* lower semicontinuity of $\mathcal{F}$:

$$f(A) = \int_Q f(A)\,dx = \mathcal{F}(Ax) \leq \liminf_{k\to\infty} \mathcal{F}(u_k) = \liminf_{k\to\infty} \int_Q f(A + D\varphi(kx))\,dx.$$

To conclude the proof, we evaluate the integral on RHS. We can decompose the unit cube Q into k^d disjoint subcubes Q_i of side length $1/k$. Using the change of variables $y = kx$ mapping each Q_i to Q, and exploiting the Q-periodicity of $D\varphi$, we obtain:

$$\begin{aligned}\int_Q f(A + D\varphi(kx))\,dx = \sum_{i=1}^{k^d} \int_{Q_i} f(A + D\varphi(kx))\,dx &= \sum_{i=1}^{k^d} k^{-d} \int_Q f(A + D\varphi(y))\,dy \\ &= k^d \left(k^{-d} \int_Q f(A + D\varphi(y))\,dy \right) \\ &= \int_Q f(A + D\varphi(y))\,dy.\end{aligned}$$

The sequence of integrals is therefore constant for all k. Substituting this back into our lower semicontinuity inequality yields:

$$f(A) \leq \fint_Q f(A + D\varphi(y))\,dy.$$

Since A and φ were arbitrary, f is quasiconvex. ∎

With analogous argument, we can also deduce the following result.

Theorem 11.1.2
Let $\Omega \subset \mathbb{R}^d$ be a bounded open set and $f : \Omega \times \mathbb{R} \times \mathbb{R}^\ell \otimes \mathbb{R}^d \to \mathbb{R}$ be a Carathéodory function. Assume that $\mathcal{F}(u) = \int_\Omega f(x, u, (x), Du(x))\, dx$ is l.s.c. with respect to the weak* topology of $W^{1,\infty}(\Omega; \mathbb{R}^\ell)$, i.e.

$$\mathcal{F}(u) \leq \liminf_{k\to\infty} \mathcal{F}(u_k)$$

for any $u \in W^{1,\infty}(\Omega; \mathbb{R}^\ell)$ and any $\{u_k\} \subseteq W^{1,\infty}(\Omega; \mathbb{R}^\ell)$ such that $u_k \xrightarrow{L^\infty} u$ and $Du_k \rightharpoonup^* Du$. Then, for a.e. $x \in \Omega$ and for every $z \in \mathbb{R}^\ell$, the mapping $p \mapsto f(x, z, p)$ is quasiconvex.

§11.2 Lower Semicontinuity for Polyconvex Functions

A result given by Ball provides lower semicontinuity for the case of polyconvex functions as follows.

Theorem 11.2.1 (Ball's Theorem)
Let f be a Carathéodory and polyconvex function and Ω be an open bounded set with Lipschitz boundary. Then, the functional $\mathcal{F}(u) = \int_\Omega f(Du)\, dx$ is sequentially lower semicontinuous (l.s.c.) with respect to weak $W^{1,p}$ convergence for $p > \ell \wedge d$.

Proposition 11.2.2
Let $p > \ell \wedge d$ and suppose $u_k \rightharpoonup u$ weakly in $W^{1,p}(\Omega; \mathbb{R}^l)$. Then for the mapping of minors $\mathcal{M}$, we have

$$\mathcal{M}(Du_k) \rightharpoonup \mathcal{M}(Du) \quad \text{weakly in } L^{p/d}(\Omega).$$

Remark 11.2.3. Actually, a more general statement is true: If $u_k \rightharpoonup u$ weakly in $W^{1,p}$ and $m \leq \min\{p, \ell, d\}$, then the M-th order minors converge weakly, i.e.,

$$\mathcal{M}_m(Du_k) \rightharpoonup \mathcal{M}_m(Du) \quad \text{weakly in } L^{p/m}(\Omega)$$

even when $p/m = 1$. However, for the critical case $p/m = 1$, one need different approach from the proof presented here.

Proof of Ball's Theorem (Assuming Proposition 11.2.2 is True). Since f is polyconvex, there exists a convex function g such that $f(\xi) = g(\mathcal{M}(\xi))$ for all matrices ξ. To ensure the integrals are well-defined and Fatou's lemma can be applied, we assume g is bounded from below.

Let $u_k \rightharpoonup u$ weakly in $W^{1,p}(\Omega; \mathbb{R}^l)$ for $p > d$. By the assumed Proposition, the sequence of minors converges weakly in the Banach space $L^{p/d}(\Omega)$:

$$\mathcal{M}(Du_k) \rightharpoonup \mathcal{M}(Du) \quad \text{weakly in } L^{p/d}(\Omega).$$

Now, we define the functional $\mathcal{G}(v) = \int_\Omega g(v(x))\,dx$ for $v \in L^{p/d}(\Omega)$. We analyze the properties of $\mathcal{G}$ to justify the weak lower semicontinuity:

1. **Convexity:** Since the integrand g is a convex function, the functional $\mathcal{G}$ is strictly convex on $L^{p/d}(\Omega)$.

2. **Strong Lower Semicontinuity:** Because g is convex on a finite-dimensional space, it is continuous. If a sequence $v_k \to v$ strongly in $L^{p/d}(\Omega)$, there exists a subsequence v_{k_j} that converges to v pointwise almost everywhere and it attains the limit inferior of $\{G(v_k)\}$. Applying Fatou's Lemma (using the lower bound of g), we get:

$$\liminf_{k\to\infty} \mathcal{G}(v_k) = \liminf_{j\to\infty} \int_\Omega g(v_{k_j}(x))\,dx \geq \int_\Omega g(v(x))\,dx = \mathcal{G}(v).$$

   Since this holds for any converging subsequence, the functional $\mathcal{G}$ is strongly lower semicontinuous in $L^{p/d}(\Omega)$.

Therefore, we can use Theorem 4.2.3 to have weak lower semicontinuity of $\mathcal{G}$ on $L^{p/d}(\Omega)$.

We can now apply the weak lower semicontinuity of $\mathcal{G}$ to our weakly converging sequence of minors, $v_k = \mathcal{M}(Du_k) \rightharpoonup \mathcal{M}(Du) = v$:

$$\liminf_{k\to\infty} \int_\Omega g(\mathcal{M}(Du_k))\,dx \geq \int_\Omega g(\mathcal{M}(Du))\,dx.$$

Substituting back our original definition $f(Du) = g(\mathcal{M}(Du))$, this inequality becomes:

$$\liminf_{k\to\infty} \int_\Omega f(Du_k)\,dx \geq \int_\Omega f(Du)\,dx,$$

which rigorously concludes the proof that the functional is sequentially weakly lower semicontinuous in $W^{1,p}$. ■

12 Beyond Convexity - 4

§12.1 Proof on Lower Semicontinuity of the Minors

Proof of Proposition 11.2.2. We focus on the case $m = 2$, without loss of generality, we can assume $\ell = d = 2$ as well. Let $p > 2$ and suppose $u_k \rightharpoonup u$ weakly in $W^{1,p}(\Omega; \mathbb{R}^2)$.

Step 1: Integration by Parts (Piola Identity) for $m = 2$. For any smooth function $v \in C^2(\Omega; \mathbb{R}^2)$, the determinant can be rewritten as a divergence:

$$\det Dv = \partial_1 v^1 \partial_2 v^2 - \partial_2 v^1 \partial_1 v^2 = \partial_1 (v^1 \partial_2 v^2) - \partial_2 (v^1 \partial_1 v^2).$$

By multiplying by a test function $\varphi \in C_c^1(\Omega; \mathbb{R})$ and integrating by parts, we shift the derivative onto φ:

$$\int_\Omega \varphi \det Dv \, dx = -\int_\Omega \left(v^1 \partial_2 v^2 \partial_1 \varphi - v^1 \partial_1 v^2 \partial_2 \varphi \right) dx. \tag{12.1}$$

By a standard density approximation argument, this identity (12.1) holds for all $v \in W^{1,p}$ with $p \geq 2$.

Step 2: Passing to the Limit on Dense Set of $L^{p/2}$. Applying (12.1) to our weakly converging sequence u_k:

$$\int_\Omega \varphi \det Du_k \, dx = -\int_\Omega u_k^1 \partial_2 u_k^2 \partial_1 \varphi \, dx + \int_\Omega u_k^1 \partial_1 u_k^2 \partial_2 \varphi \, dx.$$

We need to justify why this converges to the corresponding expression for u. Because $u_k \rightharpoonup u$ weakly in $W^{1,p}(\Omega)$, the Rellich-Kondrachov Compactness Theorem (Theorem 6.1.9) guarantees that $u_k \to u$ *strongly* in $L^2_{\text{loc}}(\Omega)$. Since φ has compact support, we are integrating over a bounded domain, so $u_k^1 \to u^1$ strongly in $L^2(\operatorname{supp} \varphi)$.

Simultaneously, weak convergence in $W^{1,p}$ implies that the derivatives converge weakly: $\partial_j u_k^2 \rightharpoonup \partial_j u^2$ weakly in L^p, and thus weakly in L^2_{loc} (since $p \geq 2$).

In functional analysis, the inner product of a strongly convergent sequence and a weakly convergent sequence passes to the limit. Therefore:

$$\int_\Omega (u_k^1 \partial_1 \varphi) \partial_2 u_k^2 \, dx \xrightarrow{k \to \infty} \int_\Omega (u^1 \partial_1 \varphi) \partial_2 u^2 \, dx.$$

Applying this to both terms yields:

$$\int_\Omega \varphi \det Du_k \, dx \xrightarrow{k\to\infty} -\int_\Omega \left(u^1 \partial_2 u^2 \partial_1 \varphi - u^1 \partial_1 u^2 \partial_2 \varphi \right) dx = \int_\Omega \varphi \det Du \, dx.$$

This proves that $\det Du_k \to \det Du$ in the sense of distributions (for all $\varphi \in C_c^1(\Omega)$).

Step 3: Extending to Weak Convergence. We have established convergence against C_c^1 test functions, but we must upgrade this to weak convergence in $L^{p/2}(\Omega)$.

Notice that the sequence of determinants is uniformly bounded in the $L^{p/2}$ norm by the L^p norm of the gradients:

$$\| \det Du_k \|_{L^{p/2}} \leq \| Du_k \|_{L^p}^2 \leq K.$$

Since $p > 2$ (because $p > d = 2$), it follows that $p/2 > 1$, which makes $L^{p/2}(\Omega)$ a reflexive Banach space.

Since we have uniform boundedness and convergence with respect to the dense subset $C_c^1(\Omega) \subset L^{p/2}(\Omega)$. We can invoke Theorem 5.1.1 to assert $\det Du_k \rightharpoonup \det Du$ weakly in $L^{p/2}(\Omega)$.

■

Proof of Proposition (General Case for m-th Order Minors). We proceed by induction on the order of the minor, m.

Base Case ($m = 1$): This is trivially the weak convergence of the gradient $Du_k \rightharpoonup Du$ in $L^p(\Omega)$, which holds by assumption. The case $m = 2$ was also fully established in the previous step.

Induction Hypothesis: Assume that for a given integer m, the claim holds for all $(m-1)$-th order minors. Specifically, if $u_k \rightharpoonup u$ weakly in $W^{1,p}$ with $p > m-1$, then the $(m-1)$-th order minors converge weakly:

$$\mathcal{M}_{m-1}(Du_k) \rightharpoonup \mathcal{M}_{m-1}(Du) \quad \text{weakly in } L^{p/(m-1)}(\Omega).$$

Inductive Step: We now show the result holds for m-th order minors assuming $p > m$. Let $\mathcal{M}_{\alpha\beta}(Du)$ be the specific $m \times m$ minor corresponding to rows $\alpha = (\alpha_1, \dots, \alpha_m)$ and columns $\beta = (\beta_1, \dots, \beta_m)$. In the language of differential forms, this minor satisfies the relation:

$$\mathcal{M}_{\alpha\beta}(Du)\, dx^1 \wedge \cdots \wedge dx^d = du^{\alpha_1} \wedge \cdots \wedge du^{\alpha_m} \wedge \hat{dx^\beta},$$

where $\hat{dx^\beta}$ denotes the wedge product of all standard basis differentials $dx^1 \wedge \cdots \wedge dx^d$ *excluding* $dx^{\beta_1}, \dots, dx^{\beta_m}$.

Because the exterior derivative of a differential is zero ($d^2 = 0$), we can express the m-form as an exact differential:

$$du^{\alpha_1} \wedge \cdots \wedge du^{\alpha_m} = d(u^{\alpha_1} du^{\alpha_2} \wedge \cdots \wedge du^{\alpha_m}).$$

Multiplying by a smooth test function $\varphi \in C_c^1(\Omega)$ and integrating by parts (which is equivalent to applying Stokes' Theorem for forms), the derivative shifts onto φ:

$$\int_\Omega \varphi\, du^{\alpha_1} \wedge \cdots \wedge du^{\alpha_m} = -\int_\Omega u^{\alpha_1} d\varphi \wedge (du^{\alpha_2} \wedge \cdots \wedge du^{\alpha_m}).$$

Evaluating this identity along our weakly converging sequence $u_k \rightharpoonup u$:

$$\int_\Omega \varphi\, du_k^{\alpha_1} \wedge \cdots \wedge du_k^{\alpha_m} = -\int_\Omega u_k^{\alpha_1} d\varphi \wedge \left(du_k^{\alpha_2} \wedge \cdots \wedge du_k^{\alpha_m}\right).$$

We analyze the components on the right-hand side as $k \to \infty$:

1. **Strong Convergence:** By the Rellich-Kondrachov Theorem, $u_k \rightharpoonup u$ in $W^{1,p}$ implies $u_k^{\alpha_1} \to u^{\alpha_1}$ *strongly* in $L^p_{\text{loc}}(\Omega)$.
2. **Smooth Test Function:** The exterior derivative $d\varphi$ is a fixed, smooth, and bounded 1-form.

3. **Weak Convergence (Induction):** The term $\left(du_k^{\alpha_2} \wedge \cdots \wedge du_k^{\alpha_m}\right)$ represents an $(m-1)$-th order minor. Since $p > m > m-1$, our induction hypothesis guarantees this sequence converges *weakly* to $(du^{\alpha_2} \wedge \cdots \wedge du^{\alpha_m})$ in $L^{p/(m-1)}(\Omega)$.

To pass to the limit, we rely on the fact that the product of a strongly converging sequence and a weakly converging sequence converges in the sense of distributions, provided their dual exponents are compatible. Here, $\frac{1}{p} + \frac{m-1}{p} = \frac{m}{p} < 1$ (since $p > m$). Therefore, the integral safely passes to the limit:

$$\int_\Omega \varphi\, \mathcal{M}_{\alpha\beta}(Du_k)\, dx \xrightarrow{k\to\infty} \int_\Omega \varphi\, \mathcal{M}_{\alpha\beta}(Du)\, dx.$$

This establishes distributional convergence against $C_c^1(\Omega)$ functions. Because $p > m$, the sequence $\mathcal{M}_{\alpha\beta}(Du_k)$ is uniformly bounded in $L^{p/m}(\Omega)$. Using the exact same density argument as in the $m = 2$ case, Since we have uniform boundedness and convergence with respect to the dense subset $C_c^1(\Omega) \subset L^{p/m}(\Omega)$. We can invoke Theorem 5.1.1 to assert $\mathcal{M}_m(Du_k) \rightharpoonup \mathcal{M}_m(Du)$ weakly in $L^{p/m}(\Omega)$. This completes the induction step and finishes the proof. ■

Lemma 12.1.1 (Div-Curl Lemma)
Let $B_1 \subset \mathbb{R}^d$ be a ball, and let the vector fields $U_k, V_k \rightharpoonup U, V$ weakly in $L^2(B_1;\mathbb{R}^d)$. Assume that:

1. $\operatorname{div} U_k = 0$ in the sense of distributions,
2. $\operatorname{curl} V_k = 0$ in the sense of distributions (i.e., $V_k = \nabla f_k$ for some scalar functions f_k).

Then, the inner product converges in the sense of distributions:

$$\int_{B_1} \varphi\langle V_k, U_k\rangle\, dx \xrightarrow{k\to\infty} \int_{B_1} \varphi\langle V, U\rangle\, dx \quad \forall \varphi \in C_c^\infty(B_1).$$

Remark 12.1.2 (Connection to Weak Continuity of Minors). The weak continuity of the 2×2 determinant, which we proved in the previous proposition, is actually the classic, motivating example for the Div-Curl Lemma.

§12.2 Lower Semicontinuity for Quasiconvex Functions at $1 < p < +\infty$

Theorem 12.2.1 (Marcellini)
Let $\Omega \subset \mathbb{R}^d$ be an open bounded set with Lipschitz boundary. Let $f \colon \mathbb{R}^l \otimes \mathbb{R}^d \to \mathbb{R}$ be a continuous and quasiconvex function satisfying the polynomial growth condition:

$$0 \leq f(\xi) \leq C(1 + |\xi|^p)$$

for some $1 \leq p < \infty$ and a constant $C > 0$. Then, the functional $\mathcal{F}(u) = \int_\Omega f(Du)\,dx$ is sequentially lower semicontinuous (l.s.c.) with respect to the weak $W^{1,p}$ convergence.

Theorem 12.2.2 (Acerbi-Fusco)
The same lower semicontinuity result is true also for integrands depending explicitly on the spatial variable and the function value, i.e., $f(x, z, p)$.

Remark 12.2.3. The polynomial growth theorem above is not completely satisfactory since it excludes energy functionals which are highly relevant in physical applications, such as in nonlinear elasticity. Examples of such excluded energies are:

- Energies for Neo-Hookean materials:

  $$\int_\Omega \left(|Du|^2 + g(\det Du)\right) dx,$$

  where g is a convex function.

- More general polyconvex energies:

  $$\int_\Omega g(Du, \mathcal{M}_2(Du), \det Du)\,dx,$$

  where $Du \in \mathbb{R}^3 \otimes \mathbb{R}^3$.

In these physical models, the function $g(A, M, t)$ (where $t = \det Du$) typically exhibits singular behavior to prevent infinite compression or infinite expansion. As sketched in the margins, we require the energy to blow up at the extremes:

$$g(A, M, t) \to \infty \quad \text{as } t \to \infty$$

and

$$g(A, M, t) \to \infty \quad \text{as } t \to 0^+$$

This required singular behavior as $\det Du \to 0^+$ clearly violates the global polynomial upper bound $|f(\xi)| \leq C(1 + |\xi|^p)$ assumed in Marcellini's Theorem.

We split the proof into three steps with the first two steps provided in this lecture. The third step is contained in Lecture 13.

Step 1 Proof of Theorem 12.2.1: Establishing the local Lipschitz bound. Since f is quasiconvex, it is necessarily rank-one convex. Rank-one convexity means $t \mapsto f(A + t(a \otimes b))$

is convex for any vectors a, b. In particular, by choosing the standard basis vectors e_i and e_j, f is convex with respect to each individual matrix component (component-wise convex).

To see how the growth condition $|f(\xi)| \le C(1+|\xi|^p)$ yields the bound on the difference quotient, we first isolate the mechanism in one dimension.

Lemma 12.2.4 (1D Growth to Derivative Bound)
Let $h : \mathbb{R} \to \mathbb{R}$ be a convex function satisfying $|h(t)| \le M(1+|t|^p)$ for some $p \ge 1$. Then there exists a constant $c > 0$ such that for all $s, t \in \mathbb{R}$:

$$|h(t) - h(s)| \le c(1 + |t|^{p-1} + |s|^{p-1})|t - s|.$$

Proof of Lemma. By symmetry, we can assume without loss of generality that $0 \le s < t$ (the reader is invited to be assured our assumption is enough). We split this into cases based on the location of s and t:

Case 1 ($1 \le s < t$): By the monotonicity of slopes for convex functions, the difference quotient between s and t is bounded above by the slope of the secant line from t to $2t$:

$$\frac{h(t)-h(s)}{t-s} \le \frac{h(2t)-h(t)}{2t-t} = \frac{h(2t)-h(t)}{t}.$$

Applying the polynomial growth bound to the numerator, and noting that $\frac{1}{t} \le 1 \le t^{p-1}$ since $t \ge 1$:

$$\frac{h(2t)-h(t)}{t} \le \frac{M(1+2^p t^p) + M(1+t^p)}{t} \le M' \frac{t^p}{t} = M' t^{p-1}.$$

To bound the absolute value, we also need a lower bound. Looking to the left towards 0:

$$\frac{h(t)-h(s)}{t-s} \ge \frac{h(s)-h(0)}{s-0} \ge \frac{-M(1+s^p) - M}{s} \ge -M'' s^{p-1}.$$

Therefore, $\left|\dfrac{h(t)-h(s)}{t-s}\right| \le \max(M' t^{p-1}, M'' s^{p-1}) \le c(t^{p-1} + s^{p-1})$.

Case 2 ($0 \le s < t \le 1$): Convex functions are locally Lipschitz on bounded domains. Since h is bounded on $[-1, 2]$ by our growth condition, it is Lipschitz continuous on the compact subinterval $[0, 1]$ with some constant L. Hence:

$$|h(t) - h(s)| \le L|t - s| \le L(1 + t^{p-1} + s^{p-1})|t - s|.$$

Case 3 ($0 \le s \le 1 \le t$): We simply bridge the gap by adding and subtracting $h(1)$:

$$|h(t) - h(s)| \le |h(t) - h(1)| + |h(1) - h(s)|,$$

and apply Case 1 to the first term and Case 2 to the second. The lemma is proved. ■

Returning to Step 1: Let $\xi, \eta \in \mathbb{R}^\ell \otimes \mathbb{R}^d$. Because f is component-wise convex, we can move from η to ξ by changing exactly one matrix component at a time. This creates a path of $N = \ell \times d$ steps: $\eta = M_0, M_1, \ldots, M_N = \xi$.

For each step k, the matrices M_k and M_{k-1} differ only in one coordinate. Let's call the scalar value of this coordinate t and s, respectively. The function restricted to this line, $h(\tau) = f(M_{k-1} + (\tau - s)e_i \otimes e_j)$, is a 1D convex function.

Because the other components are fixed along this line, the growth condition becomes $|h(\tau)| \leq C'(1 + |\tau|^p)$. Applying our 1D Lemma to each of the N steps and summing them up (using the triangle inequality and the equivalence of norms in finite-dimensional spaces) immediately yields the desired multidimensional bound:

$$|f(\xi) - f(\eta)| \leq c(1 + |\xi|^{p-1} + |\eta|^{p-1})|\xi - \eta|.$$

■

Step 2 Proof of Theorem 12.2.1: Proving l.s.c. for the case u affine. Let $u(x) = x_0 + Ax$ be an affine function, and suppose $u_k \rightharpoonup u$ weakly in $W^{1,p}(\Omega; \mathbb{R}^\ell)$.

Step 2.1: Particular Case with Additional Assumption.
We further assume that $u_k \equiv u$ on $\partial\Omega$. We define the perturbation $\varphi_k = u_k - u$, which belongs to $W_0^{1,p}(\Omega; \mathbb{R}^\ell)$.

Since Ω has a Lipschitz boundary, we can approximate φ_k strongly in $W_0^{1,p}$ by smooth functions with compact support in $W_0^{1,\infty}$. Furthermore, the polynomial growth condition $|f(\xi)| \leq C(1 + |\xi|^p)$ guarantees that the integral functional $v \mapsto \int_\Omega f(Dv)\, dx$ is strongly continuous with respect to the $W^{1,p}$ norm following the same argument for Part (a) of Proposition 6.2.4.

Because the functional is strongly continuous, the quasiconvexity inequality, which inherently holds for $W_0^{1,\infty}$ test functions, extends directly to $\varphi_k \in W_0^{1,p}(\Omega)$:

$$\int_\Omega f(A)\, dx \leq \int_\Omega f(A + D\varphi_k)\, dx = \int_\Omega f(Du_k)\, dx.$$

Since u is affine, we know $Du = A$ $\int_\Omega f(Du)\, dx = \int_\Omega f(A)\, dx$. Taking the limit infimum as $k \to \infty$ immediately yields the desired lower semicontinuity for this specific case.

Step 2.2: The General Case (De Giorgi Averaging Method).
We want to show $\int_\Omega f(Du)\, dx \leq \liminf_{k\to\infty} \int_\Omega f(Du_k)\, dx$. Let $\Omega_0 \Subset \Omega$ be an arbitrary compactly contained subdomain, and let $d = \operatorname{dist}(\Omega_0, \partial\Omega) > 0$.

For a fixed integer $N \in \mathbb{N}$, we construct a sequence of $N + 1$ strictly nested subdomains:

$$\Omega_i := \left\{ x \in \Omega : \operatorname{dist}(x, \Omega_0) < \frac{i}{N} d \right\} \quad \text{for } i = 0, 1, \ldots, N.$$

Notice that $\Omega_0 \Subset \Omega_1 \Subset \cdots \Subset \Omega_N \Subset \Omega$.

For each layer $i \in \{0, 1, \ldots, N-1\}$, we define a cutoff function $\varphi_i \in C_c^\infty(\Omega_{i+1}; [0, 1])$ satisfying:

- $\varphi_i \equiv 1$ on Ω_i and $\varphi_i \equiv 0$ outside Ω_{i+1}.
- $\operatorname{supp}(D\varphi_i) \subset \overline{\Omega}_{i+1} \setminus \Omega_i$.
- $|D\varphi_i| \leq \dfrac{CN}{d}$ for some absolute constant $C > 0$.

We define the mixed sequence $v_{k,i} = \varphi_i u_k + (1-\varphi_i)u = u + \varphi_i(u_k - u)$. Because $v_{k,i} = u$ outside Ω_{i+1}, $v_{k,i} - u \in W_0^{1,p}(\Omega)$. As established in Step 2.1, the lower semicontinuity holds for functions matching u on the boundary, we have:

$$\int_\Omega f(Du)\,dx \le \int_\Omega f(Dv_{k,i})\,dx.$$

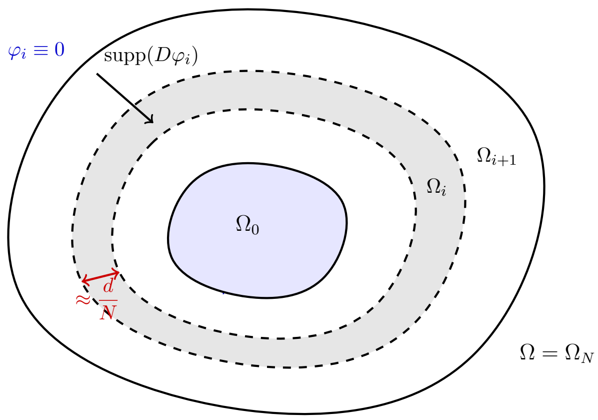


Figure 12.1: Construction of the nested subdomains Ω_i for the De Giorgi averaging method. The cutoff function φ_i transitions from 1 to 0 strictly within the shaded annulus $\Omega_{i+1} \setminus \Omega_i$, which has a thickness proportional to d/N.

We split the integral on the right into three disjoint regions: Ω_i, the transition layer $\Omega_{i+1} \setminus \Omega_i$, and the exterior $\Omega \setminus \Omega_{i+1}$, and noting $f \ge 0$.

$$\begin{aligned}\int_\Omega f(Du)\,dx &\le \int_{\Omega_i} f(Du_k)\,dx + \int_{\Omega_{i+1}\setminus\Omega_i} f(Dv_{k,i})\,dx + \int_{\Omega\setminus\Omega_{i+1}} f(Du)\,dx \\ &\le \int_\Omega f(Du_k)\,dx + \int_{\Omega_{i+1}\setminus\Omega_i} f(Dv_{k,i})\,dx + \int_{\Omega\setminus\Omega_0} f(Du)\,dx.\end{aligned}$$

In the transition layer, $Dv_{k,i} = \varphi_i Du_k + (1-\varphi_i)Du + (u_k - u) \otimes D\varphi_i$. Using the growth condition $|f(\xi)| \le C(1+|\xi|^p)$ and the convexity of $t \mapsto t^p$, we bound the transition energy:

$$\int_{\Omega_{i+1}\setminus\Omega_i} f(Dv_{k,i})\,dx \le C \int_{\Omega_{i+1}\setminus\Omega_i} \left(1 + |Du_k|^p + |Du|^p + \left(\frac{N}{d}\right)^p |u_k - u|^p\right) dx.$$

Now, we sum the inequalities over all $i = 0, \dots, N-1$. Because the transition layers $\Omega_{i+1} \setminus \Omega_i$ are pairwise disjoint, summing them yields an integral strictly over $\Omega_N \setminus \Omega_0 \subset \Omega \setminus \Omega_0$:

$$\begin{aligned}N \int_\Omega f(Du)\,dx \le N \int_\Omega f(Du_k)\,dx &+ C \int_{\Omega\setminus\Omega_0} (1 + |Du_k|^p + |Du|^p)\,dx \\ &+ C\left(\frac{N}{d}\right)^p \int_\Omega |u_k - u|^p\,dx + N \int_{\Omega\setminus\Omega_0} f(Du)\,dx.\end{aligned}$$

Divide by N to find the average:

$$\begin{aligned}\int_\Omega f(Du)\,dx \le \int_\Omega f(Du_k)\,dx &+ \frac{C}{N} \int_{\Omega\setminus\Omega_0} (1 + |Du_k|^p + |Du|^p)\,dx \\ &+ \frac{CN^{p-1}}{d^p} \int_\Omega |u_k - u|^p\,dx + \int_{\Omega\setminus\Omega_0} f(Du)\,dx.\end{aligned}$$

Finally, we take limits in the correct order to eliminate the error terms:

1. **Limit as $k \to \infty$:** By weak convergence of $W^{1,p}$, $u_k \to u$ strongly in $L^p(\Omega)$. Thus, $\displaystyle\int |u_k - u|^p \to 0$ and the sequence $\{Du_k\}$ is uniformly bounded in L^p.
2. **Limit as $N \to \infty$:** The term divided by N vanishes.
3. **Limit as $\Omega_0 \nearrow \Omega$:** The Lebesgue measure of $\Omega \setminus \Omega_0$ goes to zero, so
$$\int_{\Omega \setminus \Omega_0} f(Du)\, dx \to 0.$$

This sequence of limits leaves exactly the desired inequality:

$$\int_{\Omega} f(Du)\, dx \leq \liminf_{k\to\infty} \int_{\Omega} f(Du_k)\, dx.$$

■

13 Hilbert's 19th Problem - 0: Regularity of Minimizers

§13.1 Continuation on Proof of Lower Semicontinuity for Quasiconvex Functions at $1 < p < +\infty$

Step 3 Proof of Theorem 12.2.1: Proving l.s.c. for the general case. We now drop the assumption that u is affine. Let $u_k \rightharpoonup u$ weakly in $W^{1,p}(\Omega)$.

Phase 1: Grid Approximation. Fix a large integer $N \gg 1$. Since Ω is open, there exists a finite family $\mathcal{F}$ of disjoint open cubes Q_i with side length $1/N$ contained entirely in Ω, such that the leftover measure is small: $|\Omega \setminus \bigcup_{Q_i \in \mathcal{F}} Q_i| \leq \frac{1}{N}$. Let $\Omega_N = \bigcup_{Q_i \in \mathcal{F}} Q_i$.

On each cube Q_i, we define the constant matrix $z_i = ⨍_{Q_i} Du \, dx$, which is the average of the limit gradient. We define the global piecewise constant approximation $Z_N = \sum_{Q_i \in \mathcal{F}} z_i \mathbf{1}_{Q_i}$.

Remark 13.1.1 (The Egorov/Lebesgue Exercise). By the density of continuous functions, for any fixed $\varepsilon > 0$, there exists a uniformly continuous function $g \in C_c(\Omega; \mathbb{R}^{l \times d})$ such that $\|Du - g\|_{L^p(\Omega)} < \varepsilon$.

By the triangle inequality, we can bound the L^p error on Ω_N by three terms:

$$\|Du - Z_N(Du)\|_{L^p(\Omega_N)} \leq \|Du - g\|_{L^p(\Omega_N)} + \|g - Z_N(g)\|_{L^p(\Omega_N)} + \|Z_N(g) - Z_N(Du)\|_{L^p(\Omega_N)}.$$

We estimate each term independently:

1. **The density bound:** By our choice of g, $\|Du - g\|_{L^p(\Omega_N)} \leq \|Du - g\|_{L^p(\Omega)} < \varepsilon$.
2. **The contraction bound:** Because $t \mapsto |t|^p$ is convex, Jensen's inequality guarantees that the averaging operator Z_N is a contraction in L^p. Thus:

$$\begin{aligned}\|Z_N(g - Du)\|^p_{L^p(\Omega_N)} = \sum_{Q_i} \int_{Q_i} \left| ⨍_{Q_i} (g - Du) \, dy \right|^p &\leq \sum_{Q_i} \int_{Q_i} |g - Du|^p \, dx \\ &\leq \|g - Du\|^p_{L^p(\Omega)} \\ &< \varepsilon^p.\end{aligned}$$

   Therefore, the third term is bounded by ε.

3. **The uniform continuity bound:** Since g is uniformly continuous on Ω, the oscillation of g within any cube Q_i vanishes as the grid size $1/N \to 0$. Specifically, $\sup_{x \in Q_i} |g(x) - ⨍_{Q_i} g| \to 0$ uniformly across all cubes in $\mathcal{F}$. Therefore, for N sufficiently large, we have $\|g - Z_N(g)\|_{L^p(\Omega_N)} < \varepsilon$.

Combining these bounds, for N large enough, $\|Du - Z_N(Du)\|_{L^p(\Omega_N)} < 3\varepsilon$. Because $\varepsilon > 0$ is arbitrary, we conclude that $\lim_{N \to \infty} \|Du - Z_N(Du)\|_{L^p(\Omega_N)} = 0$.

Phase 2: Localized Affine Perturbations. The intention is to approximate u locally on each cube by its first-order Taylor expansion: $u(x) \approx \fint_{Q_i} u + z_i(x - x_i)$, where x_i is the center of Q_i. For each Q_i, we define the sequence:

$$v_{k,i}(x) = u_k(x) - u(x) + \fint_{Q_i} u\, dx + z_i(x - x_i).$$

Taking the gradient, we get $Dv_{k,i} = Du_k - Du + z_i$. Note that $|Dv_{k,i} - Du_k| = |Du - z_i|$. As $k \to \infty$, since $u_k - u \rightharpoonup 0$ weakly in $W^{1,p}$, we have $v_{k,i} \rightharpoonup \fint_{Q_i} u + z_i(x - x_i)$ weakly in $W^{1,p}(Q_i)$. Because the weak limit is an affine function, we can apply the result from Step 2 on each cube:

$$\liminf_{k\to\infty} \int_{Q_i} f(Dv_{k,i})\, dx \geq \int_{Q_i} f(z_i)\, dx. \quad \forall Q_i \in \mathcal{F}.$$

Phase 3: Telescoping Decomposition. Since $f \geq 0$, we can restrict our integral to the sub-domain Ω_N. We decompose the integral algebraically by adding and subtracting terms:

$$\begin{aligned}
\int_\Omega f(Du_k)\, dx \geq \int_{\Omega_N} f(Du_k)\, dx &= \sum_{Q_i \in \mathcal{F}} \int_{Q_i} f(Du_k)\, dx \\
&= \sum_{Q_i \in \mathcal{F}} \int_{Q_i} \big(f(Du_k) - f(Dv_{k,i})\big)\, dx && \text{(Error I)} \\
&\quad + \sum_{Q_i \in \mathcal{F}} \int_{Q_i} \big(f(Dv_{k,i}) - f(z_i)\big)\, dx && \text{(``Error'' II)} \\
&\quad + \sum_{Q_i \in \mathcal{F}} \int_{Q_i} \big(f(z_i) - f(Du)\big)\, dx && \text{(Error III)} \\
&\quad + \int_{\Omega_N} f(Du)\, dx. && \text{(Error IV related)}
\end{aligned}$$

Phase 4: Bounding the Errors. We analyze the limit of each term as $k \to \infty$, and then as $N \to \infty$.

- **Term I:** Using the local Lipschitz bound from Step 1 and Hölder's inequality:

$$\begin{aligned}
|\mathrm{I}| &\leq \sum_{Q_i} \int_{Q_i} C\big(1 + |Du_k|^{p-1} + |Dv_{k,i}|^{p-1}\big)|Du_k - Dv_{k,i}|\, dx \\
&\leq \int_{\Omega_N} C'\big(1 + |Du_k|^{p-1} + |Du|^{p-1} + |Z_N|^{p-1}\big)|Du - Z_N|\, dx \\
&\leq C' \underbrace{\big\|1 + |Du_k|^{p-1} + |Du|^{p-1} + |Z_N|^{p-1}\big\|_{L^{\frac{p}{p-1}}(\Omega_N)}}_{\text{Uniformly bounded } O_k(1)} \underbrace{\|Du - Z_N\|_{L^p(\Omega_N)}}_{=o_N(1)\to 0 \text{ as } N\to\infty}.
\end{aligned}$$

- **Term II:** By the affine L.S.C. property established above, $\liminf_{k\to\infty} \mathrm{II} \geq 0$.
- **Term III:** By the exact same Lipschitz argument used for Term I (replacing Du_k with Z_N and $Dv_{k,i}$ with Du), we find $|\mathrm{III}| \leq C'\|Du - Z_N\|_{L^p(\Omega_N)}$, which goes to 0 as $N \to \infty$.

- **Term IV:** We rewrite $\int_{\Omega_N} f(Du)\,dx = \int_\Omega f(Du)\,dx - \int_{\Omega\setminus\Omega_N} f(Du)\,dx$. Because $f(Du) \in L^1(\Omega)$ and $|\Omega \setminus \Omega_N| \le 1/N \to 0$, the subtracted integral goes to 0 by the absolute continuity of the Lebesgue integral.

Conclusion: Fix $\varepsilon > 0$. Choose $N \gg 1$ large enough such that the spatial errors $|\mathrm{I}| + |\mathrm{III}| + \int_{\Omega\setminus\Omega_N} f(Du) < \varepsilon$ independent of k.

Taking the limit infimum as $k \to \infty$ of our decomposition yields:

$$\liminf_{k\to\infty} \int_\Omega f(Du_k)\,dx \ge -\varepsilon + 0 - \varepsilon + \int_\Omega f(Du)\,dx - \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, taking $\varepsilon \to 0^+$ concludes the proof:

$$\liminf_{k\to\infty} \int_\Omega f(Du_k)\,dx \ge \int_\Omega f(Du)\,dx.$$

■

§13.2 Regarding De Giorgi-Nash-Moser Regularity Theorem

Hilbert's 19th Problem: *Are minimizers of regular variational problems regular?* We consider a scalar variational problem where the Lagrangian $f\colon \mathbb{R}^d \to \mathbb{R}$ satisfies the following conditions:

1. f is smooth.
2. f is strictly convex. More specifically, we assume f is uniformly elliptic, meaning there exist constants $0 < \lambda \le \Lambda < \infty$ such that:

$$\lambda I_d \le D^2 f(\xi) \le \Lambda I_d \quad \forall \xi \in \mathbb{R}^d.$$

Exercise 13.2.1 (Growth Bounds Exercise)

The uniform ellipticity assumption implies the quadratic growth bounds:

$$f(\xi) \ge c|\xi|^2 - C \quad \text{and} \quad |f(\xi)| + |\xi||Df(\xi)| \le C(1 + |\xi|^2).$$

Solution. **Lower Bound:**
By the Fundamental Theorem of Calculus, we integrate twice:

$$f(\xi) - f(0) = \int_0^1 \frac{d}{dt} f(t\xi)\,dt = \int_0^1 \langle Df(t\xi), \xi\rangle\,dt,$$
$$Df(t\xi) - Df(0) = \int_0^t \frac{d}{ds} Df(s\xi)\,ds = \int_0^t D^2 f(s\xi)\xi\,ds.$$

Substituting the second equation into the first yields the exact Taylor expansion:

$$f(\xi) = f(0) + \langle Df(0), \xi\rangle + \int_0^1 \int_0^t \langle D^2 f(s\xi)\xi, \xi\rangle\,ds\,dt.$$

By the uniform ellipticity assumption, we have $\langle D^2 f(s\xi)\xi, \xi\rangle \geq \lambda|\xi|^2$. Therefore:

$$f(\xi) \geq f(0) + \langle Df(0), \xi\rangle + \frac{\lambda}{2}|\xi|^2.$$

Next, applying the Cauchy-Schwarz and Young's inequalities (with $\epsilon = \frac{\lambda}{2}$), we bound the linear term:

$$\langle Df(0), \xi\rangle \geq -|Df(0)||\xi| \geq -\frac{|Df(0)|^2}{2\epsilon} - \frac{\epsilon|\xi|^2}{2} = -\frac{|Df(0)|^2}{\lambda} - \frac{\lambda}{4}|\xi|^2.$$

Plugging this back in, we obtain the desired lower bound with $c = \frac{\lambda}{4}$:

$$f(\xi) \geq \frac{\lambda}{4}|\xi|^2 + \left(f(0) - \frac{|Df(0)|^2}{\lambda}\right).$$

Upper Bound:
The uniform ellipticity also provides an upper bound on the Hessian: $D^2 f(\xi) \leq \Lambda I$. Integrating the Hessian once gives a bound on the gradient:

$$|Df(\xi)| \leq |Df(0)| + \int_0^1 |D^2 f(t\xi)\xi|\,dt \leq |Df(0)| + \Lambda|\xi|.$$

Multiplying by $|\xi|$ and applying Young's inequality yields:

$$|\xi||Df(\xi)| \leq |Df(0)||\xi| + \Lambda|\xi|^2 \leq \frac{1}{2}|Df(0)|^2 + \left(\Lambda + \frac{1}{2}\right)|\xi|^2.$$

Integrating the gradient bound provides a bound on the function itself:

$$\begin{aligned} |f(\xi)| \leq |f(0)| + \int_0^1 |Df(t\xi)||\xi|\,dt &\leq |f(0)| + |Df(0)||\xi| + \frac{\Lambda}{2}|\xi|^2 \\ &\leq |f(0)| + \frac{1}{2}|Df(0)|^2 + \left(\frac{\Lambda+1}{2}\right)|\xi|^2. \end{aligned}$$

Summing the bounds for $|f(\xi)|$ and $|\xi||Df(\xi)|$, we can absorb all constants evaluated at 0 and the coefficients of $|\xi|^2$ into a single generic constant C, concluding:

$$|f(\xi)| + |\xi||Df(\xi)| \leq C(1 + |\xi|^2).$$

■

By the direct method in the Calculus of Variations (Theorem 6.2.1), for any boundary data $g \in W^{1,2}(\Omega)$, there exists a unique minimizer $u \in W^{1,2}(\Omega)$ such that:

$$\int_\Omega f(Du)\,dx = \min_{v \in g + W_0^{1,2}(\Omega)} \int_\Omega f(Dv)\,dx.$$

This minimizer u weakly solves the Euler-Lagrange (EL) equation:

$$\int_\Omega Df(Du) \cdot D\varphi\,dx = 0 \quad \forall \varphi \in W_0^{1,2}(\Omega) \implies -\operatorname{div}(Df(Du)) = 0 \quad \text{in the weak sense.}$$

Interior Regularity and Linearization

Global regularity on $\overline{\Omega}$ depends strictly on the smoothness of both the boundary $\partial\Omega$ and the boundary data g. Instead, we focus on **interior regularity** on subdomains $\Omega' \Subset \Omega$. By a standard covering argument, it suffices to understand the regularity when the domain is a ball $B_R(x)$.

To investigate the regularity of u, let us assume u is smooth enough to differentiate the EL equation. Differentiating with respect to an arbitrary directional vector $e \in \mathbb{S}^{d-1}$ yields:

$$\partial_e\big(\operatorname{div}(Df(Du))\big) = 0 \implies \operatorname{div}(D^2 f(Du) D(\partial_e u)) = 0.$$

If we define $v = \partial_e u$ and the matrix $A(x) = D^2 f(Du(x))$, we see that the directional derivative v solves a linear elliptic equation:

$$\operatorname{div}(A(x)Dv) = 0.$$

Because f is uniformly elliptic, the coefficient matrix $A(x)$ is also uniformly elliptic: $\langle A(x)\xi, \xi\rangle \geq \lambda|\xi|^2$ for all $\xi \in \mathbb{R}^d$.

The Schauder Bootstrapping Mechanism

The core idea of Schauder Theory is that Hölder continuous coefficients elevate the regularity of the solution:

$$A \in C^{k,\alpha}(B_R(x)) \implies v \in C^{k+1,\alpha}(B_{R/2}(x)).$$

If we can somehow prove that the gradient is Hölder continuous (i.e., $Du \in C^\alpha$, which implies $A \in C^\alpha$ since f is smooth), this triggers a powerful bootstrapping loop:

$$\begin{aligned} A \in C^\alpha &\implies v \in C^{1,\alpha} \\ &\implies u \in C^{2,\alpha} \quad (\text{since } v = Du) \\ &\implies A \in C^{1,\alpha} \quad (\text{since } A(x) = D^2 f(Du(x))) \\ &\implies v \in C^{2,\alpha} \\ &\implies u \in C^{3,\alpha} \\ &\implies A \in C^{2,\alpha} \implies \dots \end{aligned}$$

Through this iteration, the solution bootstraps to C^∞. The critical, non-trivial missing link is establishing that initial C^α regularity for $A(x)$, which is exactly what the De Giorgi-Nash-Moser theorem provides.

14 Hilbert's 19th Problem - 1

§14.1 Statement and Outline of Proof for the Theorem

From the previous lecture, we understood our preliminary ingredients on deducting the regularity of a minimizer. We shall state the result of De Giorgi-Nash-Moser which will be our main goal for this part of the lectures.

Theorem 14.1.1 (De Giorgi-Schauder)
Let $f\colon \mathbb{R}^d \to \mathbb{R}$ such that:

1. f is smooth (analytic).
2. f is strictly convex. More specifically, we assume f is uniformly elliptic, meaning there exist constants $0 < \lambda \leq \Lambda < \infty$ such that:

$$\lambda I_d \leq D^2 f(\xi) \leq \Lambda I_d \quad \forall \xi \in \mathbb{R}^d.$$

Then, for any bounded open set $\Omega \subset \mathbb{R}^d$ and $g \in W^{1,2}(\Omega)$, there exists a unique minimizer $u \in g + W_0^{1,2}(\Omega)$ such that

$$\int_\Omega f(Du)\, dx = \min_{v \in g + W_0^{1,2}(\Omega)} \int_\Omega f(Dv)\, dx$$

and $u \in C^\infty(\Omega)$ (analytic).

Remark 14.1.2. Utilizing the result of Exercise 13.2.1, we have f coercive in the sense of Theorem 6.2.1 and thus we have a minimizer $u \in g + W_0^{1,2}(\Omega)$. Moreover, from strict convexity, the results of Proposition 6.2.4 indeed ensure the uniqueness of the minimizer and any minimizer satisfies the property:

$$u \text{ is a minimizer} \iff u \text{ solves } -\operatorname{div}(Df(Du)) = 0 \quad \text{in the weak sense (EL)}.$$

This theorem focuses on the regularity of u, which will be proven analytic.

Remark 14.1.3. The statement of smoothness for u is false in the vectorial case (systems of equations). Partial regularity theory is required instead.

Note that we are dealing with interior regularity. By a standard covering argument, we can assume $\Omega = B_R(x_0)$. Up to the change of variable and rescaling $u(x) \in W^{1,2}(B_R(x_0))$ into $\tilde{u}(x) = \dfrac{u(Rx + x_0)}{R} \in W^{1,2}(B_1(0))$, we can assume our domain is the unit ball $\Omega = B_1(0)$.

Our proof will be based on the Schauder bootstrapping mechanism. Our strategy relies on differentiating the Euler-Lagrange equation with respect to an arbitrary directional vector $e \in \mathbb{S}^{d-1}$:

$$\partial_e(\operatorname{div}(Df(Du))) = 0 \implies \operatorname{div}(D^2 f(Du) D(\partial_e u)) = 0.$$

If we denote $v = \partial_e u$, we see that v solves a linear elliptic equation:

$$-\operatorname{div}(A(x)Dv) = 0$$

where $A(x) = D^2 f(Du(x))$. Because f is uniformly elliptic, the coefficient matrix $A(x)$ is uniformly elliptic as well ($\lambda I_d \le A(x) \le \Lambda I_d$). The result of Schauder, which we use strictly for bootstrapping (stated here without proof), is as follows.

Theorem 14.1.4 (Schauder Theory)
Let $A \in C^{k,\alpha}(B_1)$ be a coefficient matrix satisfying $\lambda I_d \le A \le \Lambda I_d$ for some $0 < \lambda < \Lambda$. Then, every weak solution of $-\operatorname{div}(A(x)Dv) = 0$ belongs to $C^{k+1,\alpha}(B_{1/2})$ and satisfies the estimate:

$$\|v\|_{C^{k+1,\alpha}(B_{1/2})} \le C(\lambda, \Lambda, [A]_{C^{k,\alpha}(B_1)}) \cdot \|v\|_{W^{1,2}(B_1)}.$$

Remark 14.1.5. Schauder Theory is fundamentally a *perturbative theory.* By freezing the coefficients at a point x_0, one can rewrite the operator as $-\Delta v = \operatorname{div}((A(x) - I_d)Dv)$. If $A(x)$ is continuous, the perturbation term $(A(x) - I_d)$ is arbitrarily small in a small neighborhood, allowing the smoothness of the Laplace operator to transfer to v. However, if $A(x)$ is only bounded and measurable, this error is $O(1)$, and the perturbation argument fails.

Following Schauder's result, the bootstrapping mechanism activates formally:

Corollary 14.1.6
Let f and u be as in Theorem 14.1.1. If $u \in C^{1,\alpha}(B_1)$, then $u \in C^\infty(B_{1/2})$.

Indeed, Theorem 14.1.1 is rigorously proved if we can establish the missing link: that $v = \partial_e u \in C^{0,\alpha}$ for any direction e. We will reach this final goal by proving the following monumental theorem.

Theorem 14.1.7 (De Giorgi-Nash-Moser Theorem)
Let $A(x)$ be a measurable matrix satisfying $\lambda I_d \le A(x) \le \Lambda I_d$. If $v \in W^{1,2}(B_1)$ is a weak solution of $-\operatorname{div}(A(x)Dv) = 0$, then there exist constants $\beta = \beta(d, \lambda, \Lambda) > 0$ and $C = C(d, \lambda, \Lambda) > 0$ such that $v \in C^{0,\beta}(B_{1/2})$ and

$$\sup_{\substack{x,y \in B_{1/2} \\ x \ne y}} \frac{|v(x) - v(y)|}{|x - y|^\beta} = [v]_{C^{0,\beta}(B_{1/2})} \le C \cdot \|v\|_{L^2(B_1)}.$$

Moreover, by rescaling on $B_R(x_0)$, we have:

$$R^\beta [v]_{C^{0,\beta}(B_{R/2}(x_0))} \le C \left(\fint_{B_R(x_0)} v^2 \, dx \right)^{\frac{1}{2}}.$$

To reach this theorem, we will navigate through the following technicalities:

1. Deducing the Caccioppoli inequality to control local gradient energy.

2. For dimension $d = 2$, establishing Hölder continuity directly via the hole-filling technique and Campanato spaces.
3. For higher dimensions ($d \geq 3$), asserting Hölder continuity by working on three main intermediate results: Caccioppoli inequality on level sets, interior boundedness of subsolutions ($L^2 - L^\infty$ estimates), and the Oscillation Lemma.

Remark 14.1.8 (The One-Dimensional Case). In dimension $d = 1$, the advanced machinery of De Giorgi, Caccioppoli, and Schauder is entirely unnecessary. The domain reduces to an interval $\Omega = (a, b)$, and the weak Euler-Lagrange equation becomes a standard ordinary differential equation:

$$-\frac{d}{dx}\big(f'(u'(x))\big) = 0 \implies f'(u'(x)) = c$$

for some constant $c \in \mathbb{R}$. The uniform ellipticity (strict convexity) assumption guarantees that $f''(\xi) \geq \lambda > 0$. This implies the first derivative f' is strictly monotonically increasing and therefore globally invertible. Applying the inverse yields a constant gradient:

$$u'(x) = (f')^{-1}(c) = m$$

Integrating this shows that any minimizer must be an affine function, $u(x) = mx + k$. Since affine functions are trivially C^∞ (and analytic), the interior regularity problem is immediately solved without requiring energy bounds or bootstrapping.

§14.2 Caccioppoli Inequality and the Hole-Filling Technique

The cornerstone of De Giorgi's method is the strict control of the gradient's L^2 energy by the function's own L^2 mass.

Theorem 14.2.1 (Caccioppoli Inequality)
Let $v \in W^{1,2}(B_1)$ be a weak solution of $-\operatorname{div}(A(x)Dv) = 0$ with $\lambda I_d \leq A(x) \leq \Lambda I_d$. Then, for any constant $c \in \mathbb{R}$, there exists a constant $C = C(\lambda, \Lambda) > 0$ such that:

$$\int_{B_{1/2}} |Dv|^2 \, dx \leq C(\lambda, \Lambda) \int_{B_1} (v - c)^2 \, dx.$$

By rescaling on balls $B_R \subset B_1$, this generalizes to

$$\int_{B_{R/2}} |Dv|^2 \, dx \leq \frac{C}{R^2} \int_{B_R} (v - c)^2 \, dx.$$

Proof. Choose a smooth cutoff function $\varphi \in C_c^\infty(B_1)$ such that $0 \leq \varphi \leq 1$, $\varphi \equiv 1$ on $B_{1/2}$, and $|D\varphi| \leq 4$. We test the weak formulation of the PDE using the test function $\psi = \varphi^2(v - c) \in W_0^{1,2}(B_1)$:

$$\int_{B_1} \langle A(x)Dv, D(\varphi^2(v - c))\rangle \, dx = 0.$$

Applying the product rule $D\psi = \varphi^2 Dv + 2\varphi(v - c)D\varphi$, we obtain:

$$\int_{B_1} \varphi^2 \langle A(x)Dv, Dv\rangle \, dx = -2 \int_{B_1} \varphi(v - c)\langle A(x)Dv, D\varphi\rangle \, dx.$$

We apply the uniform ellipticity lower bound on the LHS and the Cauchy-Schwarz inequality combined with the upper bound Λ on the RHS ($|A(x)Dv| \le \Lambda |Dv|$):

$$\lambda \int_{B_1} \varphi^2 |Dv|^2 \, dx \le 2\Lambda \int_{B_1} (\varphi |Dv|)(|v-c||D\varphi|) \, dx.$$

Applying Young's inequality with Peter-Paul scaling ($2ab \le \epsilon a^2 + \frac{1}{\epsilon} b^2$) on the RHS:

$$\lambda \int_{B_1} \varphi^2 |Dv|^2 \, dx \le \Lambda \int_{B_1} \left(\epsilon \varphi^2 |Dv|^2 + \frac{1}{\epsilon} |v-c|^2 |D\varphi|^2 \right) dx.$$

By choosing $\epsilon = \frac{\lambda}{2\Lambda}$, we can absorb the gradient term into the LHS:

$$\frac{\lambda}{2} \int_{B_1} \varphi^2 |Dv|^2 \, dx \le \frac{2\Lambda^2}{\lambda} \int_{B_1} |v-c|^2 |D\varphi|^2 \, dx.$$

Since $\varphi \equiv 1$ on $B_{1/2}$ and $|D\varphi| \le 4$, we restrict the LHS to $B_{1/2}$ and estimate the RHS:

$$\int_{B_{1/2}} |Dv|^2 \, dx \le \frac{64\Lambda^2}{\lambda^2} \int_{B_1} (v-c)^2 \, dx.$$

■

Remark 14.2.2. A stronger inequality holds from our choice of φ. Notice that $\varphi \equiv 1$ on $B_{1/2}$. Therefore, $\operatorname{supp}(D\varphi)$ is contained in the annulus $B_1 \setminus B_{1/2}$ and by the bound $|D\varphi| \le 4$ we have

$$\int_{B_{1/2}} |Dv|^2 \, dx \le \frac{64\Lambda^2}{\lambda^2} \int_{B_1 \setminus B_{1/2}} (v-c)^2 \, dx.$$

The true power of the Caccioppoli inequality emerges when we select c strategically to exploit the Poincaré inequality.

Corollary 14.2.3 (Gradient Bound on Annulus)
If v is a weak solution as above, then

$$\int_{B_{1/2}} |Dv|^2 \, dx \le C \int_{B_1 \setminus B_{1/2}} |Dv|^2 \, dx.$$

Proof. In the Caccioppoli inequality, choose $c = \bar{v}_A = \fint_{B_1 \setminus B_{1/2}} v \, dx$, the integral average of v over the annulus. Because $v - \bar{v}_A$ has zero mean on the annulus, we apply the Poincaré-Sobolev inequality to bound the L^2 norm of the function strictly by the L^2 norm of its gradient on the annulus:

$$\int_{B_1} (v - \bar{v}_A)^2 \, dx \le \tilde{C} \int_{B_1 \setminus B_{1/2}} |Dv|^2 \, dx.$$

Substituting this into Theorem 14.2.1 after considering the sharper estimate yields the result. ■

This structural bound allows us to perform the **Hole-Filling Technique** (or Widman's trick). By adding $C \int_{B_{1/2}} |Dv|^2$ to both sides of the inequality from Corollary 14.2.3, we "fill the hole" in the domain of integration on the RHS:

$$(C+1) \int_{B_{1/2}} |Dv|^2 \, dx \leq C \left(\int_{B_1 \setminus B_{1/2}} |Dv|^2 \, dx + \int_{B_{1/2}} |Dv|^2 \, dx \right) = C \int_{B_1} |Dv|^2 \, dx.$$

Defining $\theta = \dfrac{C}{C+1} < 1$, we observe a geometric decay of the gradient energy:

$$\int_{B_{1/2}} |Dv|^2 \, dx \leq \theta \int_{B_1} |Dv|^2 \, dx.$$

By iterating this decay for arbitrary radii $r < R$, we deduce that v belongs to a Morrey space, satisfying:

$$\int_{B_r} |Dv|^2 \, dx \leq C \left(\frac{r}{R} \right)^{\alpha} \int_{B_R} |Dv|^2 \, dx \quad \text{where } \alpha = \log_2(1/\theta) > 0.$$

Remark 14.2.4 (The Two-Dimensional Case). By applying the Poincaré inequality to the Morrey decay above, we can bound the oscillation of v:

$$\int_{B_r} (v - \bar{v}_r)^2 \, dx \leq C_P r^2 \int_{B_r} |Dv|^2 \, dx \leq \tilde{C} r^{2+\alpha}.$$

This places v within the Campanato space $\mathcal{L}^{2,2+\alpha}$. Campanato's Isomorphism Theorem guarantees that if the scaling exponent strictly exceeds the dimension ($2 + \alpha > d$), the function is locally Hölder continuous with exponent $\beta = \dfrac{2+\alpha-d}{2}$.

If $d = 2$, the strict inequality $2 + \alpha > 2$ is trivially satisfied since $\alpha > 0$. Thus, $v \in C^{0,\alpha/2}$, and the regularity problem is completely solved for the two-dimensional case via hole-filling. However, for $d \geq 3$, the algebraic constant α (which depends entirely on Λ/λ) is generally too small to satisfy $2 + \alpha > d$. The failure of hole-filling in higher dimensions necessitates the complex De Giorgi iteration scheme on level sets.

Exercise 14.2.5

Let $v \in W^{1,2}(B_1)$ be a weak solution to $-\operatorname{div}(A(x)Dv) = 0$ with $\lambda I \leq A(x) \leq \Lambda I$. Show that for all $R < S \leq 1$,

$$\int_{B_R} |Dv|^2 \, dx \leq \frac{C(\Lambda/\lambda)}{(S-R)^2} \inf_c \int_{B_S} (v-c)^2 \, dx.$$

Solution. Let $c \in \mathbb{R}$ be arbitrary and choose a smooth cut-off function $\varphi \in C_c^\infty(B_S)$ such that $0 \leq \varphi \leq 1$, $\varphi \equiv 1$ on B_R, and $|D\varphi| \leq \dfrac{2}{S-R}$.

We test the weak formulation of the PDE with the test function $\psi = \varphi^2(v-c) \in W_0^{1,2}(B_S)$. Applying the product rule $D\psi = \varphi^2 Dv + 2\varphi(v-c)D\varphi$, we obtain:

$$\int_{B_S} \varphi^2 \langle A(x)Dv, Dv \rangle \, dx = -2 \int_{B_S} \varphi(v-c) \langle A(x)Dv, D\varphi \rangle \, dx.$$

Using the uniform ellipticity $\langle A(x)Dv, Dv\rangle \geq \lambda|Dv|^2$ and the spectral bound $|A(x)Dv| \leq \Lambda|Dv|$:

$$\lambda \int_{B_S} \varphi^2|Dv|^2\,dx \leq 2\Lambda \int_{B_S} (\varphi|Dv|)(|v-c||D\varphi|)\,dx.$$

Applying Young's inequality with Peter-Paul scaling ($\epsilon = \frac{\lambda}{2\Lambda}$) to the right-hand side yields:

$$\lambda \int_{B_S} \varphi^2|Dv|^2\,dx \leq \frac{\lambda}{2}\int_{B_S} \varphi^2|Dv|^2\,dx + \frac{2\Lambda^2}{\lambda}\int_{B_S} |v-c|^2|D\varphi|^2\,dx.$$

Absorbing the gradient term into the left-hand side and utilizing the properties of φ on the right-hand side gives:

$$\frac{\lambda}{2}\int_{B_R} |Dv|^2\,dx \leq \frac{2\Lambda^2}{\lambda}\int_{B_S} |v-c|^2\frac{4}{(S-R)^2}\,dx.$$

Multiplying by $\frac{2}{\lambda}$, we establish the bound:

$$\int_{B_R} |Dv|^2\,dx \leq \frac{16(\Lambda/\lambda)^2}{(S-R)^2}\int_{B_S} (v-c)^2\,dx.$$

Since the inequality holds for any constant $c \in \mathbb{R}$, taking the infimum over c yields the desired result with $C(\Lambda/\lambda) = 16(\Lambda/\lambda)^2$. ■

Exercise 14.2.6

Let $\mathcal{F}_C = \left\{u \in W^{1,2}(B_1) : \int_{B_1} |Du|^2\,dx \leq C\int_{B_1} u^2\,dx\right\}$. Show that any vector subspace $V \subset \mathcal{F}_C$ is finite-dimensional.

Solution. Let V be a vector subspace of $W^{1,2}(B_1)$ strictly contained in $\mathcal{F}_C$. We equip V with the standard L^2 norm and consider its closed unit ball:

$$B_V = \{u \in V : \|u\|_{L^2(B_1)} \leq 1\}.$$

Because $V \subset \mathcal{F}_C$, every function $u \in B_V$ satisfies the gradient energy bound:

$$\|Du\|^2_{L^2(B_1)} \leq C\|u\|^2_{L^2(B_1)} \leq C.$$

Consequently, the Sobolev norm of any $u \in B_V$ is uniformly bounded:

$$\|u\|^2_{W^{1,2}} = \|u\|^2_{L^2} + \|Du\|^2_{L^2} \leq 1 + C.$$

This shows that B_V is a uniformly bounded set in $W^{1,2}(B_1)$. By the Rellich-Kondrachov Theorem, the embedding $W^{1,2}(B_1) \hookrightarrow L^2(B_1)$ is compact. Therefore, the bounded set B_V must be precompact in $L^2(B_1)$.

Since V is a normed vector space (under the L^2 norm) whose unit ball B_V is precompact, Riesz's Lemma guarantees that V must be finite-dimensional. ■

15 Hilbert's 19th Problem - 2

§15.1 Higher Regularity for Solutions to $-\operatorname{div}(Df(Du)) = 0$

In the previous lecture, we established the existence of a minimizer $u \in W^{1,2}(\Omega)$ for the functional $\int_\Omega f(Du)\,dx$ and observed that if the gradient Du is locally Hölder continuous, Schauder's perturbative theory automatically bootstraps the regularity to C^∞.

To bridge the gap between $u \in W^{1,2}$ and $Du \in C^{0,\alpha}_{\text{loc}}$, we decompose our main objective (Theorem 14.1.7) into two distinct, highly non-trivial regularity results.

Theorem 15.1.1 ($W^{2,2}$ Regularity)
Let f be an uniformly elliptic Lagrangian such that $\lambda I_d \leq D^2 f(\xi) \leq \Lambda I_d$. If $u \in W^{1,2}(B_1)$ is a weak solution to the non-linear Euler-Lagrange equation

$$-\operatorname{div}(Df(Du)) = 0,$$

then $u \in W^{2,2}_{\text{loc}}(B_1)$. Consequently, the gradient Du admits weak derivatives in L^2_{loc}.

Theorem 15.1.2 (De Giorgi for Linear Equations)
Let $A(x)$ be a bounded, measurable coefficient matrix satisfying the uniform ellipticity condition $\lambda I_d \leq A(x) \leq \Lambda I_d$. If $v \in W^{1,2}(B_1)$ is a weak solution to the linear equation

$$-\operatorname{div}(A(x)Dv) = 0,$$

then $v \in C^{0,\alpha}_{\text{loc}}(B_1)$ for some exponent $\alpha \in (0,1)$ depending only on d, λ, and Λ.

The full regularity of the minimizer follows immediately from the concatenation of these two theorems.

Assume $u \in W^{1,2}(B_1)$ is our minimizer. By Theorem 15.1.1, we gain the exact regularity needed to differentiate the Euler-Lagrange equation. Specifically, since $u \in W^{2,2}_{\text{loc}}(B_1)$, the directional derivative $v = \partial_e u$ is well-defined as a function in $W^{1,2}_{\text{loc}}(B_1)$.

Differentiating the weak formulation $-\operatorname{div}(Df(Du)) = 0$ with respect to an arbitrary direction $e \in \mathbb{S}^{d-1}$ yields:

$$-\operatorname{div}(D^2 f(Du) D(\partial_e u)) = 0 \implies -\operatorname{div}(A(x)Dv) = 0,$$

where $A(x) = D^2 f(Du(x))$. Because f is uniformly elliptic, $A(x)$ strictly satisfies $\lambda I_d \leq A(x) \leq \Lambda I_d$.

At this stage, $v = \partial_e u \in W^{1,2}_{\text{loc}}$ solves a linear uniformly elliptic equation with measurable coefficients. We may therefore apply Theorem 15.1.2, which immediately grants $v \in C^{0,\alpha}_{\text{loc}}$. Since this holds for any direction e, the full gradient is Hölder continuous: $Du \in C^{0,\alpha}_{\text{loc}}$.

Once $Du \in C^{0,\alpha}_{\text{loc}}$, the coefficient matrix $A(x) = D^2 f(Du(x))$ is also $C^{0,\alpha}_{\text{loc}}$. This activates the Schauder bootstrapping mechanism introduced in the previous lecture, elevating u to $C^{2,\alpha}$, then $C^{3,\alpha}$, and ultimately C^∞.

The remainder of this section is strictly devoted to proving Theorem 15.1.1.

To rigorously prove Theorem 15.1.1, we must show that the second weak derivatives of u exist. However, we cannot directly differentiate the Euler-Lagrange equation because we do not yet know if D^2u exists. To bypass this circular dependency, we must utilize Nirenberg's method of finite difference quotients.

For a function $u(x)$, a unit vector $e \in \mathbb{S}^{d-1}$, and a non-zero scalar $h \in \mathbb{R}$, we define the finite difference quotient as:

$$\delta_{e,h}u(x) = \frac{u(x+he) - u(x)}{h}.$$

We take h to be sufficiently small in absolute value so that the shifted points remain inside the domain. The following foundational lemma from functional analysis establishes that bounding these difference quotients is perfectly equivalent to possessing a weak derivative.

Lemma 15.1.3 (Translation Lemma / Difference Quotients)
Let $p \in (1, +\infty]$ and let B_1 be the unit ball centered at the origin. For any function $u \in L^p(B_1)$, the following statements are equivalent:

1. $u \in W^{1,p}(B_{1/2})$.
2. There exists a constant $C > 0$ such that for any direction $e \in \mathbb{S}^{d-1}$ and any scalar $0 < |h| < 1/2$, the difference quotient satisfies $\|\delta_{e,h}u\|_{L^p(B_{1/2})} \leq C$.

Furthermore, if statement (1) holds, the constant in (2) can be taken as $C = \|Du\|_{L^p(B_1)}$.

Remark 15.1.4. For $1 < p \leq \infty$, the proof that (2) $\implies$ (1) relies fundamentally on the weak compactness of bounded sets in L^p. For the boundary case $p = 1$, the space L^1 lacks this weak compactness property. In that scenario, condition (2) instead characterizes the space of functions of Bounded Variation, $u \in BV(B_{1/2})$.

Proof. **Proof of (1) $\implies$ (2):**
Assume $u \in W^{1,p}(B_1)$. By the density of smooth functions (and the absolutely continuous on lines property of Sobolev functions), we can write the difference as the integral of the weak derivative along the segment connecting x and $x + he$:

$$u(x+he) - u(x) = h \int_0^1 \langle Du(x+the), e\rangle \, dt.$$

Dividing by h and taking the absolute value yields:

$$|\delta_{e,h}u(x)| \leq \int_0^1 |Du(x+the)| \, dt.$$

Applying Jensen's inequality (or Hölder's inequality) to the convex function $t \mapsto |t|^p$ with respect to the probability measure dt on $[0, 1]$:

$$|\delta_{e,h}u(x)|^p \leq \int_0^1 |Du(x+the)|^p \, dt.$$

We now integrate both sides with respect to x over the ball $B_{1/2}$. Using Fubini's Theorem to swap the integrals and performing the simple change of variables $y = x + the$:

$$\int_{B_{1/2}} |\delta_{e,h} u(x)|^p \, dx \le \int_0^1 \left(\int_{B_{1/2}} |Du(x + the)|^p \, dx \right) dt$$
$$\le \int_0^1 \left(\int_{B_{1/2+|h|}} |Du(y)|^p \, dy \right) dt.$$

Since $|h| < 1/2$, the shifted domain $B_{1/2+|h|}$ is strictly contained within B_1. The inner integral is independent of t, leaving us with:

$$\|\delta_{e,h} u\|^p_{L^p(B_{1/2})} \le \|Du\|^p_{L^p(B_1)} .$$

Taking the p-th root yields the desired bound with $C = \|Du\|_{L^p(B_1)}$.

Proof of $(2) \implies (1)$**:**
Assume that the difference quotients are uniformly bounded: $\|\delta_{e,h} u\|_{L^p(B_{1/2})} \le C$ for all $0 < |h| < 1/2$.

Because $1 < p \le \infty$, the space $L^p(B_{1/2})$ is reflexive (or is the dual space of L^1, for $p = \infty$). By the Banach-Alaoglu Theorem, any bounded sequence in L^p admits a weakly convergent subsequence.

Therefore, for any fixed direction e, we can extract a sequence $h_k \to 0$ and find a function $v_e \in L^p(B_{1/2})$ such that:

$$\delta_{e,h_k} u \rightharpoonup v_e \quad \text{weakly in } L^p(B_{1/2}) \text{ as } k \to \infty.$$

To prove that v_e is the weak derivative of u, we must show it satisfies the integration by parts formula. Let $\varphi \in C^\infty_c(B_{1/2})$ be a smooth test function. Using the discrete integration by parts identity (which simply shifts the summation/integral index):

$$\int_{B_{1/2}} \delta_{e,h_k} u(x)\varphi(x) \, dx = -\int_{B_{1/2}} u(x)\delta_{e,-h_k}\varphi(x) \, dx.$$

Note: When performing the change of variables $y = x + h_k e$ for the discrete integration by parts, the region of integration technically shifts:

$$\int_{B_{1/2}(0)} u(x + h_k e)\varphi(x) \, dx = \int_{B_{1/2}(h_k e)} u(y)\varphi(y - h_k e) \, dy$$

However, because $\text{supp}(\varphi) \subset B_{1/2}(0)$, we have $\text{dist}(\text{supp}(\varphi), \partial B_{1/2}(0)) > 0$. By taking h_k sufficiently small, the shifted support $\text{supp}(\varphi) + h_k e$ remains entirely contained within $B_{1/2}(0)$, allowing us to safely write the domain of integration as $B_{1/2}(0)$ on both sides.

We now pass to the limit as $k \to \infty$ on both sides. On the left side, the weak convergence of $\delta_{e,h_k} u$ directly gives the integral against v_e. On the right side, because φ is smooth, the difference quotient $\delta_{e,-h_k}\varphi$ converges uniformly to the strong directional derivative $\partial_e \varphi$. Thus:

$$\int_{B_{1/2}} v_e(x)\varphi(x) \, dx = -\int_{B_{1/2}} u(x)\partial_e\varphi(x) \, dx.$$

This holds for all $\varphi \in C_c^\infty(B_{1/2})$, which is the exact definition of v_e being the weak directional derivative $\partial_e u$. Since $v_e \in L^p(B_{1/2})$ for every basis direction, we conclude $u \in W^{1,p}(B_{1/2})$.

Finally, by the weak lower semicontinuity of the L^p norm, we obtain the required bound:

$$\|Du\|_{L^p(B_{1/2})} \le \liminf_{k\to\infty} \|\delta_{e,h_k} u\|_{L^p(B_{1/2})} \le C.$$

■

Proof of Theorem 15.1.1. Let $u \in W^{1,2}(B_1)$ be a weak solution to $-\operatorname{div}(Df(Du)) = 0$. For a fixed unit vector $e \in \mathbb{S}^{d-1}$ and a small scalar $h > 0$, we consider

$$v_{e,h}(x) = \delta_{e,h} u(x) = \frac{u(x+he) - u(x)}{h}.$$

Since the Euler-Lagrange equation holds everywhere in the domain, it also holds for the shifted function $u(x+he)$. Subtracting the weak formulation of the original equation from the shifted equation and dividing by h yields:

$$-\operatorname{div}\left(\frac{Df(Du(x+he)) - Df(Du(x))}{h}\right) = 0.$$

Because the Lagrangian f is assumed to be smooth, we can apply the Fundamental Theorem of Calculus to write the difference of the gradients Df as an integral of the Hessian $D^2 f$ along the line segment connecting $Du(x)$ and $Du(x+he)$:

$$Df(Du(x+he)) - Df(Du(x)) = \\ \left[\int_0^1 D^2 f\big((1-t)Du(x) + tDu(x+he)\big)\,dt\right](Du(x+he) - Du(x)).$$

Let $A_{e,h}(x)$ denote the bracketed integral matrix above. Dividing both sides by h, we get exactly $A_{e,h}(x)Dv_{e,h}(x)$. Thus, the difference quotient $v_{e,h}$ weakly solves a linear elliptic equation:

$$-\operatorname{div}(A_{e,h}(x)Dv_{e,h}) = 0 \quad \text{in } B_{3/4}.$$

Because f is uniformly elliptic ($\lambda I \le D^2 f \le \Lambda I$), integrating $D^2 f$ over t preserves this property. Therefore, $A_{e,h}(x)$ is strictly uniformly elliptic with the exact same constants λ and Λ.

Now, because $v_{e,h}$ solves a linear uniformly elliptic equation, we can safely apply the standard Caccioppoli inequality (Theorem 14.2.1) to it between the balls $B_{1/2}$ and $B_{3/4}$:

$$\|Dv_{e,h}\|^2_{L^2(B_{1/2})} \le C(\lambda,\Lambda)\|v_{e,h}\|^2_{L^2(B_{3/4})}.$$

Following the approach of Lemma 15.1.3, we can bound $v_{e,h}$ by the norm of the gradient Du:

$$\|v_{e,h}\|^2_{L^2(B_{3/4})} = \left\|\frac{u(\cdot+he) - u(\cdot)}{h}\right\|^2_{L^2(B_{3/4})} \le \|Du\|^2_{L^2(B_1)}.$$

Substituting this into our Caccioppoli bound gives:

$$\|Dv_{e,h}\|^2_{L^2(B_{1/2})} \le C(\lambda,\Lambda)\|Du\|^2_{L^2(B_1)}.$$

Notice that the right-hand side is a constant completely independent of h. This means the difference quotients of the gradient, $Dv_{e,h} = \delta_{e,h}(Du)$, are uniformly bounded in $L^2(B_{1/2})$.

By Lemma 15.1.3, this uniform bound implies that the weak derivative of Du exists. Thus, $Du \in W^{1,2}(B_{1/2})$ (which implies $u \in W^{2,2}(B_{1/2})$). Additionally, we can pass the bound to the limit $h \to 0$ which yields the desired energy estimate:

$$\int_{B_{1/2}} |D^2u|^2\,dx \le \tilde{C}(\lambda,\Lambda)\int_{B_1} |Du|^2\,dx.$$

■

§15.2 Caccioppoli Inequality on Level Sets

To prove Theorem 15.1.2 (Hölder continuity for linear equations), we cannot rely on the standard Caccioppoli inequality. As we saw in the previous lecture, the hole-filling technique derived from the global Caccioppoli inequality fails to yield Hölder continuity in dimensions $d \ge 3$.

To bypass this dimensional barrier, De Giorgi introduced a profound shift in perspective: instead of bounding the energy of the entire function v, we must bound the energy of its super-level sets (the parts of the function that exceed a certain height k).

Definition 15.2.1 (Truncation)
For any constant $k \in \mathbb{R}$, we define the positive truncated function as:

$$(v-k)_+ = \max(v-k, 0) = (v-k) \vee 0.$$

By Stampacchia's Lemma (chain rule for Sobolev space), if $v \in W^{1,2}(\Omega)$, then $(v-k)_+ \in W^{1,2}(\Omega)$ and its weak gradient is given by:

$$D(v-k)_+ = \begin{cases} Dv & \text{almost everywhere on } \{x \in \Omega : v(x) > k\} \\ 0 & \text{almost everywhere on } \{x \in \Omega : v(x) \le k\} \end{cases}$$

This allows us to establish a localized Caccioppoli inequality specifically for subsolutions.

Theorem 15.2.2 (Level-Set Caccioppoli Inequality)
Let $A(x)$ be a measurable matrix satisfying the uniform ellipticity condition $\lambda I_d \le A(x) \le \Lambda I_d$. Let $v \in W^{1,2}(B_1)$ be a weak subsolution to the linear equation, meaning:

$$-\operatorname{div}(A(x)Dv) \le 0 \quad \text{in } B_1.$$

Then, for any $k \in \mathbb{R}$ and any concentric balls $B_r \subset B_R \Subset B_1$, the following energy estimate holds:

$$\int_{B_r} |D(v-k)_+|^2\,dx \le \frac{C(\lambda,\Lambda)}{(R-r)^2}\int_{B_R} (v-k)_+^2\,dx.$$

Proof. Choose a smooth cut-off function $\eta \in C_c^\infty(B_R)$ such that $0 \leq \eta \leq 1$, $\eta \equiv 1$ on B_r, and the gradient is bounded by $|D\eta| \leq \dfrac{2}{R-r}$.

We select our test function to target the region where v exceeds k:

$$\psi = \eta^2 (v-k)_+ \in W_0^{1,2}(B_R).$$

Because $\psi \geq 0$ and v is a subsolution, testing the weak formulation yields:

$$\int_{B_R} \langle A(x)Dv, D(\eta^2(v-k)_+)\rangle \, dx \leq 0.$$

Applying the product rule, $D\psi = 2\eta(v-k)_+ D\eta + \eta^2 D(v-k)_+$. Expanding the integral:

$$\int_{B_R} \eta^2 \langle A(x)Dv, D(v-k)_+\rangle \, dx + 2\int_{B_R} \eta(v-k)_+ \langle A(x)Dv, D\eta\rangle \, dx \leq 0.$$

At this stage, we exploit the structure of the truncation. Because $(v-k)_+$ and its gradient vanish exactly where $v \leq k$, the integrands are entirely supported on the set $A_{k,R} = \{x \in B_R : v(x) > k\}$. On this active set, we have $Dv = D(v-k)_+$. Thus, we can replace the original gradient with the truncated gradient everywhere:

$$\int_{B_R} \eta^2 \langle A(x)D(v-k)_+, D(v-k)_+\rangle \, dx \leq -2\int_{B_R} \eta(v-k)_+ \langle A(x)D(v-k)_+, D\eta\rangle \, dx.$$

From here, the algebraic machinery is identical to the standard Caccioppoli inequality. We apply the uniform ellipticity lower bound on the left side, and the Cauchy-Schwarz inequality with the spectral bound Λ on the right side:

$$\lambda \int_{B_R} \eta^2 |D(v-k)_+|^2 \, dx \leq 2\Lambda \int_{B_R} \big(\eta |D(v-k)_+|\big)\big((v-k)_+ |D\eta|\big) \, dx.$$

Applying Young's inequality with Peter-Paul scaling ($\epsilon = \dfrac{\lambda}{2\Lambda}$) to the right side allows us to split the product:

$$\lambda \int_{B_R} \eta^2 |D(v-k)_+|^2 \, dx \leq \frac{\lambda}{2}\int_{B_R} \eta^2 |D(v-k)_+|^2 \, dx + \frac{2\Lambda^2}{\lambda}\int_{B_R} (v-k)_+^2 |D\eta|^2 \, dx.$$

Absorbing the gradient term into the left-hand side gives:

$$\frac{\lambda}{2}\int_{B_R} \eta^2 |D(v-k)_+|^2 \, dx \leq \frac{2\Lambda^2}{\lambda}\int_{B_R} (v-k)_+^2 |D\eta|^2 \, dx. \tag{15.1}$$

Finally, we utilize the specific properties of our cut-off function η. Because $\eta \equiv 1$ on B_r, we can restrict the integration domain on the left side to B_r. On the right side, we substitute the bound $|D\eta| \leq \dfrac{2}{R-r}$:

$$\int_{B_r} |D(v-k)_+|^2 \, dx \leq \frac{16\Lambda^2}{\lambda^2}\frac{1}{(R-r)^2}\int_{B_R} (v-k)_+^2 \, dx.$$

Setting $C(\lambda, \Lambda) = \dfrac{16\Lambda^2}{\lambda^2}$ concludes the proof. ∎

16 Hilbert's 19th Problem - 3

§16.1 Symmetries and the $L^2 - L^\infty$ Bound

Before we proceed to the De Giorgi iteration, we establish several fundamental stability properties of solutions and subsolutions under elementary transformations. These symmetries allow us to assume, without loss of generality, that our domains are unit balls and our solutions are non-negative and properly scaled.

Recall that $v \in W^{1,2}$ is a subsolution to our linear equation if $-\operatorname{div}(A(x)Dv) \le 0$ in the weak sense, meaning:

$$\int \langle A(x)Dv, D\varphi\rangle \, dx \le 0 \quad \forall \varphi \in W_0^{1,2}, \varphi \ge 0.$$

Exercise 16.1.1

Assume $A(x)$ is uniformly elliptic ($\lambda \, \mathrm{Id} \le A(x) \le \Lambda \, \mathrm{Id}$). Prove the following symmetries:

1. **Translation, Scaling, and Multiplication:** If v is a subsolution (resp. supersolution) in $B_R(x_0)$, then for any $\alpha > 0$, the rescaled function $\tilde{v}(y) = \alpha v(x_0 + Ry)$ is a subsolution (resp. supersolution) in $B_1(0)$ with respect to the rescaled matrix $\tilde{A}(y) = A(x_0 + Ry)$.
2. **Sign Reversal:** v is a subsolution if and only if $-v$ is a supersolution.
3. **Composition (Subsolutions):** If v is a subsolution, then $f(v)$ is a subsolution for any function $f \in C^2(\mathbb{R})$ that is monotonically increasing ($f' \ge 0$) and convex ($f'' \ge 0$).
4. **Composition (Solutions):** If v is a exact solution, then $f(v)$ is a subsolution for any convex function $f \in C^2(\mathbb{R})$ (monotonicity is no longer required).

Solution.

1. Let $\varphi \in W_0^{1,2}(B_1(0))$ with $\varphi \ge 0$. We change variables $x = x_0 + Ry$. Note that $D_y\tilde{v}(y) = \alpha R Dv(x)$, $D_y\varphi(y) = RD_x\varphi(x)$, and $dy = R^{-d}dx$. Substituting these into the weak formulation for $\tilde{v}$ on $B_1(0)$ yields:

   $$\int_{B_1(0)} \tilde{A}(y)D_y\tilde{v}(y) \cdot D_y\varphi(y) \, dy = \alpha R^{2-d} \int_{B_R(x_0)} A(x)Dv(x) \cdot D_x\varphi(x) \, dx \le 0,$$

   confirming $\tilde{v}$ is a subsolution.
2. This follows trivially from the linearity of the integral and the gradient: $D(-v) = -Dv$. Multiplying the weak inequality by -1 flips the inequality sign, exchanging subsolutions for supersolutions.
3. To show $f(v)$ is a subsolution, we must show $\displaystyle\int AD(f(v)) \cdot D\varphi \le 0$ for any $\varphi \ge 0$. Using the chain rule, $D(f(v)) = f'(v)Dv$.

Because v is a subsolution, we can test its weak formulation with the specific non-negative test function $\psi = \varphi f'(v) \in W_0^{1,2}$. (Note: $f'(v) \geq 0$ by monotonicity, so $\psi \geq 0$). Applying the product rule gives $D\psi = f'(v)D\varphi + \varphi f''(v)Dv$. Plugging ψ into the subsolution inequality for v:

$$\int ADv \cdot \big(f'(v)D\varphi + \varphi f''(v)Dv\big)\, dx \leq 0.$$

Rearranging the terms:

$$\int f'(v)ADv \cdot D\varphi\, dx \leq -\int \varphi f''(v)ADv \cdot Dv\, dx.$$

The left side is exactly $\int AD(f(v)) \cdot D\varphi\, dx$. On the right side, $\varphi \geq 0$, $f''(v) \geq 0$ (by convexity), and $ADv \cdot Dv \geq \lambda |Dv|^2 \geq 0$ (by uniform ellipticity). Thus, the entire right side is ≤ 0, which completes the proof.

4. If v is an exact solution, then testing with $\psi = \varphi f'(v)$ gives an exact equality:

$$\int ADv \cdot \big(f'(v)D\varphi + \varphi f''(v)Dv\big)\, dx = 0.$$

Since the right-hand side of our rearranged inequality is strictly 0 initially, we no longer need the sign of ψ (and thus the sign of f') to guarantee the inequality direction. The convexity $f'' \geq 0$ alone guarantees the right side is negative after rearrangement, proving $f(v)$ is a subsolution.

■

§16.1.1 The Maximum Principle and the Necessity of the Local Bound

To understand why the De Giorgi iteration is necessary, consider a variational perspective. The weak solution v is a minimizer of the energy functional

$$\int \langle A(x)Dv, Dv\rangle\, dx \simeq \int |Dv|^2\, dx$$

over the space $g + W_0^{1,2}(B_1)$, where g is fixed boundary data.

Suppose the boundary data is globally bounded by a constant: $\sup_{\partial B_1} g \leq K$. Consider the truncated function $w = v \wedge K$. Since w matches v on the boundary, it is a valid competitor. Because the gradient of w vanishes wherever $v > K$, its total energy is strictly less than v unless $v \leq K$ everywhere. Since v is the minimizer, we immediately conclude that $\sup_{B_1} v \leq K$.

However, this classical Maximum Principle requires explicit knowledge that the boundary data is bounded (L^∞). If we only know that $v \in W^{1,2}_{\text{loc}}$, we have no boundary control. De Giorgi's monumental realization was that the Level-Set Caccioppoli inequality is sufficient to force the function to be locally bounded strictly in the interior, relying only on its L^2 mass.

Proposition 16.1.2 (Local $L^2 - L^\infty$ Estimate)
Let $v \in W^{1,2}(B_1)$ be a subsolution ($-\operatorname{div}(A(x)Dv) \le 0$ in B_1). Then v is locally bounded from above, and there exists a constant $C = C(d, \Lambda/\lambda) > 0$ such that:

$$\operatorname{ess\,sup}_{B_{1/2}} v \le C\|v_+\|_{L^2(B_1)}.$$

Corollary 16.1.3 (Oscillation Bound)
If v is a solution in $B_R(x_0)$, applying Proposition 16.1.2 to both v and $-v$ (which are both subsolutions) and rescaling yields the oscillation bound:

$$\operatorname{osc}_{B_{R/2}} v := \left(\sup_{B_{R/2}} v - \inf_{B_{R/2}} v\right) \le C(d, \Lambda/\lambda) \left(⨍_{B_R} v^2\,dx\right)^{1/2}.$$

Proof of Proposition 16.1.2. The proof relies on an infinite recurrence relation built from the Level-Set Caccioppoli inequality.

First, we can assume without loss of generality that $v \ge 0$ everywhere by replacing v with $v_+ = \max\{v, 0\}$. (By Exercise 16.1.1, v_+ is a subsolution because $f(s) = \max\{s, 0\}$ is convex and monotonically increasing).

Furthermore, by the scaling symmetry (Exercise 16.1.1), if we can prove the theorem for functions with a very small L^2 norm, we can prove it for any function. Specifically, it is enough to prove that there exists a small universal constant $\delta = \delta(d, \Lambda/\lambda) > 0$ such that:

$$\text{If } \int_{B_1} v^2\,dx \le \delta^2, \quad \text{then } v \le 1 \text{ a.e. in } B_{1/2}.$$

If we can prove this statement, we can simply apply it to the normalized function $\tilde{v} = \dfrac{\delta v}{\|v\|_{L^2(B_1)}}$ to recover the general bound.

To prove that $v \le 1$ on $B_{1/2}$, we define an infinite sequence of shrinking radii R_n and increasing height levels k_n that interpolate between the outer ball B_1 at height 0, and the inner ball $B_{1/2}$ at height 1.

$$R_n = \frac{1}{2} + \frac{1}{2^n} \quad \Longrightarrow \quad R_1 = 1, \quad R_\infty = \frac{1}{2}.$$
$$k_n = 1 - \frac{1}{2^{n-1}} \quad \Longrightarrow \quad k_1 = 0, \quad k_\infty = 1.$$

For each step n, we construct a smooth cut-off function $\eta_n \in C_c^\infty(B_{R_n})$ such that $\eta_n \equiv 1$ on the inner ball $B_{R_{n+1}}$. The gradient bound for this cut-off is strictly controlled by the width of the annulus between the successive balls:

$$|D\eta_n| \le \frac{2}{R_n - R_{n+1}} = \frac{2}{\frac{1}{2^n} - \frac{1}{2^{n+1}}} = 2^{n+2}.$$

Finally, we define the truncated sequence:

$$v_n = (v - k_n)_+.$$

Our goal is to show that the L^2 energy of v_n on the balls B_{R_n} collapses to 0 as $n \to \infty$.

Let us define the energy sequence as the L^2 norm of our truncated, localized functions:

$$a_n = \int_{B_1} (\eta_n v_n)^2 \, dx.$$

Our goal is to bound a_n in terms of a_{n-1} to create a recurrence relation. We will achieve this by combining four fundamental inequalities.

Step 1: Sobolev Embedding.
Since $\eta_n v_n \in W_0^{1,2}(B_1)$, we can apply the Sobolev inequality (Theorem 6.1.11 1.). There exists a constant $C_S > 0$ such that:

$$\left(\int_{B_1} (\eta_n v_n)^{2^*} \, dx \right)^{\frac{2}{2^*}} \leq C_S \int_{B_1} |D(\eta_n v_n)|^2 \, dx.$$

(Note: If $d = 2$, we replace 2^* with any finite $p > 2$ and the proof follows identically).

Step 2: Level-Set Caccioppoli and the Product Rule.
To apply the Sobolev embedding, we must bound the gradient of the localized function $\eta_n v_n$. Using the product rule $D(\eta_n v_n) = \eta_n D v_n + v_n D\eta_n$ and the inequality from convexity: $|a+b|^2 \leq 2|a|^2 + 2|b|^2$, we expand the energy:

$$\int_{B_1} |D(\eta_n v_n)|^2 \, dx \leq 2 \int_{B_{R_n}} \eta_n^2 |D v_n|^2 \, dx + 2 \int_{B_{R_n}} v_n^2 |D\eta_n|^2 \, dx.$$

For the first term, we recall the intermediate step in the proof of the Level-Set Caccioppoli Inequality (Theorem 15.2.2, inequality (15.1)), which directly bounds the localized gradient of the subsolution by the gradient of the cut-off function:

$$\int_{B_{R_n}} \eta_n^2 |D v_n|^2 \, dx \leq \frac{4\Lambda^2}{\lambda^2} \int_{B_{R_n}} v_n^2 |D\eta_n|^2 \, dx.$$

Substituting this back, both terms are now entirely bounded by the integral of $v_n^2 |D\eta_n|^2$. We use our explicit gradient bound $|D\eta_n| \leq \dfrac{2}{R_n - R_{n+1}} = 2^{n+2}$, which implies $|D\eta_n|^2 \leq 16 \cdot 4^n$. Absorbing the 16, the 2, and the ellipticity ratio Λ^2/λ^2 into a single generic constant $C(\lambda, \Lambda)$, we obtain:

$$\int_{B_1} |D(\eta_n v_n)|^2 \, dx \leq C(\lambda, \Lambda) \cdot 4^n \int_{B_{R_n}} v_n^2 \, dx.$$

Notice that on the domain B_{R_n}, the previous cut-off function satisfies $\eta_{n-1} \equiv 1$. Furthermore, since $k_n > k_{n-1}$, we strictly have $v_n \leq v_{n-1}$. Thus, we can bound this integral entirely by the previous sequence term:

$$\int_{B_{R_n}} v_n^2 \, dx \leq \int_{B_{R_n}} (\eta_{n-1} v_{n-1})^2 \, dx = a_{n-1}.$$

Combining Step 1 and Step 2 yields:

$$\left(\int_{B_1} (\eta_n v_n)^{2^*} \, dx \right)^{\frac{2}{2^*}} \leq C \cdot 4^n a_{n-1}. \tag{16.1}$$

Step 3: Hölder Interpolation.
We want to bound a_n, which is an L^2 norm, using the L^{2^*} norm we just bounded. We use Hölder's inequality over the active support set $A_n = \{x \in B_1 : \eta_n v_n > 0\}$:

$$a_n = \int_{A_n} (\eta_n v_n)^2 \cdot 1 \, dx \leq \left(\int_{B_1} (\eta_n v_n)^{2^*} \, dx \right)^{\frac{2}{2^*}} |A_n|^{1-\frac{2}{2^*}}.$$

Let $\varepsilon = 1 - \frac{2}{2^*} = \frac{2}{d}$ (which is strictly positive). Substituting (16.1) into this interpolation gives:

$$a_n \leq C \cdot 4^n a_{n-1} |A_n|^{\varepsilon}. \tag{16.2}$$

Step 4: Bounding the Support via Markov's Inequality.
We must control the measure of the support $|A_n|$. If $x \in A_n$, then $v(x) > k_n$. Observe the gap between successive levels: $k_n - k_{n-1} = \frac{1}{2^{n-1}}$. Therefore, on the set A_n, we have:

$$v_{n-1}(x) = (v(x) - k_{n-1})_+ > k_n - k_{n-1} = \frac{1}{2^{n-1}}.$$

Since $\eta_{n-1} \equiv 1$ on B_{R_n} (which contains A_n), we have $\eta_{n-1} v_{n-1} > \frac{1}{2^{n-1}}$. By Markov's (Chebyshev's) inequality:

$$|A_n| \leq \left| \left\{ x : \eta_{n-1} v_{n-1} > \frac{1}{2^{n-1}} \right\} \right| \leq \frac{\int_{B_1} (\eta_{n-1} v_{n-1})^2 \, dx}{(2^{-(n-1)})^2} = 4^{n-1} a_{n-1}.$$

Step 5: The Recurrence Relation.
Substitute the bound for $|A_n|$ back into (16.2):

$$a_n \leq C \cdot 4^n a_{n-1} \left(4^{n-1} a_{n-1}\right)^{\varepsilon} \leq C \cdot (4^{1+\varepsilon})^n a_{n-1}^{1+\varepsilon}.$$

By absorbing the constants into a generic base $C_0 = 4^{1+\varepsilon} > 1$, we arrive at the fundamental De Giorgi nonlinear recurrence relation:

$$a_n \leq C \cdot C_0^n a_{n-1}^{1+\varepsilon}.$$

From Exercise 16.1.4, we can establish that if $a_1 = \int_{B_1} (\eta_1 v_1)^2 \leq \bar{\delta}^2$, then $a_n \to 0$.

Notice that $v_1 = (v - k_1)_+ = v_+$, since $0 \leq \eta_1 \leq 1$ in B_1, we have

$$a_1 \leq \int_{B_1} (v_+)^2.$$

So, if we have $\int_{B_1} v_+^2 \leq \bar{\delta}^2$, then our desired result holds.

As $n \to \infty$, the radii $R_n \to 1/2$ and the heights $k_n \to 1$. By Fatou's Lemma (or the Monotone Convergence Theorem), the limit of the energies bounds the energy of the limit:

$$\int_{B_{1/2}} (v-1)_+^2 \, dx \leq \lim_{n \to \infty} \int_{B_{R_n}} (\eta_n (v - k_n)_+)^2 \, dx = \lim_{n \to \infty} a_n = 0.$$

Because the integral of the non-negative function $(v-1)_+^2$ over $B_{1/2}$ is strictly zero, the integrand must be zero almost everywhere.

$$(v-1)_+ = 0 \text{ a.e. in } B_{1/2} \implies v \leq 1 \text{ a.e. in } B_{1/2}.$$

This completes the proof of Proposition 16.1.2, cementing the critical local L^∞ bound required for the De Giorgi iteration. ■

Exercise 16.1.4 (Fast Convergence Lemma)
Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of non-negative real numbers satisfying the nonlinear recurrence:

$$a_n \leq C \cdot C_0^n a_{n-1}^{1+\varepsilon} \quad \forall n \geq 2$$

where $C_0 > 1$ and C, $\varepsilon > 0$. Prove that there exists a strictly positive constant $\bar{\delta} = \bar{\delta}(C, C_0, \varepsilon) > 0$ such that if $a_1 \leq \bar{\delta}^2$, then $\lim_{n\to\infty} a_n = 0$.

Solution. We will prove by induction that if a_1 is sufficiently small, the sequence decays exponentially: $a_n \leq \frac{a_1}{K^n}$ for some constant $K > 1$.

Assume $a_{n-1} \leq \frac{a_1}{K^{n-1}}$. Substituting this into the recurrence relation yields:

$$a_n \leq C \cdot C_0^n \left(\frac{a_1}{K^{n-1}}\right)^{1+\varepsilon} = C \cdot C_0^n \frac{a_1^{1+\varepsilon}}{K^{(n-1)(1+\varepsilon)}}.$$

We want to naïvely enforce this upper bound is less than or equal to $\frac{a_1}{K^n}$. Thus, we require:

$$C \cdot C_0^n \frac{a_1^{1+\varepsilon}}{K^{(n-1)(1+\varepsilon)}} \leq \frac{a_1}{K^n}.$$

Dividing by a_1 and multiplying by $K^{(n-1)(1+\varepsilon)}$ gives:

$$C \cdot C_0^n a_1^{\varepsilon} \leq K^{(n-1)(1+\varepsilon)-n} = K^{n\varepsilon-1-\varepsilon}.$$

To manage the exponential growth of n, we choose K such that $K^{\varepsilon} = C_0 > 1$. This perfectly cancels the n-dependent terms ($C_0^n = K^{n\varepsilon}$):

$$C \cdot K^{n\varepsilon} a_1^{\varepsilon} \leq K^{n\varepsilon} K^{-1-\varepsilon} \implies C \cdot a_1^{\varepsilon} \leq K^{-1-\varepsilon}.$$

Therefore, the induction step holds for all n if we simply require the initial value to satisfy:

$$a_1 \leq \frac{K^{-\frac{1+\varepsilon}{\varepsilon}}}{C^{\frac{1}{\varepsilon}}} = \frac{C_0^{-\frac{1+\varepsilon}{\varepsilon^2}}}{C^{\frac{1}{\varepsilon}}} := \bar{\delta}^2.$$

Since $K > C_0^{\frac{1}{\varepsilon}} > 1$, the sequence $a_n \leq \frac{a_1}{K^n}$ converges to 0 as $n \to \infty$. ■

Remark 16.1.5 (Historical Literature). The techniques developed here trace back to the foundational works in regularity theory. For the original proof, see De Giorgi's seminal 1957 paper [DG57]. For modern, highly streamlined treatments applying these methods to fluid dynamics and general PDEs, excellent resources include the notes by Caffarelli and Vasseur [CV10], and the comprehensive text by Ambrosio, Carlotto, and Massaccessi [ACM18].

§16.2 Oscillation Lemma - 1

To finalize the proof of De Giorgi's Theorem (Theorem 15.1.2), it is sufficient to prove that the oscillation of the solution strictly drops when passing to a smaller inner ball. We state this as the following central claim.

Claim 16.2.1 — There exists a constant $\eta = \eta(d, \Lambda/\lambda) \in (0,1)$ such that if $v \in W^{1,2}(B_1)$ is a weak solution in B_1 satisfying the normalization:

$$\operatorname{ess\,sup}_{B_1} v = 1 \quad \text{and} \quad \operatorname{ess\,inf}_{B_1} v = -1,$$

then the oscillation strictly decreases in $B_{1/4}$. Specifically, either:

$$\sup_{B_{1/4}} v \le 1 - \eta \quad \text{or} \quad \inf_{B_{1/4}} v \ge -1 + \eta.$$

§16.2.1 Sufficiency of the Claim for Hölder Continuity

Before proving the claim, we must demonstrate why this specific drop in oscillation rigorously implies C^α regularity.

Assume Claim 16.2.1 holds. Because the total oscillation on B_1 is $1 - (-1) = 2$, the claim guarantees that the new oscillation on $B_{1/4}$ is bounded by:

$$\operatorname{osc}_{B_{1/4}} v = \sup_{B_{1/4}} v - \inf_{B_{1/4}} v \le 2 - \eta = \left(1 - \frac{\eta}{2}\right) \operatorname{osc}_{B_1} v.$$

This provides a geometric decay factor $\theta = (1 - \eta/2) < 1$. To iterate this decay, for any arbitrary ball $B_R(x_0)$ where v is a solution, we define the affine rescaling:

$$\tilde{v}(x) = \frac{2\left[v(x_0 + Rx) - \frac{1}{2}\left(\sup_{B_R(x_0)} v + \inf_{B_R(x_0)} v\right)\right]}{\sup_{B_R(x_0)} v - \inf_{B_R(x_0)} v}.$$

By the symmetry properties established in Exercise 16.1.1, $\tilde{v}$ is a valid solution in $B_1(0)$. Furthermore, the translation and scaling guarantee that $\sup\limits_{B_1} \tilde{v} = 1$ and $\inf\limits_{B_1} \tilde{v} = -1$. Applying the claim to $\tilde{v}$ yields $\operatorname{osc}_{B_{1/4}} \tilde{v} \le (1 - \eta/2) \operatorname{osc}_{B_1} \tilde{v}$, which pulls back directly to the original function v:

$$\operatorname{osc}_{B_{R/4}(x_0)} v \le \left(1 - \frac{\eta}{2}\right) \operatorname{osc}_{B_R(x_0)} v.$$

From Discrete Decay to Continuous Oscillation

Iterating this inequality k times over shrinking radii $R_k = 4^{-k}$ centered at an arbitrary point x_0 yields:

$$\operatorname{osc}_{B_{4^{-k}}(x_0)} v \le \left(1 - \frac{\eta}{2}\right)^k \operatorname{osc}_{B_1(x_0)} v.$$

We can express this discrete decay factor as a spatial power law. Define the Hölder exponent $\alpha = -\log_4(1 - \eta/2) > 0$. By definition, this means $(1 - \eta/2)^k = (4^{-k})^\alpha$.

Now, let r be any arbitrary continuous radius such that $0 < r \le 1/4$. We can trap r between two successive dyadic radii by finding an integer $k \ge 1$ such that $4^{-(k+1)} < r \le 4^{-k}$. Because the oscillation of a function monotonically decreases as the domain shrinks, we have:

$$\operatorname{osc}_{B_r(x_0)} v \le \operatorname{osc}_{B_{4^{-k}}(x_0)} v \le (4^{-k})^\alpha \operatorname{osc}_{B_1(x_0)} v.$$

Using the lower bound $4^{-k} = 4 \cdot 4^{-(k+1)} < 4r$, we obtain the continuous oscillation bound:

$$\operatorname{osc}_{B_r(x_0)} v \le (4r)^\alpha \operatorname{osc}_{B_1(x_0)} v = Cr^\alpha.$$

Remark 16.2.2 (From Oscillation to Pointwise Hölder Continuity). To see why this local oscillation bound forces the standard pointwise Hölder condition everywhere in $B_{1/2}$, consider any two points $x, y \in B_{1/2}$.

- **Close points ($|x-y| \le 1/4$):** Let $r = |x-y|$. Because $x \in B_{1/2}$, its distance to ∂B_1 is at least $1/2$, so the "sliding microscope" ball $B_{1/4}(x)$ safely fits inside the domain. Both x and y lie in $B_r(x)$, meaning their difference is bounded by the oscillation on that ball:

$$|v(x) - v(y)| \le \operatorname{osc}_{B_r(x)} v \le Cr^\alpha = C|x-y|^\alpha.$$

- **Far points ($|x-y| > 1/4$):** The $L^2 - L^\infty$ estimate guarantees v is bounded. Let $M = \sup_{B_{1/2}} |v|$. The maximum possible difference is $2M$. Since $4|x-y| > 1$, we trivially have $1 < (4|x-y|)^\alpha$, meaning:

$$|v(x) - v(y)| \le 2M < 2M4^\alpha |x-y|^\alpha.$$

Taking the maximum of the constants from both cases rigorously proves $v \in C^\alpha(B_{1/2})$, officially completing the proof of Theorem 15.1.2.

Furthermore, we can explicitly bound the final Hölder constant by the initial energy of the system. Recall from our continuous decay formula that the Hölder constant C depends directly on the initial oscillation and the supremum of v. By applying the $L^2 - L^\infty$ estimate (Proposition 16.1.2) and the Oscillation Bound Corollary 16.1.3, we can control both $\sup v$ and $\operatorname{osc} v$ entirely by the L^2 average of v.

Restoring the arbitrary radius R via standard scaling symmetry (Exercise 16.1.1), we obtain the complete, scale-invariant form of the De Giorgi regularity estimate:

$$\|v\|_{L^\infty(B_{R/2}(x_0))} + R^\alpha [v]_{C^{0,\alpha}(B_{R/2}(x_0))} \le C(d, \lambda, \Lambda) \left(\fint_{B_R(x_0)} v^2 \, dx \right)^{\frac{1}{2}}.$$

This successfully recovers the exact quantitative bound stated in the De Giorgi-Nash-Moser theorem (Theorem 14.1.7).

§16.2.2 Bridging the Gap: The Intermediate Measure Lemma

To prove Claim 16.2.1, we recall the specific consequence of the $L^2 - L^\infty$ bound (Fast Convergence Lemma) established in the previous section.

Remark 16.2.3 (The "Tiny Measure" Condition). From Proposition 16.1.2 and its recurrence relation, we know there exists a strictly positive constant $\delta = \delta(d, \Lambda/\lambda)$ such that if v is a subsolution, $v \le 1$ in B_1, and the measure of its positive part is sufficiently tiny:

$$|\{v > 0\} \cap B_{1/2}| \le \delta,$$

then the supremum strictly drops in the inner ball: $\sup_{B_{1/4}} v \le 1/2$.

Derivation of δ: We apply the scale-invariant $L^2 - L^\infty$ estimate (Proposition 16.1.2) starting from the ball $B_{1/2}$ instead of B_1. Because $v \le 1$, its energy on this domain is trivially bounded by its active support area:

$$\|v_+\|^2_{L^2(B_{1/2})} \le \int_{\{v>0\} \cap B_{1/2}} 1 \, dx = |\{v > 0\} \cap B_{1/2}|.$$

Substituting this into the estimate gives $\sup_{B_{1/4}} v \le C\sqrt{|\{v>0\}\cap B_{1/2}|}$. Forcing this upper bound to equal $1/2$ explicitly dictates our choice of threshold: $\delta = \dfrac{1}{4C^2}$.

In practice, we cannot guarantee that the measure of the positive part is less than δ. We only know that since v oscillates between -1 and 1, at least half of the domain must be ≤ 0, or at least half must be ≥ 0. We need an intermediate lemma to improve the assumption from a "tiny measure" δ to a "half measure" $|B_{1/2}|/2$.

Lemma 16.2.4 (Measure-to-Pointwise Lemma)
For every $\varepsilon > 0$, there exists $\eta = \eta(\varepsilon, d, \Lambda/\lambda) \in (0,1)$ such that if $v \in W^{1,2}(B_1)$ is a subsolution to $-\operatorname{div}(A(x)Dv) \le 0$ satisfying:

1. $v \le 1$ in B_1
2. $|\{v \le 0\}\cap B_{1/2}| \ge \dfrac{1}{2}|B_{1/2}|$

Then the measure of the set where v is close to its maximum is tiny:

$$|\{v \ge 1-\eta\}\cap B_{1/2}| \le \varepsilon.$$

The proof of this lemma requires the Measure/Poincaré inequality (which we will cover in the next lecture). For now, we will assume Lemma 16.2.4 holds and use it to formally construct the proof of the main claim.

Proof of Claim 16.2.1. Let $\delta = \delta(d, \Lambda/\lambda)$ be the universal constant from Remark 16.2.3. We invoke Lemma 16.2.4 by strictly choosing our tolerance to be this exact constant: $\varepsilon = \delta$. This provides us with a specific height parameter $\eta = \eta(\delta, d, \Lambda/\lambda) > 0$.

Since $\sup_{B_1} v = 1$ and $\inf_{B_1} v = -1$, the domain $B_{1/2}$ is partitioned by the sign of v. We certainly have either:

(a) $|\{v \le 0\}\cap B_{1/2}| \ge \dfrac{1}{2}|B_{1/2}|$, or

(b) $|\{v \ge 0\}\cap B_{1/2}| \ge \dfrac{1}{2}|B_{1/2}|$.

Case (a): Assume the function is mostly negative or zero. By Lemma 16.2.4, the measure of the set where v is near its peak is strictly controlled:

$$|\{v \ge 1-\eta\}\cap B_{1/2}| \le \varepsilon = \delta.$$

We now construct a specialized function to feed into Remark 16.2.3. Let:

$$\tilde{w} = \frac{(v-(1-\eta))_+}{\eta}.$$

Because v is a subsolution, the truncation and positive scaling guarantee that $\tilde{w}$ is also a subsolution. Furthermore:

- Since $v \le 1$, we have $(v-(1-\eta))_+ \le \eta$, meaning $\tilde{w} \le 1$ in B_1.
- The set where $\tilde{w} > 0$ is exactly the set where $v > 1-\eta$. We just established that the measure of this set in $B_{1/2}$ is bounded by δ.

Thus, $\tilde{w}$ perfectly satisfies all conditions of Remark 16.2.3. We immediately conclude that:

$$\sup_{B_{1/4}} \tilde{w} \leq \frac{1}{2}.$$

Unpacking the definition of $\tilde{w}$ on the ball $B_{1/4}$:

$$\frac{v-(1-\eta)}{\eta} \leq \frac{1}{2} \implies v \leq (1-\eta) + \frac{\eta}{2} = 1 - \frac{\eta}{2}.$$

By setting our final claim constant to $\eta' = \eta/2$, we achieve $\sup_{B_{1/4}} v \leq 1 - \eta'$.

Case (b): If the function is mostly positive, we exploit the sign symmetry (Exercise 16.1.1). We define $\hat{v} = -v$, which is still a subsolution satisfying $\hat{v} \leq 1$ and $|\{\hat{v} \leq 0\} \cap B_{1/2}| \geq \frac{1}{2}|B_{1/2}|$. Applying the exact same machinery to $\hat{v}$ yields $\sup_{B_{1/4}} \hat{v} \leq 1 - \eta/2$. Since $\sup(-\hat{v}) = -\inf(\hat{v})$, this is equivalent to $\inf_{B_{1/4}} v \geq -1 + \eta/2$.

In either case, the oscillation strictly drops, proving the claim. ■

17 Hilbert's 19th Problem - 4

§17.1 Oscillation Lemma - 2: The Measure/Poincaré Inequality

In the previous lecture, the entire proof of Hölder continuity was reduced to a single missing link: the Measure-to-Pointwise Lemma (Lemma 16.2.4). To rigorously prove that lemma, we must first establish a variation of the Poincaré inequality that controls the L^1 norm of a function based on the size of its zero-set.

Remark 17.1.1 (Integration Notation). Throughout this section, we will occasionally adopt a shorthand notation commonly found in physics for iterated integrals. Specifically, we will write $\int dx \int dy\, f(x,y)$ to represent the double integral $\iint f(x,y)\,dy\,dx$.

Lemma 17.1.2 (Measure/Poincaré Inequality)
Let $v \in W^{1,1}(B_R)$. Assume there is a subset of the ball where v vanishes, and its measure is bounded strictly from below:

$$|\{v=0\} \cap B_R| \geq \nu |B_R| \quad \text{for some fraction } \nu > 0.$$

Then, there exists a dimensional constant $C(d) > 0$ such that:

$$\int_{B_R} |v|\,dx \leq \frac{C(d)}{\nu} R \int_{B_R} |Dv|\,dx.$$

Proof. By a standard scaling argument, we can assume without loss of generality that $R = 1$. Furthermore, by density and smooth extension, we can assume $v \in C^1(B_1)$.

Let $y \in \{v = 0\}$ and $x \in B_1$. By the Fundamental Theorem of Calculus along the line segment connecting x and y:

$$|v(x)| = |v(x) - v(y)| = \left| \int_0^{|y-x|} \left\langle Du\left(x + t\frac{y-x}{|y-x|}\right), \frac{y-x}{|y-x|} \right\rangle dt \right|$$

Because the directional vector has length 1, we can bound this by the integral of the gradient's magnitude:

$$|v(x)| \leq \int_0^{|y-x|} \left| Du\left(x + t\frac{y-x}{|y-x|}\right) \right| dt.$$

We now integrate this inequality. We integrate x over the entire ball B_1, and y strictly over the zero-set $\{v = 0\}$:

$$|\{v=0\}| \int_{B_1} |v(x)|\,dx \leq \int_{B_1} dx \int_{\{v=0\}} dy \int_0^{|y-x|} dt \left| Du\left(x + t\frac{y-x}{|y-x|}\right) \right|.$$

To resolve the right-hand side, we perform a change of variables to spherical coordinates centered at x. Let $y = x + \rho\omega$, where $\omega = \dfrac{y-x}{|y-x|} \in \mathbb{S}^{d-1}$ and $\rho = |y - x| \in (0, 2)$. The volume element becomes $dy = \rho^{d-1} d\rho\, d\omega$. (We use same technique as we did to solve Problem 2 on Exercise 17.1.3)

Enlarging the integration domain for y to the full ball B_1 allows us to write:

$$\begin{aligned}\text{RHS} &\le \int_{B_1} dx \int_0^2 dt \left[\iint_{\mathbb{S}^{d-1}\times(0,2)} \rho^{d-1} d\rho\, d\omega\, |Du(x+t\omega)| \right] \\ &\le \int_{B_1} dx \int_0^2 dt \int_0^2 d\rho \cdot \rho^{d-1} \int_{\mathbb{S}^{d-1}} d\omega \frac{t^{d-1}|Du(x+t\omega)|}{t^{d-1}}.\end{aligned}$$

Notice that $z = x + t\omega$ is simply a polar change of variables in reverse, where $t = |z - x|$. Thus, $t^{d-1} dt\, d\omega = dz$. Absorbing the ρ integral into a dimensional constant $C(d)$, we get:

$$\text{RHS} \le C(d) \int_{B_1} dx \int_{B_1} dz \, \frac{|Du(z)|}{|x-z|^{d-1}}.$$

Applying Fubini's Theorem to swap the integrals:

$$\text{RHS} \le C(d) \left(\int_{B_1} |Du(z)|\, dz \right) \cdot \sup_{w \in B_1} \int_{B_1} \frac{dx}{|x-w|^{d-1}}.$$

The supremum of this Riesz potential is bounded by a dimensional constant $C^(d)$. Dividing the original left-hand side by $|\{v = 0\}| \ge \nu |B_1|$ completes the proof. ■

Exercise 17.1.3

Prove the following three inequalities utilizing the Riesz Transformation $I_1(f)(x) = \int \frac{f(z)}{|x-z|^{d-1}}\, dz$:

1. For all $u \in C_c^1(\mathbb{R}^d)$, $|u(x)| \le C(d) \int \frac{|Du(z)|}{|x-z|^{d-1}}\, dz$.
2. For all $u \in C^1(B_1)$, $\left| u(x) - \fint_{B_1} u \right| \le C(d) \int_{B_1} \frac{|Du(z)|}{|x-z|^{d-1}}\, dz$.
3. For all $u \in C^1(B_1)$, $|u(x) - u(y)| \le C(d) \left[\int_{B_1} \frac{|Du(z)|}{|x-z|^{d-1}}\, dz + \int_{B_1} \frac{|Du(z)|}{|y-z|^{d-1}}\, dz \right]$.

Solution.

1. Since $u \in C_c^1(\mathbb{R}^d)$, it vanishes at infinity. By the Fundamental Theorem of Calculus along a ray starting at x in direction $\omega \in \mathbb{S}^{d-1}$:

$$|u(x)| = \left| - \int_0^\infty \frac{d}{dt} u(x+t\omega)\, dt \right| \le \int_0^\infty |Du(x+t\omega)|\, dt.$$

Integrating both sides over the unit sphere $\mathbb{S}^{d-1}$ with respect to $d\omega$:

$$\int_{\mathbb{S}^{d-1}} |u(x)|\, d\omega \le \int_{\mathbb{S}^{d-1}} \int_0^\infty |Du(x+t\omega)|\, dt\, d\omega.$$

The left side evaluates to $d\omega_d |u(x)|$, where $d\omega_d$ is the surface area of $\mathbb{S}^{d-1}$. For the right side, we apply the change of variables $z = x + t\omega$. By multiplying and dividing by t^{d-1}, we reconstruct the Cartesian volume element $dz = t^{d-1} dt\, d\omega$:

$$d\omega_d |u(x)| \le \int_{\mathbb{S}^{d-1}} \int_0^\infty \frac{|Du(x+t\omega)|}{t^{d-1}} t^{d-1} dt\, d\omega = \int_{\mathbb{R}^d} \frac{|Du(z)|}{|x-z|^{d-1}}\, dz.$$

Dividing by $d\omega_d$ completes the proof.

2. Expressing the difference via the average and applying the FTC yields the initial bound:
$$\left|u(x) - \fint_{B_1} u\right| \leq \frac{1}{\omega_d}\int_{B_1} dy \int_0^{|x-y|} dt \left|Du\left(x + t\frac{y-x}{|y-x|}\right)\right|.$$
We immediately apply a spherical change of variables centered at x. Let $y = x + \rho\omega$, where $\rho \in (0,2)$ and $\omega \in \mathbb{S}^{d-1}$, yielding $dy = \rho^{d-1}d\rho\, d\omega$. Substituting this transforms the integral:
$$\text{LHS} \leq \frac{1}{\omega_d}\int_{\mathbb{S}^{d-1}} d\omega \int_0^2 d\rho\, \rho^{d-1}\int_0^\rho dt\, \big|Du(x+t\omega)\big|.$$
Applying Fubini's Theorem to evaluate the $d\rho$ integral first ($\int_t^2 \rho^{d-1}d\rho \leq \frac{2^d}{d}$):
$$\text{LHS} \leq \frac{2^d}{d\omega_d}\int_{\mathbb{S}^{d-1}} d\omega \int_0^2 dt\, \big|Du(x+t\omega)\big|.$$
Using the same substitution $z = x + t\omega$ and $dz = t^{d-1}dt\, d\omega$ as before, we arrive at the final bound with $C(d) = \frac{2^d}{d\omega_d}$:
$$\text{LHS} \leq \frac{2^d}{d\omega_d}\int_{\mathbb{S}^{d-1}}\int_0^2 \frac{\big|Du(x+t\omega)\big|}{t^{d-1}} t^{d-1}dt\, d\omega = C(d)\int_{B_1} \frac{|Du(z)|}{|x-z|^{d-1}}\, dz.$$

3. This is an immediate corollary of the second inequality. By the triangle inequality, we insert the average $\fint_{B_1} u$ as a stepping stone:
$$|u(x)-u(y)| = \left|\left(u(x) - \fint_{B_1} u\right) - \left(u(y) - \fint_{B_1} u\right)\right| \leq \left|u(x) - \fint_{B_1} u\right| + \left|u(y) - \fint_{B_1} u\right|.$$
Applying the pointwise bound from Exercise 2 directly to each term yields:
$$|u(x) - u(y)| \leq C(d)\left[\int_{B_1} \frac{|Du(z)|}{|x-z|^{d-1}}\, dz + \int_{B_1} \frac{|Du(z)|}{|y-z|^{d-1}}\, dz\right].$$

■

§17.1.1 Energy Bounds on Intermediate Level Sets

Using the Measure/Poincaré inequality, we can directly relate the gradient energy between two height levels to the measure of the super-level sets.

Corollary 17.1.4

Let $u \in W^{1,2}(B_R)$ and assume $|\{u \leq 0\} \cap B_R| \geq \nu|B_R|$. Then for any two height levels $0 < \ell < h$:
$$(h-\ell)|\{u \geq h\}| \cdot \frac{|\{u \leq \ell\}|}{|B_R|} \leq C(d)R|\{\ell < u < h\}|^{1/2}\left(\int_{\{\ell<u<h\}} |Du|^2\, dx\right)^{1/2}.$$

Proof. Define the truncated function $v = \min\{(u-\ell)_+, h-\ell\}$. Notice that $v = 0$ exactly where $u \leq \ell$. Since $|\{u \leq \ell\}| \geq |\{u \leq 0\}| \geq \nu |B_R|$, v satisfies the conditions of Lemma 17.1.2.

Applying the Measure/Poincaré inequality to v yields:

$$\int_{B_R} |v|\, dx \leq \frac{C(d)R}{|\{u \leq \ell\}|/|B_R|} \int_{B_R} |Dv|\, dx.$$

On the left side, v reaches its maximum constant value of $(h-\ell)$ whenever $u \geq h$. Thus:

$$\int_{B_R} |v|\, dx \geq \int_{\{u \geq h\}} (h-\ell)\, dx = (h-\ell)|\{u \geq h\}|.$$

On the right side, the gradient Dv is non-zero only strictly between the two levels $\ell < u < h$, where it perfectly matches Du. Applying the Cauchy-Schwarz inequality to this specific integration domain:

$$\int_{B_R} |Dv|\, dx = \int_{\{\ell<u<h\}} |Du| \cdot 1\, dx \leq \left(\int_{\{\ell<u<h\}} |Du|^2\, dx \right)^{1/2} |\{\ell < u < h\}|^{1/2}.$$

Combining these bounds and rearranging the measure ratio $|\{u \leq \ell\}|/|B_R|$ to the left side proves the corollary. ■

§17.2 Closing the Proof: The Measure-to-Pointwise Lemma

We now possess all the required machinery to rigorously prove the aforementioned Lemma 16.2.4, officially resolving Hilbert's 19th Problem.

Proof of Lemma 16.2.4. We define an infinite sequence of height levels $k_n \to 1$:

$$k_n = 1 - 2^{-n} \quad \text{for } n = 0, 1, 2, \ldots$$

Notice that $k_0 = 0$, $k_{n+1} - k_n = 2^{-(n+1)}$, and $1 - k_n = 2^{-n}$.

We apply Corollary 17.1.4 to the function v on the ball $B_{1/2}$, using the height levels $\ell = k_n$ and $h = k_{n+1}$. By assumption, $|\{v \leq 0\}| \geq \frac{1}{2}|B_{1/2}|$, so $|\{v \leq k_n\}|/|B_{1/2}| \geq 1/2$. The corollary gives:

$$(k_{n+1} - k_n) \cdot |\{v \geq k_{n+1}\}| \cdot \frac{1}{2} \leq C(d)|\{k_n < v \leq k_{n+1}\}|^{1/2} \left(\int_{B_{1/2}} |D(v-k_n)_+|^2\, dx \right)^{1/2}.$$

We now control the gradient energy term using the Level-Set Caccioppoli Inequality (Theorem 15.2.2). Because $v \leq 1$, we have $(v-k_n)_+ \leq 1 - k_n$. Therefore:

$$\int_{B_{1/2}} |D(v-k_n)_+|^2\, dx \leq C(d, \lambda, \Lambda) \int_{B_1} (v-k_n)_+^2\, dx \leq \tilde{C}(d, \lambda, \Lambda)(1-k_n)^2.$$

Substituting this energy bound back into our inequality and squaring both sides:

$$\frac{(k_{n+1}-k_n)^2}{4} |\{v \geq k_{n+1}\}|^2 \leq C|\{k_n < v \leq k_{n+1}\}|(1-k_n)^2.$$

Isolating the super-level set measure:

$$|\{v \geq k_{n+1}\}|^2 \leq C|\{k_n < v \leq k_{n+1}\}| \frac{(1-k_n)^2}{4(k_{n+1}-k_n)^2}.$$

By the properties of our chosen sequence, the ratio $\frac{1-k_n}{k_{n+1}-k_n} = \frac{2^{-n}}{2^{-(n+1)}} = 2$. Thus, the square of the ratio is uniformly bounded by 16, which is absorbed into the constant:

$$|\{v \geq k_{n+1}\}|^2 \leq C(d,\lambda,\Lambda)|\{k_n < v \leq k_{n+1}\}|.$$

We now sum this inequality over $n = 0$ to $N-1$. Because the sets $E_n = \{k_n < v \leq k_{n+1}\}$ are mutually disjoint and all contained within $B_{1/2}$, the sum of their measures cannot exceed the volume of the ball: $\sum |E_n| \leq |B_{1/2}|$.

Furthermore, since the sequence of super-level sets $\{v \geq k_{n+1}\}$ is strictly shrinking, every term in the sum on the left is bounded from below by the smallest set, $\{v \geq k_N\}$. Thus:

$$N|\{v \geq k_N\}|^2 \leq \sum_{n=0}^{N-1} |\{v \geq k_{n+1}\}|^2 \leq C \sum_{n=0}^{N-1} |\{k_n < v \leq k_{n+1}\}| \leq C|B_{1/2}|.$$

Dividing by N, we establish a strict rate of decay for the measure of the super-level set:

$$|\{v \geq k_N\}|^2 \leq \frac{C(d,\lambda,\Lambda)}{N}.$$

To satisfy the lemma for any given $\varepsilon > 0$, we simply choose N large enough such that $\sqrt{C/N} \leq \varepsilon$. We then define our required height parameter as $\eta = 1 - k_N = 2^{-N}$. This forces:

$$|\{v \geq 1-\eta\} \cap B_{1/2}| \leq \varepsilon,$$

completing the proof. ■

Remark 17.2.1 (The End of the Bootstrapping Loop). With Lemma 16.2.4 proven, Claim 16.2.1 is completely rigorous. The oscillation of the directional derivative $v = \partial_e u$ geometrically decays, proving $v \in C^\alpha$. Schauder theory activates, bootstrapping the minimizer $u \in W^{1,2}$ to $u \in C^\infty$.

Hilbert's 19th Problem is officially solved.

§17.3 The Vectorial Case: De Giorgi's Counterexample

Throughout our proof of Hilbert's 19th Problem, we assumed that our minimizer u is a scalar-valued function ($u\colon \mathbb{R}^d \to \mathbb{R}$). A natural question arises: does this magnificent regularity theory hold for systems of equations, where u is vector-valued ($u\colon \mathbb{R}^d \to \mathbb{R}^\ell$ for $\ell > 1$)?

In 1968, De Giorgi proved that the answer is definitively **no** for dimensions $\ell \geq 3$. He constructed a brilliant counterexample demonstrating that weak solutions to linear elliptic systems with bounded measurable coefficients can be unbounded, meaning the crucial $L^2 - L^\infty$ estimate fails completely in the vectorial case.

Theorem 17.3.1 (De Giorgi's Counterexample (1968))
Let $\ell \geq 3$. There exists a tensor of coefficients $A^{ij}_{\alpha\beta}(x)$ (for spatial indices $i, j = 1, \dots, d$ and vectorial indices $\alpha, \beta = 1, \dots, m$) that is symmetric and uniformly elliptic:

$$\lambda \,\mathrm{Id} \leq A^{ij}_{\alpha\beta}(x) \leq \Lambda \,\mathrm{Id}$$

and a vector-valued function $u \in W^{1,2}(B_1; \mathbb{R}^\ell)$ that weakly solves the linear elliptic system:

$$\mathrm{div}(A(x)Du) = 0 \quad \text{in } B_1$$

such that $u \notin L^\infty_{\mathrm{loc}}(B_1; \mathbb{R}^\ell)$.

In particular, u is a minimizer of the strictly convex quadratic functional

$$\int_{B_1} \langle A(x)Du, Du\rangle \, dx,$$

yet it fails to be continuous.

Remark 17.3.2 (Given by Gemini). In his original paper, De Giorgi takes $\ell = d$ and explicitly builds the unbounded solution in the form:

$$u(x) = \frac{x}{|x|^\gamma}$$

for a carefully chosen exponent $\gamma > 0$. Because the origin acts as a singularity where the function blows up to infinity, it is impossible to establish the Oscillation Lemma. For vectorial problems, one must abandon the hope of everywhere-regularity and instead rely on *partial regularity theory*, which guarantees that solutions are smooth everywhere except on a singular set of measure zero.

18 Convergence of Variational Problems

Consider a sequence of problems (P_m) concerning the minimization of a functional F_m among competitors $u \in X_m$. We would like to understand in which sense these variational problems approximate a limit problem (P_∞). Specifically, we are interested in integral functionals of the form:

$$F_m(u) = \int f_m(x, u(x), Du(x))\, dx$$

§18.1 Motivations

A primary motivation for this study comes from homogenization theory (for a comprehensive reference, see [Bra02].

§18.1.1 Periodic Homogenization

Consider the elliptic boundary value problem associated with the operator $\operatorname{div}(A(x)Du) = f$. Provided that $A \in \mathrm{Sym}_{d\times d}$ and $A(x) \geq 0$, this PDE is formally equivalent to the minimization problem:

$$\min \left\{ \frac{1}{2} \int \langle A(x)Du, Du\rangle \; dx - \int uf\, dx \right\}$$

Here, the matrix field $A(x)$ describes the local anisotropy and inhomogeneity of the material around a point x. In most physical situations, A varies at a scale strictly smaller than the observation scale. One might hope that there exists an "effective" or "macroscopic" problem that describes the behavior of the solution at large scales.

It is convenient to introduce a microscopic length scale $\varepsilon > 0$. By setting $A_\varepsilon(x) = A\left(\frac{x}{\varepsilon}\right)$, we consider the sequence of scaled problems (P_ε) given by:

$$\min \left\{ \frac{1}{2} \int \langle A_\varepsilon(x)Du, Du\rangle \; dx - \int uf\, dx \right\}$$

We would like to find an effective matrix field $\overline{A}_{\text{eff}}(x)$ such that the minimizers u_ε of (P_ε) converge in a suitable sense, $u_\varepsilon \sim \bar{u}$, where $\bar{u}$ is the minimizer of the effective problem:

$$\min \left\{ \frac{1}{2} \int \langle \overline{A}_{\text{eff}}(x)Du, Du\rangle \; dx - \int uf\, dx \right\}$$

We will mostly consider the setting in which $A(x)$ is strictly positive definite and 1-periodic (i.e., $\mathbb{Z}^d$-periodic). Consequently, the scaled matrix $A\left(\frac{x}{\varepsilon}\right)$ is ε-periodic (i.e., $\varepsilon\mathbb{Z}^d$-periodic). In this case, we expect the effective matrix to be a constant tensor, meaning $\overline{A}_{\text{eff}}(x) \equiv A_{\text{eff}}$ is independent of x. This specific framework is known as **periodic homogenization**.

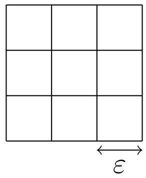


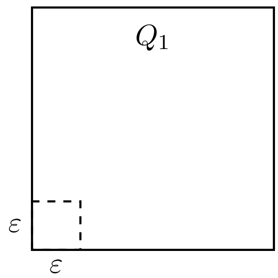
Figure 18.1: The material domain partitioned into ε-periodic microscopic cells.

Remark 18.1.1. There is also a broader case in which the matrix A is chosen randomly according to some probability distribution. This leads to the theory of **stochastic homogenization**.

Note that in the convex case, the equivalence between the minimization problem and the Euler-Lagrange (EL) equations implies that this is exactly the same as understanding in which sense the solutions of the PDE $-\operatorname{div}\left(A\left(\frac{x}{\varepsilon}\right)Du\right) = f$ converge.

§18.1.2 Homogenization of Riemannian Metrics (Hamilton-Jacobi Equations)

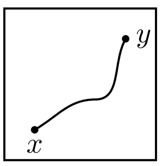
Let $g_{\alpha\beta}$ be a Riemannian metric on, say, the unit cube Q_1. We define the highly oscillatory scaled metric as $g_\varepsilon(x) = g_{\alpha\beta}\left(\frac{x}{\varepsilon}\right)$.

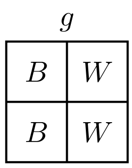


We would like to understand the global behavior of the minimal length problem:

$$\min\left\{\int_0^1 g_{\alpha\beta}\left(\frac{\gamma(t)}{\varepsilon}\right)\dot{\gamma}^\alpha(t)\dot{\gamma}^\beta(t)\,dt \;:\; \gamma:[0,1]\to\mathbb{R}^d,\ \gamma(0)=x,\ \gamma(1)=y\right\}$$

The typical setting assumes that g is 1-periodic, making g_ε strictly ε-periodic (though g can also be a random "1-independent" metric).

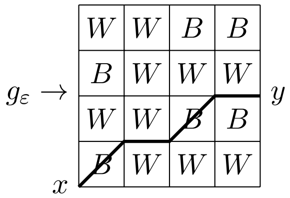




For a concrete example, suppose $g_{ij}(x)$ takes two distinct phases:

$$g_{ij}(x) = \begin{cases} \alpha\delta_{ij} & \text{in } B \\ \beta\delta_{ij} & \text{in } W \end{cases}$$

where $\alpha \ll \beta$ (specifically, $\alpha < \sqrt{2}\beta$).

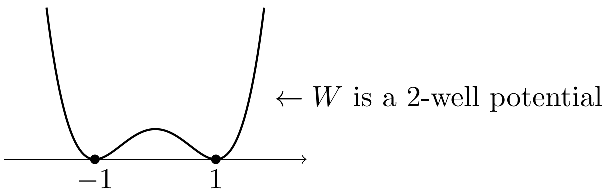


Because passing through the "cheap" regions (B) is highly preferred over the "expensive" regions (W), the optimal curves will zigzag diagonally. At a large scale, we can expect the metric to converge to a **non-Riemannian** behavior, like the L^∞ metric:

$$d(x, y) = \max\{|x^1 - y^1|, |x^2 - y^2|\}$$

§18.1.3 Gradient Theory of Phase Transitions

This framework naturally captures models such as Ginzburg-Landau, Allen-Cahn, and Modica-Mortola. We consider an order parameter $u : \Omega \to [-1, 1]$, where the set $\{u = -1\}$ represents Phase 1, and $\{u = 1\}$ represents Phase 2.



The classical energy minimization problem is given by:

$$\min \left\{ \int_\Omega W(u)\, dx : \fint_\Omega u\, dx = m \right\}$$

This problem has a massive degeneracy of minimizers. For any set $E \subset \Omega$ such that $\dfrac{|E| - |E^c|}{|\Omega|} = m$, the function $u = \chi_E - \chi_{E^c}$ is a global minimizer.

To select the physically relevant minimizer, we introduce an **expansion in gradients**:

$$W_\varepsilon(u, Du) = W(u) + \varepsilon^2 |Du|^2$$

We analyze the sequence of problems (P_ε):

$$\min \left\{ \int_\Omega \left(W(u) + \varepsilon^2 |Du|^2 \right) dx : \fint_\Omega u\, dx = m \right\}$$

For $\varepsilon \ll 1$, we expect the minimizer $u_\varepsilon \sim \chi_E - \chi_{E^c}$. Formally, we can write the asymptotic expansion of the energy:

$$(P_\varepsilon) \simeq P_0 + \varepsilon P_1 + O(\varepsilon^2)$$

where P_0 represents $\min \left\{ \fint W(u) \right\}$, and P_1 should be the problem that "reflects" an optimal selection $\gamma \in \operatorname{argmin} P_0$.

To isolate P_1, it should be characterized as the limit of the renormalized energy differences $\dfrac{\min P_\varepsilon - \min P_0}{\varepsilon}$. Thus, P_1 is the limit of the rescaled functional $(\tilde{P}_\varepsilon)$:

$$\inf \left\{ \int_\Omega \left(\frac{W(u)}{\varepsilon} + \varepsilon |Du|^2 \right) dx : \fint_\Omega u\, dx = m \right\}$$

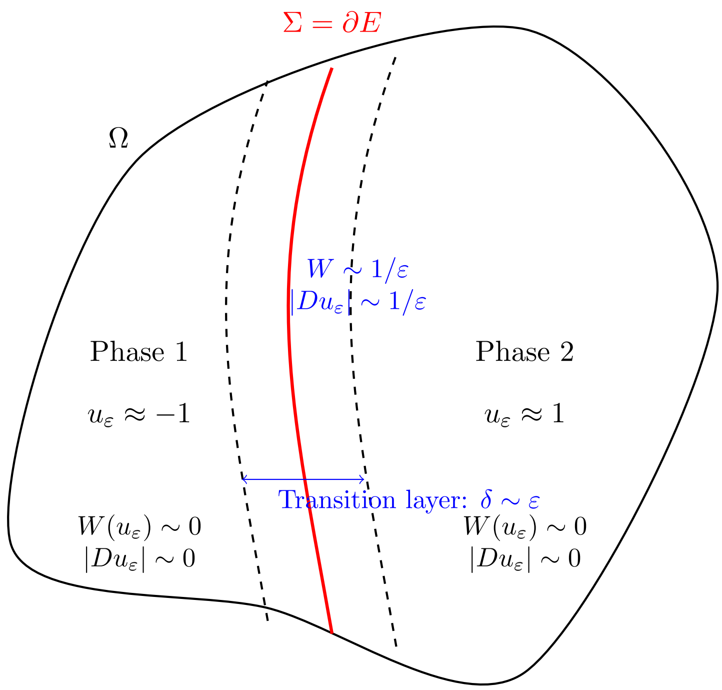


Figure 18.2: The Modica-Mortola phase transition. The macroscopic domain Ω with a transition layer of width $\delta \sim \varepsilon$ concentrated around the interface Σ.

By our physical intuition, $u_\varepsilon \sim \chi_E - \chi_{E^c}$ everywhere except on an interface $\Sigma = \partial E$. Outside a δ-neighborhood of ∂E, we expect $u_\varepsilon = \pm 1$, which implies $W(u_\varepsilon) \sim 0$ and $|Du_\varepsilon| \sim 0$.

Inside the δ-neighborhood, the state must transition, so $W(u_\varepsilon) \sim 1$. The gradient scales like $|Du_\varepsilon| \sim 1/\delta$. Furthermore, the volume of this region is proportional to the perimeter of the interface:

$$|\delta\text{-neighborhood}| \sim \delta \, \mathrm{Per}(E) = \delta \, \sigma(\Sigma)$$

If we plug this analytical guess into the energy functional, we obtain:

$$\int_\Omega \left(\frac{W(u_\varepsilon)}{\varepsilon} + \varepsilon |Du_\varepsilon|^2 \right) dx \sim \int_{\delta\text{-nbhd}} \left(\frac{W(u_\varepsilon)}{\varepsilon} + \varepsilon |Du_\varepsilon|^2 \right) dx = \delta \, \mathrm{Per}(E) \left[\frac{1}{\varepsilon} + \frac{\varepsilon}{\delta^2} \right]$$

Simplifying this expression yields:

$$= \mathrm{Per}(E) \left[\frac{\delta}{\varepsilon} + \frac{\varepsilon}{\delta} \right]$$

By optimizing this bound (which occurs at $\delta = \varepsilon$), we get:

$$\overset{\delta=\varepsilon}{\gtrsim} \mathrm{Per}(E)$$

In some sense, we expect the sequence of problems:

$$(\tilde{P}_\varepsilon) \quad \min \left\{ \int_\Omega \left(\frac{W(u)}{\varepsilon} + \varepsilon |Du|^2 \right) dx : \fint_\Omega u \, dx = m \right\}$$

$$\longrightarrow \quad (P) \quad \min \left\{ \mathrm{Per}(E, \Omega) : |E| - |E^c| = m|\Omega| \right\}$$

§18.2 Introduction to Γ-Convergence

The rigorous framework to formalize these asymptotic behaviors is Γ-convergence, introduced by De Giorgi. Let (X, d) be a metric space (or more generally, (X, τ) a topological space). Let $F_\varepsilon : X \to \mathbb{R} \cup \{+\infty\}$. We define the sequence of minimization problems and the limit problem as:

$$\begin{aligned} (P_\varepsilon): &\quad \inf\{F_\varepsilon(u) : u \in X\} \\ (P_0): &\quad \inf\{F_0(u) : u \in X\} \end{aligned}$$

Our key requirements to say that $F_\varepsilon \xrightarrow{\Gamma} F_0$ are:

1. If u_ε is an almost minimizer of (P_ε), meaning $F_\varepsilon(u_\varepsilon) = \inf_v F_\varepsilon(v) + o_\varepsilon(1)$, and $u_\varepsilon \to \bar{u}$, we want $\bar{u}$ to be a minimizer of (P_0).
2. If G is a d-continuous functional, we want the almost minimizers of $F_\varepsilon + G$ to converge to the minimizers of $F_0 + G$.

Moreover, we also consider other ingredients wehre we want the infimal values to converge properly:

1. $\forall \{v_\varepsilon\} \subset X : F_0(\bar{u}) = \inf\{F_0(v)\} = \lim_{\varepsilon \to 0} \inf_u \{F_\varepsilon(u)\} = \lim_{\varepsilon \to 0} F_\varepsilon(u_\varepsilon) \leq \liminf_{\varepsilon \to 0} F_\varepsilon(v_\varepsilon)$, where u_ε is a minimizer sequence.
2. If u_ε are a.e. minimizers, their energies should converge: $F_\varepsilon(u_\varepsilon) \to F_0(\bar{u})$.

Remark 18.2.1. Notee that for any function F and any target point $\bar{u}$, a standard penalization approach ensures:

$$u_\lambda \in \operatorname{argmin}\{F(u) + \lambda d(u, \bar{u})\} \xrightarrow{\lambda \to \infty} \bar{u}$$

To formalize this, let us properly package both requirements into the strict definition of Γ-convergence.

Definition 18.2.2 (Γ-Convergence)
Let (X, d) be a metric space (or more generally, a topological space (X, τ)). Let $F_\varepsilon : X \to \overline{\mathbb{R}}$ and $F : X \to \overline{\mathbb{R}}$. We say $F_\varepsilon \xrightarrow{\Gamma} F$ if both of the following conditions hold:

1. **(Γ-liminf inequality):** For every $\varepsilon_k \searrow 0$, for every $u \in X$, and for every sequence $u_{\varepsilon_k} \to u$, we have:
$$F(u) \leq \liminf_{k \to \infty} F_{\varepsilon_k}(u_{\varepsilon_k})$$
2. **(Γ-limsup inequality):** For every $u \in X$, there exists a sequence $u_{\varepsilon_k} \to u$ such that:
$$F(u) \geq \limsup_{k \to \infty} F_{\varepsilon_k}(u_{\varepsilon_k})$$

Remark 18.2.3. Several immediate structural consequences follow from this definition:

1. By coupling condition (1) with condition (2), the limit supremum in (2) is actually a strict limit.

2. The optimal sequence constructed in (2) is traditionally called the **recovery sequence**.
3. The definition can be compactly rewritten in terms of infima over all converging sequences:

$$F(u) = \inf \left\{ \liminf_{\varepsilon \to 0} F_\varepsilon(u_\varepsilon) : u_\varepsilon \to u \right\} \iff \text{Condition (1)}$$
$$F(u) = \inf \left\{ \limsup_{\varepsilon \to 0} F_\varepsilon(u_\varepsilon) : u_\varepsilon \to u \right\} \iff \text{Condition (2)}$$

Theorem 18.2.4 (Fundamental Theorem of Γ-Convergence)

Assume $F_\varepsilon \xrightarrow{\Gamma} F$. Let $u_\varepsilon \in X$ be a sequence such that:

a) $F_\varepsilon(u_\varepsilon) - \inf_X F_\varepsilon(u_\varepsilon) \to 0$ (i.e., u_ε is almost minimizing),

b) $u_\varepsilon \to \bar{u}$.

Then $\bar{u}$ is a global minimizer of F.

Proof. Pick any arbitrary $v \in X$ and let v_{ε_k} be a recovery sequence for v (guaranteed by the Γ-limsup inequality). We can evaluate the energy bounds as $k \to \infty$:

$$O(1) + F(v) \overset{\Gamma\text{-limsup}}{\geq} F_{\varepsilon_k}(v_{\varepsilon_k}) \geq F_{\varepsilon_k}(u_{\varepsilon_k}) + O(1) \overset{\Gamma\text{-liminf}}{\geq} F(\bar{u}) + O(1)$$

Note that the middle inequality holds because u_{ε_k} is almost minimizing. Taking the limit as $k \to \infty$, we simply obtain $F(\bar{u}) \leq F(v)$. Since $v \in X$ was arbitrary, $\bar{u}$ is a minimizer of F. ∎

Example 18.2.5

Consider the functional $F : \mathbb{R} \to \mathbb{R}$ defined by a point mass at the origin:

$$F(u) = \begin{cases} 1, & u = 0 \\ 0, & u \neq 0 \end{cases}$$

If we consider the constant sequence $F_\varepsilon = F$, its Γ-limit is **not** F itself. Instead, $F_\varepsilon \xrightarrow{\Gamma} \bar{F}$, where $\bar{F}(u) \equiv 0$.

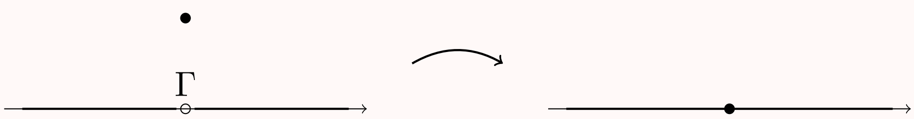


Figure 18.3: Γ-convergence of a constant sequence of a non-l.s.c. functional.

Remark 18.2.6. In general, Γ-convergence strongly lacks linearity! Specifically:

- $\Gamma\text{-}\lim(-F_\varepsilon) \neq -\Gamma\text{-}\lim(F_\varepsilon)$
- $\Gamma\text{-}\lim(F_\varepsilon + G_\varepsilon) \neq \Gamma\text{-}\lim(F_\varepsilon) + \Gamma\text{-}\lim(G_\varepsilon)$

19 Advanced Properties of Γ-Convergence and Homogenization

§19.1 Fundamental Properties of Γ-Convergence

Note that an equivalent, more compact characterization of the Γ-limit is:

$$F(x) = \Gamma\text{-}\lim_{\varepsilon\to 0} F_\varepsilon(x) \iff F(x) = \inf\left\{\liminf_{\varepsilon\to 0} F_\varepsilon(x_\varepsilon) : x_\varepsilon \to x\right\}$$
$$= \inf\left\{\limsup_{\varepsilon\to 0} F_\varepsilon(x_\varepsilon) : x_\varepsilon \to x\right\}$$

Proposition 19.1.1 (Fundamental Properties of Γ-Convergence)

Assume $F_\varepsilon \xrightarrow{\Gamma} F$.

1. **Convergence of Minimizers:** If $\{x_\varepsilon\}$ is an almost-minimizing sequence, meaning $F_\varepsilon(x_\varepsilon) = \inf_X F_\varepsilon + o_\varepsilon(1)$, and $x_\varepsilon \to \bar{x}$, then $\bar{x} \in \operatorname{argmin}_X F$.
2. **Stability under Continuous Perturbation:** If $G : X \to \mathbb{R}$ is d-continuous, then $F_\varepsilon + G \xrightarrow{\Gamma} F + G$.
3. **Lower Semicontinuity:** The Γ-limit F is always lower semicontinuous (l.s.c.).
4. **Relaxation of Constant Sequences:** If we take the constant sequence of functionals $F_\varepsilon \equiv F$, then $\Gamma\text{-}\lim F_\varepsilon = \bar{F}$, where $\bar{F}$ is the l.s.c. envelope (relaxation) of F:

$$\bar{F}(x) = \sup\{G(x) : G \leq F,\ G \text{ is l.s.c.}\} = \min_{y\to x}\left\{F(x), \liminf_{y\to x} F(y)\right\}$$

An immediate consequence of this is that the Γ-limit is **not linear**.

Proof. We address each property sequentially:

1. Let $y \in X$ be an arbitrary competitor, and let $\{y_\varepsilon\}$ be a recovery sequence for y such that $y_\varepsilon \to y$ and $F(y) \geq \limsup_{\varepsilon\to 0} F_\varepsilon(y_\varepsilon)$. We can then evaluate the energy bounds:

$$F(y) \geq \limsup_{\varepsilon\to 0} F_\varepsilon(y_\varepsilon) \geq \limsup_{\varepsilon\to 0}\left(\inf_X F_\varepsilon\right) = \limsup_{\varepsilon\to 0}\left(F_\varepsilon(x_\varepsilon) - o_\varepsilon(1)\right)$$
$$\geq \liminf_{\varepsilon\to 0} F_\varepsilon(x_\varepsilon)$$
$$\geq F(\bar{x}) \quad \text{(by the } \Gamma\text{-liminf inequality for } x_\varepsilon \to \bar{x})$$

Since $F(y) \geq F(\bar{x})$ for all $y \in X$, it follows that $\bar{x}$ is a global minimizer of F.

2. It is sufficient to prove that $F_\varepsilon \xrightarrow{\Gamma} F \implies F_\varepsilon + G \xrightarrow{\Gamma} F + G$.

- *Liminf:* Let $x \in X$ and $x_\varepsilon \to x$. By the continuity of G, $G(x_\varepsilon) \to G(x)$. Thus,

$$\liminf_{\varepsilon\to 0} \big(F_\varepsilon(x_\varepsilon) + G(x_\varepsilon)\big) = \liminf_{\varepsilon\to 0} F_\varepsilon(x_\varepsilon) + G(x) \geq F(x) + G(x)$$

- *Limsup:* Let $x \in X$. We work with the same recovery sequence $\{x_\varepsilon\}$ for F such that $x_\varepsilon \to x$ and $\limsup F_\varepsilon(x_\varepsilon) \leq F(x)$. Again by continuity of G,

$$\limsup_{\varepsilon\to 0} \big(F_\varepsilon(x_\varepsilon) + G(x_\varepsilon)\big) = \limsup_{\varepsilon\to 0} F_\varepsilon(x_\varepsilon) + G(x) \leq F(x) + G(x)$$

3. To show F is l.s.c., let $x \in X$ and consider a generic sequence $x_j \to x$. For each fixed j, let x_j^ε be a recovery sequence for x_j, meaning $x_j^\varepsilon \xrightarrow{\varepsilon\to 0} x_j$ and $F(x_j) = \lim_{\varepsilon\to 0} F_\varepsilon(x_j^\varepsilon)$.

   Thanks to a standard diagonalization argument, we can extract a sequence $\varepsilon(j) \searrow 0$ such that the diagonal sequence $x_j^{\varepsilon(j)} \to x$ as $j \to \infty$, and $F_{\varepsilon(j)}(x_j^{\varepsilon(j)}) = F(x_j) + o_j(1)$. Applying the Γ-liminf inequality to this diagonal sequence yields:

$$F(x) \leq \liminf_{j\to\infty} F_{\varepsilon(j)}(x_j^{\varepsilon(j)}) = \liminf_{j\to\infty} \big(F(x_j) + o_j(1)\big) = \liminf_{j\to\infty} F(x_j)$$

   This confirms F is lower semicontinuous.

4. Let $F_\varepsilon \equiv F$ and suppose its Γ-limit is denoted by F_Γ. We want to show $F_\Gamma = \bar{F}$.

   - By definition of the Γ-limit and the fact that F_Γ must be l.s.c. (from property 3) while satisfying $F_\Gamma \leq F$, it immediately follows that $F_\Gamma \leq \bar{F}$, since $\bar{F}$ is the supremum over all such functions.
   - To show $F_\Gamma \geq \bar{F}$, let $x \in X$ and consider the recovery sequence $x_\varepsilon \to x$ such that $F_\Gamma(x) = \lim_{\varepsilon\to 0} F(x_\varepsilon)$. Because $\bar{F} \leq F$ pointwise and $\bar{F}$ is l.s.c., we have:

$$F_\Gamma(x) = \lim_{\varepsilon\to 0} F(x_\varepsilon) \geq \liminf_{\varepsilon\to 0} \bar{F}(x_\varepsilon) \geq \bar{F}(x)$$

   Hence $F_\Gamma = \bar{F}$.

■

§19.2 Epigraphical Convergence (Kuratowski)

Definition 19.2.1 (Kuratowski Convergence of Sets)
Let (Y, d) be a metric space and $\{C_k\} \subseteq \mathcal{P}(Y)$. We say that $C_k \to C \in \mathcal{P}(Y)$ in the Kuratowski sense (denoted $C_k \xrightarrow{K} C$) if:

1. $\forall x_k \in C_k$, if $x_k \to x$, then $x \in C$.
2. $\forall x \in C$, $\exists x_k \in C_k$ such that $x_k \to x$.

Exercise 19.2.2
Prove that if $C_k \xrightarrow{K} C$, then C is automatically a closed set.

Proof. Let $\bar{x} \in \bar{C}$. There exists a sequence $x^{(j)} \in C$ such that $x^{(j)} \to \bar{x}$ as $j \to \infty$. By the second property of Kuratowski convergence, for each fixed j, there exists a sequence $x_k^{(j)} \in C_k$ such that $x_k^{(j)} \to x^{(j)}$ as $k \to \infty$. By a standard diagonalization procedure, we can extract a sequence $k(j) \to \infty$ such that $y_j := x_{k(j)}^{(j)} \to \bar{x}$. Since $y_j \in C_{k(j)}$, the first property of Kuratowski convergence implies $\bar{x} \in C$. Thus C is closed. ∎

Proposition 19.2.3

$F_k \xrightarrow{\Gamma} F \iff \operatorname{epi} F_k \xrightarrow{K} \operatorname{epi} F$, where $\operatorname{epi} F_k = \{(x,t) \in X \times \mathbb{R} : F_k(x) \le t\}$.

Proof. Assume $F_k \xrightarrow{\Gamma} F$.

- Let $(x_k, t_k) \in \operatorname{epi} F_k$ converge to (x,t). Then $F_k(x_k) \le t_k$. Taking the liminf and applying the Γ-liminf inequality gives $F(x) \le \liminf F_k(x_k) \le \liminf t_k = t$. Hence $(x,t) \in \operatorname{epi} F$.
- Let $(x,t) \in \operatorname{epi} F$, so $F(x) \le t$. By the Γ-limsup inequality, there exists a recovery sequence $x_k \to x$ with $\limsup F_k(x_k) \le F(x) \le t$. By selecting a sequence $t_k \ge F_k(x_k)$ such that $t_k \to t$, we construct $(x_k, t_k) \in \operatorname{epi} F_k$ converging to (x,t).

The reverse implication follows symmetrically by translating Kuratowski sequences into Γ-recovery sequences and liminf bounds. ∎

Remark 19.2.4.

1. Γ-convergence is the convergence of epigraphs, **not** of graphs!
2. $C_k \xrightarrow{K} C \iff \overline{C}_k \xrightarrow{K} C$.
3. $\operatorname{epi} \bar{F} = \overline{\operatorname{epi} F}$, which is exactly the l.s.c. envelope.

Example 19.2.5

Consider the highly oscillating functional on $[-1,1]$ given by $F_k(x) = x^2 + \sin(kx)$. The pointwise limit is undefined, but the Γ-limit strongly favors the minimizing wells of the oscillation:

$$F_k(x) \xrightarrow{\Gamma} F(x) = x^2 - 1$$

§19.3 Relaxation and the Double-Well Potential

Γ-convergence serves as an excellent selection criterion for minimizers in relaxed problems.

Example 19.3.1 (Phase Transition Relaxation)
Consider the double-well potential $W(t) = (1-t^2)^2$. We define the sequence of constant functionals:

$$F_k(u) = \int_0^1 W(u)\,dx$$

Let the domain space be $X_M = \{u : [0,1] \to \mathbb{R},\ |u(x)| \le M\} \subseteq L^\infty([0,1])$, equipped with the topology τ being the L^∞ weak* topology (which is metrizable on bounded sets, denoted by metric d_M).

Claim: $F_k \xrightarrow{\Gamma} \bar{F}(u) = \displaystyle\int_0^1 W^{**}(u)\,dx$, where W^{**} is the convex envelope:

$$W^{**}(t) = \begin{cases} W(t), & |t| \ge 1 \\ 0, & |t| < 1 \end{cases}$$

Sketch of Proof for Claim. First, note that $\bar{F}$ is l.s.c. with respect to weak* convergence since it is the integral of a convex and strongly continuous function and we invoke Theorem 4.2.3.

Γ-liminf inequality: By the fundamental property of the convex conjugate, $W(t) \ge W^{**}(t)$ for all t. Thus,

$$F_k(u_k) = \int_0^1 W(u_k)\,dx \ge \int_0^1 W^{**}(u_k)\,dx$$

If $u_k \overset{*}{\rightharpoonup} u$ in L^∞, the lower semicontinuity of the convex functional integral immediately yields:

$$\liminf_{k\to\infty} \int_0^1 W^{**}(u_k)\,dx \ge \int_0^1 W^{**}(u)\,dx = \bar{F}(u)$$

This establishes the liminf bound.

Γ-limsup inequality: By standard density arguments in L^∞ weak*, it is enough to construct a recovery sequence for piecewise constant simple functions u. Let $u = \sum \alpha_j 1_{E_j}$. As before, we can assume without loss of generality that the sets $E_j = (a_j, b_j)$ are pairwise disjoint intervals. We decompose $u = u_1 + u_2$, where u_1 collects the regions where $|\alpha_j| > 1$ (outside the well) and u_2 collects $|\alpha_j| \le 1$ (inside the well).

Consider a single interval (a,b) where $v = \alpha 1_{(a,b)}$ with $|\alpha| \le 1$. We partition (a,b) into two sub-intervals, I_1 of size $\lambda(b-a)$ and I_2 of size $(1-\lambda)(b-a)$. We choose λ such that the average value is exactly α:

$$\lambda(1) + (1-\lambda)(-1) = 2\lambda - 1 = \alpha \implies \lambda = \frac{\alpha+1}{2}$$

We define the microscopic oscillation profile:

$$\tilde{v}(x) = \begin{cases} 1 & \text{on } I_1 \\ -1 & \text{on } I_2 \end{cases}$$

and extend it $\mathbb{Z}$-periodically. We then define our highly oscillating recovery sequence locally as $v_k(x) := \tilde{v}(kx)$. By the Riemann-Lebesgue Lemma (Lemma 5.2.5), $v_k \overset{*}{\rightharpoonup} \alpha$ in $L^\infty(a, b)$.

Crucially, because $v_k(x) \in \{-1, 1\}$ almost everywhere, evaluating the true non-convex energy gives identically zero:

$$\int_a^b W(v_k(x))\, dx \equiv 0 = \int_a^b W^{**}(\alpha)\, dx$$

For the full function $u = u_1 + u_2$, we define the global recovery sequence by replacing the flat constant regions α_j in u_2 with these micro-oscillations:

$$u_k = u_1 + \sum_{j:|\alpha_j|\le 1} \tilde{v}_k^{(a_j,b_j)} \cdot 1_{(a_j,b_j)}$$

This construction explicitly recovers the relaxed energy $\bar{F}(u)$, satisfying the limsup inequality. ∎

§19.4 Example: Periodic Homogenization

Let $A : \mathbb{R}^d \to \mathrm{Sym}^+_{d\times d}$ (the space of symmetric positive definite matrices) be a measurable matrix field such that:

1. A is $\mathbb{Z}^d$-periodic.
2. There exist $0 < \lambda \le \Lambda$ such that $\lambda\, \mathrm{Id} \le A(x) \le \Lambda\, \mathrm{Id}$ (uniform ellipticity).

Note that for any $\varepsilon > 0$, the scaled field $A_\varepsilon(x) = A\left(\frac{x}{\varepsilon}\right)$ is $\varepsilon\mathbb{Z}^d$-periodic.

For any bounded domain Ω, given data $f \in L^2(\Omega)$ and boundary trace $g \in W^{1,2}(\Omega)$, we investigate the limit of the solutions u_ε to the Dirichlet boundary value problem:

$$\begin{cases} -\operatorname{div}(A_\varepsilon D u_\varepsilon) = f & \text{in } \Omega \\ u_\varepsilon \in g + W_0^{1,2}(\Omega) \end{cases}$$

This is equivalent to understanding the limit of the variational problem (P_ε):

$$\inf \left\{ \frac{1}{2} \int_\Omega \langle A_\varepsilon Dv, Dv \rangle\, dx - \int_\Omega f v\, dx : v \in g + W_0^{1,2}(\Omega) \right\}$$

Assume there exists an effective matrix $A_0 \in \mathrm{Sym}^+_{d\times d}$ such that if we define the functionals:

$$F_\varepsilon(u) = \begin{cases} \frac{1}{2} \int_\Omega \langle A_\varepsilon Du, Du \rangle\, dx & \text{if } u \in g + W_0^{1,2}(\Omega) \\ +\infty & \text{otherwise} \end{cases}$$

$$F(u) = \begin{cases} \frac{1}{2} \int_\Omega \langle A_0 Du, Du \rangle\, dx & \text{if } u \in g + W_0^{1,2}(\Omega) \\ +\infty & \text{otherwise} \end{cases}$$

then $F_\varepsilon \xrightarrow{\Gamma} F$ in the strong L^2 topology.

Then for all $f \in L^2(\Omega)$, the unique solution of (P_ε) weakly converges in $W^{1,2}$ to the solution of (P_0). To see this, note that the linear perturbation $G_f(u) := -\int_\Omega uf\,dx$ is continuous with respect to L^2 convergence. Thus, by stability under continuous perturbations:

$$F_\varepsilon + G_f \xrightarrow{\Gamma} F + G_f$$

Moreover, if u_ε is a minimizer of (P_ε), we can test the energy against the boundary condition g itself:

$$\begin{aligned}
\lambda\|Du_\varepsilon\|_{L^2}^2 &\le \int_\Omega \langle A_\varepsilon Du_\varepsilon, Du_\varepsilon\rangle\,dx \le \int_\Omega \langle A_\varepsilon Dg, Dg\rangle\,dx + \int_\Omega f(u_\varepsilon - g)\,dx \\
&\le \Lambda\|Dg\|_{L^2}^2 + \|f\|_{L^2}\|u_\varepsilon - g\|_{L^2} \\
&\overset{\text{Poincaré}}{\lesssim} \Lambda\|Dg\|_{L^2}^2 + \|f\|_{L^2}\|Du_\varepsilon - Dg\|_{L^2} \\
&\overset{\text{Young}}{\lesssim} \Lambda\|Dg\|_{L^2}^2 + C(\|f\|_{L^2}, \|Dg\|_{L^2}, \delta) + \delta\|Du_\varepsilon\|_{L^2}^2
\end{aligned}$$

Choosing $\delta = \lambda/2$ allows us to absorb the final term, yielding a uniform bound $\|Du_\varepsilon\|_{L^2} \le C$. This **equicoercivity** implies that u_ε are precompact in L^2. Up to a subsequence, $u_\varepsilon \xrightarrow{L^2} \bar{u}$, and by the Fundamental Theorem of Γ-convergence (Therem 18.2.4), $\bar{u} \in \operatorname{argmin}\{F(u) + G_f(u)\}$. By strict convexity, $\bar{u}$ is unique and solves the homogenized Euler-Lagrange equation:

$$\begin{cases} -\operatorname{div}(A_0 D\bar{u}) = f \\ \bar{u} \in g + W_0^{1,2}(\Omega) \end{cases}$$

§19.4.1 The 1D Homogenization Case

Let us focus the homogenization problem in the 1D case $\Omega = (0,1)$. Let a be $\mathbb{Z}$-periodic with uniform bounds $0 < \lambda \le a(y) \le \Lambda$. Let $a_\varepsilon(x) = a\left(\frac{x}{\varepsilon}\right)$. We consider the minimization of:

$$(P_\varepsilon) \quad \min\left\{\int_0^1 a\left(\frac{x}{\varepsilon}\right)|u'|^2\,dx : u \in W^{1,2}(0,1),\, u(0) = 0,\, u(1) = 1\right\}$$

Note that for any fixed u, the Riemann-Lebesgue Lemma immediately suggests a pointwise limit:

$$F_\varepsilon(u) \to \tilde{F}(u) = \int_0^1 \bar{a}|u'|^2\,dx, \quad \text{where } \bar{a} = \int_0^1 a(x)\,dx$$

However, this naive pointwise guess is incorrect for the variational limit!

Theorem 19.4.1 (1D Periodic Homogenization)
$F_\varepsilon \xrightarrow{\Gamma} F$ in L^2, where the limit functional is defined by the **harmonic mean**:

$$F(u) = \begin{cases} \displaystyle\int_0^1 \underline{a}|u'|^2\,dx & u \in W^{1,2}(0,1), \text{ b.c.} \\ +\infty & \text{otherwise} \end{cases}$$

and the effective coefficient is $\underline{a} := \left(\fint_0^1 \frac{1}{a(y)}\,dy \right)^{-1}$. Note that by the AM-HM inequality, $\underline{a} \leq \bar{a}$.

Remark 19.4.2. Our intuition for guessing $\underline{a}$ stems from the fact that the actual minimizer u_ε of F_ε exactly solves the Euler-Lagrange equation:

$$(a_\varepsilon(x)u_\varepsilon')' = 0 \implies a_\varepsilon(x)u_\varepsilon' = C \implies u_\varepsilon'(x) = \frac{C}{a_\varepsilon(x)}$$

Passing to the weak limit in L^2, the rapidly oscillating inverse coefficient converges to its average:

$$u_\varepsilon' \xrightarrow{L^2} C \fint_0^1 \frac{1}{a(y)}\,dy = \frac{C}{\underline{a}} = \bar{u}'$$

Thus, the correct macroscopic relation is $\underline{a}\bar{u}' = C$, which means the limiting displacement $\bar{u}$ solves the homogenized equation $(\underline{a}\bar{u}')' = 0$.

20 Proof of 1D Periodic Homogenization

The following sections construct the formal proof of the 1D periodic homogenization, remaining faithful to the original lecture presentation. For an alternative and comprehensive treatment of this problem, the reader is referred to [San23].

§20.1 Proof of the Γ-liminf Inequality

It is sufficient to consider sequences $u_k \xrightarrow{L^2} u$ such that $\sup_k \mathcal{F}_k(u_k) \leq C$; otherwise, the liminf inequality holds trivially.

By the uniform ellipticity of a_k, we have

$$\int_0^1 |u_k'|^2 \, dx \leq \frac{1}{\lambda} \int_0^1 a_k(x)|u_k'|^2 \, dx \leq \frac{C}{\lambda}.$$

This implies that the sequence $\{u_k\}$ is uniformly bounded in $W^{1,2}(0,1)$. By Morrey's Embedding Theorem and compactness, $u_k \to u$ uniformly on $[0,1]$.

For any arbitrary resolution $N \in \mathbb{N}$, we bisect the domain $[0,1]$ into intervals of length $1/N$ to obtain:

$$\mathcal{F}_k(u_k) = \int_0^1 a_k(x)|u_k'(x)|^2 \, dx = \sum_{j=0}^{N-1} \int_{j/N}^{(j+1)/N} a_k(x)|u_k'(x)|^2 \, dx.$$

Claim 20.1.1 — For any interval $I = [a,b]$, the following lower bound holds:

$$\int_I a_k |u_k'|^2 \, dx \geq \frac{|u_k(b) - u_k(a)|^2}{|I| \cdot \fint_I \frac{1}{a_k} \, dx}.$$

Proof of Claim 20.1.1. Using Hölder's inequality, we separate the integrand:

$$\int_I |u_k'| \, dx = \int_I \left(\sqrt{a_k}|u_k'|\right) \frac{1}{\sqrt{a_k}} \, dx \leq \left(\int_I a_k |u_k'|^2 \, dx\right)^{1/2} \left(\int_I \frac{1}{a_k} \, dx\right)^{1/2}.$$

Squaring both sides and rearranging yields:

$$\int_I a_k |u_k'|^2 \, dx \geq \frac{\left(\int_I |u_k'| \, dx\right)^2}{\int_I \frac{1}{a_k} \, dx} \geq \frac{\left|\int_I u_k' \, dx\right|^2}{\int_I \frac{1}{a_k} \, dx} = \frac{|u_k(b) - u_k(a)|^2}{|I| \cdot \fint_I \frac{1}{a_k} \, dx},$$

where equality occurs if and only if $u_k' = C/a_k$ almost everywhere on I. ■

Following the claim, we bound the functional from below over the partitioned domain:

$$\mathcal{F}_k(u_k) \geq \sum_{j=0}^{N-1} N \cdot \frac{\left|u_k\left(\frac{j+1}{N}\right) - u_k\left(\frac{j}{N}\right)\right|^2}{⨍_{j/N}^{(j+1)/N} \frac{1}{a_k}\,dx}.$$

Passing to the limit $k \to \infty$, the uniform convergence of u_k ensures point evaluations converge. Furthermore, the Riemann-Lebesgue Lemma implies $\dfrac{1}{a_k} \rightharpoonup \dfrac{1}{\underline{a}}$ weakly-$*$ in $L^\infty(0,1)$. Thus, we obtain:

$$\liminf_{k\to\infty} \mathcal{F}_k(u_k) \geq \sum_{j=0}^{N-1} \underline{a} \cdot N \cdot \left|u\left(\frac{j+1}{N}\right) - u\left(\frac{j}{N}\right)\right|^2.$$

Let u_N denote the piecewise linear interpolation of u on the nodal grid $\{0, \dfrac{1}{N}, \dots, \dfrac{N}{N}\}$. The right-hand side of the previous inequality evaluates exactly to $\mathcal{F}(u_N)$, leading to:

$$\liminf_{k\to\infty} \mathcal{F}_k(u_k) \geq \mathcal{F}(u_N) \quad \forall N \in \mathbb{N}.$$

Since $u \in W^{1,2}(0,1)$, its piecewise linear interpolants satisfy $u_N \xrightarrow{W^{1,2}} u$ strongly. By the continuity of the functional $\mathcal{F}$ with respect to the strong topology in $W^{1,2}(0,1)$, we conclude:

$$\liminf_{k\to\infty} \mathcal{F}_k(u_k) \geq \mathcal{F}(u).$$

■

§20.2 Proof of the Γ-limsup Inequality

It is sufficient to consider target functions $u \in W^{1,2}(0,1)$ that are piecewise linear, as this class is dense in energy. Specifically, we fix $N \in \mathbb{N}$ and assume u is linear on each subinterval $\left[\dfrac{j}{N}, \dfrac{j+1}{N}\right]$ of $[0,1]$ for $j = 0, \dots, N-1$.

Step 1: Local Minimizers. By scaling and translation, we can assume without loss of generality that $u(x) = x$ on $[0,1]$.

Let $v \in W^{1,2}(0,1)$ be the exact minimizer of the localized problem:

$$\int_0^1 a(y)|v'(y)|^2\,dy \quad \text{subject to } v(0) = 0,\ v(1) = 1.$$

Then, the rescaled function $v_k(x) = \dfrac{v(kx)}{k}$ is the minimizer of the localized functional over $(0, 1/k)$:

$$\min\left\{\int_0^{1/k} a(kx)|w'(x)|^2\,dx : w(0) = 0,\ w(1/k) = 1/k\right\}.$$

Applying the change of variables $y = kx$, the energy evaluated over this initial subinterval is exactly:

$$\mathcal{F}_k(v_k, (0, 1/k)) = \frac{1}{k}\int_0^1 a(y)|v'(y)|^2\,dy.$$

Similarly, by spatial translation, $v_k(x - j/k)$ represents the minimizer on the interval $[j/k, (j+1)/k]$ accommodating the boundary conditions $w(j/k) = 0$ and $w((j+1)/k) = 1/k$.

Step 2: Patching Minimizers. We construct the recovery sequence by setting $k = N$. On each subinterval $[j/k, (j+1)/k]$, we define the sequence u_k as:

$$u_k(x) = \frac{j}{k} + v_k\left(x - \frac{j}{k}\right).$$

By construction, $u_k(j/k) = j/k$, which ensures that u_k is continuous across all nodal points. The energy contribution of each subinterval is:

$$\int_{j/k}^{(j+1)/k} a_k(x)|u_k'(x)|^2\,dx = \int_{j/k}^{(j+1)/k} a(kx)\left|v'\left(k\left(x-\frac{j}{k}\right)\right)\right|^2 dx = \frac{1}{k}\int_0^1 a(y)|v'(y)|^2\,dy.$$

Summing over all k subintervals, the total sequence energy simplifies to:

$$\mathcal{F}_k(u_k) = \int_0^1 a(y)|v'(y)|^2\,dy.$$

Step 3: Convergence and Limsup. Our construction guarantees that the sequence is uniformly bounded in $W^{1,2}(0,1)$, meaning $\sup_k \|u_k\|_{W^{1,2}} \lesssim 1$, and converges strongly in L^2 to the macroscopic function $u(x) = x$. Taking the limit supremum yields:

$$\limsup_{k\to\infty} \mathcal{F}_k(u_k) = \inf\left\{\int_0^1 a(y)|w'(y)|^2\,dy : w(0) = 0,\, w(1) = 1\right\}.$$

To conclude the proof, we must demonstrate that this infimum evaluates exactly to $\underline{a} = \mathcal{F}(u)$ when $u(x) = x$.

Recalling Claim 20.1.1, for any minimizer v, the energy is strictly bounded from below:

$$\int_0^1 a(y)|v'(y)|^2\,dy \geq \frac{1}{\int_0^1 1/a(y)\,dy} = \underline{a}.$$

Equality is achieved when $v'(x) = \dfrac{C}{a(x)}$. Integrating this condition yields $v(x) = C\displaystyle\int_0^x \frac{dy}{a(y)}$, where the normalization constant C is uniquely determined by the boundary condition $C\displaystyle\int_0^1 \frac{1}{a(y)}\,dy = 1$. Consequently, v is indeed the minimizer, yielding the minimal energy $\underline{a}$.

■

Remark 20.2.1. The underlying mechanism demonstrated is that the homogenized coefficient represents the minimum local energy configuration:

$$\underline{a} = \inf\left\{\int_0^1 a(y)|w'(y)|^2\,dy : w(0) = 0,\, w(1) = 1\right\}.$$

Therefore, for a general macroscopic slope s across an arbitrary subinterval $[j/N, (j+1)/N]$, the local minimum energy naturally scales quadratically:

$$\frac{\underline{a}s^2}{N} = \inf\left\{\int_{j/N}^{(j+1)/N} a(Ny)|w'(y)|^2\,dy : w(j/N) = \alpha,\, w((j+1)/N) = \alpha + s/N\right\}.$$

§20.3 Multidimensional Homogenization

Theorem 20.3.1 (Multidimensional Homogenization)
Let $A : \mathbb{R}^d \to \mathrm{Sym}_{d\times d}$ be a $\mathbb{Z}^d$-periodic matrix-valued function satisfying uniform ellipticity. Define the highly oscillating coefficient $A_\varepsilon(x) := A(x/\varepsilon)$, and consider the functional:

$$\mathcal{F}_\varepsilon(u) = \int_\Omega \langle A_\varepsilon Du, Du\rangle\, dx.$$

Then $\mathcal{F}_\varepsilon \xrightarrow{\Gamma(L^2)} \mathcal{F}$, where the homogenized limit functional is given by:

$$\mathcal{F}(u) := \int_\Omega \langle A_{\text{eff}} Du, Du\rangle\, dx.$$

The effective matrix $A_{\text{eff}} \in \mathrm{Sym}_{d\times d}$ is explicitly determined via the cell formula over the unit hypercube $Q = [0,1]^d$:

$$\langle A_{\text{eff}}\xi, \xi\rangle := \inf\left\{\int_Q \langle A(x)Dv, Dv\rangle\, dx : v \in \langle \xi, x\rangle + W_0^{1,2}(Q)\right\}.$$

Theorem 20.3.2 (General Result)
Let $\mathcal{F}_k(u) = \displaystyle\int_\Omega F_k(x, Du)\, dx$. If $\mathcal{F}_k \xrightarrow{\Gamma} \mathcal{F}$ and the limit functional admits an integral representation $\mathcal{F}(u) = \displaystyle\int_\Omega F(x, Du)\, dx$, then by the lower semicontinuity of $\mathcal{F}$ and Morrey's Theorem, the effective integrand F must be quasi-convex. Specifically:

$$\begin{aligned} F(x,\xi) &= \inf\left\{\int_Q F(x, \xi + D\varphi)\, dy : \varphi \in W_0^{1,2}(Q)\right\} \\ &= \lim_{k\to\infty} \inf\left\{\mathcal{F}_k(v) : v \in \langle \xi, x\rangle + W_0^{1,2}(Q)\right\}. \end{aligned}$$

Exercise 20.3.3
Let the highly oscillating coefficient $a_k(x)$ be defined as:

$$a_k(x) := \begin{cases} 1/k & \text{if } |x| \le 1/k, \\ 1 & \text{if } |x| > 1/k. \end{cases}$$

Consider the corresponding functional sequence:

$$\mathcal{F}_k(u) = \int_{-1}^{1} a_k(x)|u'(x)|^2\, dx.$$

Find the explicit Γ-limit of $\mathcal{F}_k$ as $k \to \infty$.

§20.4 Alternative Proof of the Γ-liminf Inequality

Let $u_\varepsilon \xrightarrow{L^2} u$. We seek to establish that $\liminf_{\varepsilon\to 0} \mathcal{F}_\varepsilon(u_\varepsilon) \geq \mathcal{F}(u)$.

Without loss of generality, we may assume $\liminf_{\varepsilon\to 0} \mathcal{F}_\varepsilon(u_\varepsilon) < +\infty$. Extracting a subsequence that realizes this limit, there exists a uniform constant C such that:

$$\int_0^1 a_\varepsilon(x)(u_\varepsilon')^2\,dx \leq C.$$

Invoking the uniform ellipticity condition ($a_\varepsilon(x) \geq \lambda$), we obtain the uniform bound $\|u_\varepsilon'\|_{L^2}^2 \leq C/\lambda$. This bound dictates that u_ε is precompact in the weak topology of $W^{1,2}(0,1)$. Given the strong convergence $u_\varepsilon \xrightarrow{L^2} u$, it unconditionally follows that $u_\varepsilon \xrightharpoonup{W^{1,2}} u$, meaning $u_\varepsilon' \xrightharpoonup{L^2} u'$.

We encounter an analytical obstruction at this stage: passing to the limit directly in the product of a weakly converging sequence $(u_\varepsilon')^2$ and a highly oscillating sequence $a_\varepsilon(x)$ is generally invalid.

To circumvent this, we utilize Young's inequality. For any arbitrary function $v \in L^2(0,1)$ and any real value $p \in \mathbb{R}$, we have the pointwise bound:

$$a_\varepsilon p^2 \geq 2pv - \frac{v^2}{a_\varepsilon}.$$

Evaluating this inequality pointwise with $p = u_\varepsilon'(x)$ and integrating over the domain $(0,1)$ yields:

$$\int_0^1 a_\varepsilon(x)(u_\varepsilon')^2\,dx \geq \int_0^1 2u_\varepsilon' v\,dx - \int_0^1 \frac{v^2}{a_\varepsilon(x)}\,dx.$$

From the Riemann-Lebesgue Lemma, the highly oscillating coefficient reciprocal satisfies $\frac{1}{a_\varepsilon(x)} \xrightharpoonup{*} \frac{1}{\underline{a}}$ weakly-$*$ in $L^\infty(0,1)$. Taking the limit infimum as $\varepsilon \to 0$, the weak convergence of u_ε' and the weak-$*$ convergence of $1/a_\varepsilon$ permit us to pass to the limit in both linear terms smoothly:

$$\liminf_{\varepsilon\to 0} \int_0^1 a_\varepsilon(x)(u_\varepsilon')^2\,dx \geq \int_0^1 2u'v\,dx - \int_0^1 \frac{v^2}{\underline{a}}\,dx.$$

Crucially, this lower bound is valid for *any* test function $v \in L^2(0,1)$. To extract the sharpest bound, we optimize the right-hand side by choosing $v = \underline{a}u'$. Substituting this optimal choice back into the limit yields:

$$\liminf_{\varepsilon\to 0} \mathcal{F}_\varepsilon(u_\varepsilon) \geq \int_0^1 \left(2u'(\underline{a}u') - \frac{(\underline{a}u')^2}{\underline{a}}\right)dx = \int_0^1 \underline{a}(u')^2\,dx = \mathcal{F}(u),$$

which rigorously completes the proof of the liminf inequality.

■

21 Gradient Theory of Phase Transitions

§21.1 Introduction and the Unscaled Functional

Consider a material occupying a domain $\Omega \subset \mathbb{R}^d$ that can exist in two distinct phases. We describe the state of the material by a phase parameter $u : \Omega \to [-1, 1]$, where the pure phases correspond to the sets $\{u = -1\}$ and $\{u = 1\}$.

Typically, the phase separation is driven by a double-well potential. A standard choice is:

$$W(t) = \frac{(1 - t^2)^2}{4}.$$

Notice that the minima are precisely at the pure phases $t = \pm 1$. Near these minima, the potential behaves quadratically: $W(t) \sim C(t \mp 1)^2 + \mathcal{O}((t \mp 1)^3)$.

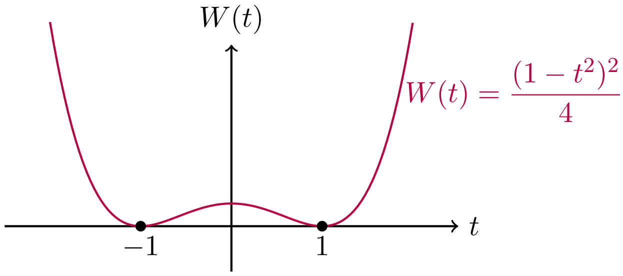


Figure 21.1: The double-well potential

We consider the following minimization problem over Ω subject to a mass constraint $\fint_\Omega u = m$:

$$\min \left\{ \int_\Omega W(u) + \varepsilon^2 |Du|^2 \, dx : \fint_\Omega u = m \right\}.$$

Here, $\varepsilon > 0$ represents a characteristic length scale of the transition layer between the two phases. We formally define the unscaled functional $\mathcal{G}_\varepsilon : L^2(\Omega) \to \mathbb{R} \cup \{+\infty\}$ as:

$$\mathcal{G}_\varepsilon(u) = \begin{cases} \displaystyle\int_\Omega \left(W(u) + \varepsilon^2 |Du|^2 \right) \, dx & \text{if } u \in W^{1,2}(\Omega; [-1, 1]), \\ +\infty & \text{otherwise.} \end{cases}$$

Remark 21.1.1. The Γ-limit of $\mathcal{G}_\varepsilon$ in the L^2-topology as $\varepsilon \to 0$ is simply $\displaystyle\int_\Omega W^{**}(u) \, dx$, where W^{**} is the convex envelope (the lower convex l.s.c. envelope) of W. For our specific potential:

$$W^{**}(t) = \begin{cases} 0 & \text{if } |t| \le 1, \\ W(t) & \text{if } |t| > 1. \end{cases}$$

Exercise 21.1.2

Try to prove the Γ-limit of $\mathcal{G}_\varepsilon$ is exactly $\int_\Omega W^{**}(u)\,dx$.

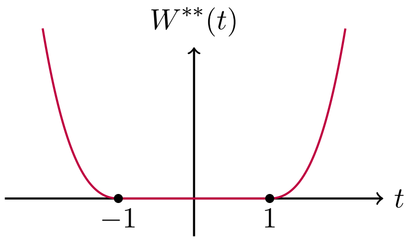


Figure 21.2: The convex envelope of double-well potential

§21.2 The Scaled Functional and Heuristic Approximation

Since the unscaled functional trivializes the gradient term in the limit, we rescale the energy by $\frac{1}{\varepsilon}$ to capture the precise energetic cost of the phase interfaces. We define $\mathcal{F}_\varepsilon = \frac{1}{\varepsilon}\mathcal{G}_\varepsilon$:

$$\mathcal{F}_\varepsilon(u) = \int_\Omega \left(\frac{1}{\varepsilon}W(u) + \varepsilon|Du|^2\right) dx.$$

Let us attempt to approximate the energy of a sharp jump. Assume Ω is partitioned into two regions, E and $\Omega \setminus E$, separated by a smooth hypersurface $\Sigma = \partial E \cap \Omega$. We construct a tubular neighborhood around Σ:

$$\Sigma_\delta = \{x \in \Omega : \text{dist}(x, \Sigma) \le \delta\}.$$

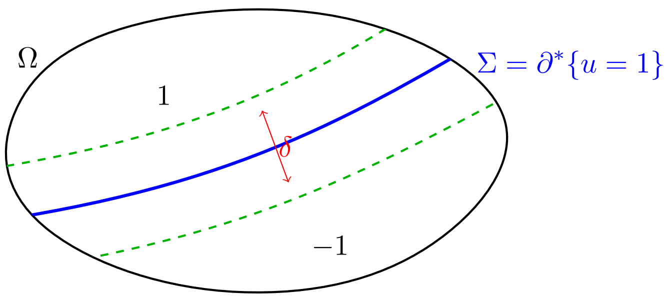


Figure 21.3: A heuristic representation of the interface Σ and its δ-tubular neighborhood separating the pure phases $\{u = 1\}$ and $\{u = -1\}$ within the domain Ω.

Let $sd(x, \Sigma)$ denote the signed distance function to Σ. We postulate a 1D transition profile $f : [-1, 1] \to \mathbb{R}$ with $f(-1) = -1$ and $f(1) = 1$. The macroscopic function is regularized as:

$$u_\delta(x) = \begin{cases} -1 & \text{if } sd(x,\Sigma) < -\delta, \\ f\left(\dfrac{sd(x,\Sigma)}{\delta}\right) & \text{if } |sd(x,\Sigma)| \le \delta, \\ 1 & \text{if } sd(x,\Sigma) > \delta. \end{cases}$$

Evaluating the energy of this constructed transition layer, noting that $|\nabla sd(x,\Sigma)| = 1$ almost everywhere near the boundary, we have:

$$\mathcal{F}_\varepsilon(u_\delta) = \int_{\Sigma_\delta} \left[\frac{1}{\varepsilon} W\left(f\left(\frac{sd(x,\Sigma)}{\delta} \right) \right) + \frac{\varepsilon}{\delta^2} \left| f'\left(\frac{sd(x,\Sigma)}{\delta} \right) \right|^2 |\nabla sd(x,\Sigma)|^2 \right] dx.$$

To evaluate this integral over level sets of the distance function, we rigorously rely on the Coarea Formula.

Theorem 21.2.1 (Coarea Formula)
Let $\Omega \subset \mathbb{R}^d$ be an open set and $v \in \mathrm{Lip}(\Omega)$ (or $W^{1,1}(\Omega)$). For any non-negative measurable function $g : \Omega \to \mathbb{R}$, the following holds:

$$\int_\Omega g(x)|\nabla v(x)|\, dx = \int_{-\infty}^{\infty} \left(\int_{\{v=s\}} g(\sigma)\, d\mathcal{H}^{d-1}(\sigma) \right) ds,$$

where $\mathcal{H}^{d-1}$ is the $(d-1)$-dimensional Hausdorff measure.

Applying the Coarea Formula with $v(x) = sd(x,\Sigma)$ and ds being the integration variable for the level sets:

$$\mathcal{F}_\varepsilon(u_\delta) = \int_{-\delta}^{\delta} \mathcal{H}^{d-1}(\{sd(x,\Sigma) = s\}) \left[\frac{1}{\varepsilon} W\left(f\left(\frac{s}{\delta} \right) \right) + \frac{\varepsilon}{\delta^2} \left| f'\left(\frac{s}{\delta} \right) \right|^2 \right] ds.$$

Applying a change of variables $z = s/\delta$, implying $ds = \delta\, dz$, we obtain:

$$\mathcal{F}_\varepsilon(u_\delta) = \int_{-1}^{1} \mathcal{H}^{d-1}(\{sd(x,\Sigma) = \delta z\}) \left[\frac{\delta}{\varepsilon} W(f(z)) + \frac{\varepsilon}{\delta} |f'(z)|^2 \right] dz.$$

Choosing the optimal transition width $\delta = \varepsilon$ and exploiting the continuity of the area measure $\mathcal{H}^{d-1}(\{sd = \varepsilon z\}) \approx \mathcal{H}^{d-1}(\Sigma)$ for small ε, we arrive at:

$$\mathcal{F}_\varepsilon(u_\varepsilon) \approx \mathcal{H}^{d-1}(\Sigma) \int_{-1}^{1} \left[W(f(z)) + |f'(z)|^2 \right] dz.$$

§21.2.1 The Leap of Faith and Sets of Finite Perimeter

Minimizing over all possible profiles f, the sharp lower bound is dictated by the optimal 1D energy cost, $C_W = \inf \left\{ \int_{-\infty}^{\infty} \left(W(v) + |v'|^2 \right) dz : v(\pm\infty) = \pm 1 \right\}$. This yields the expected scaling $\mathcal{F}_\varepsilon(u_\varepsilon) \gtrsim C_W \mathcal{H}^{d-1}(\Sigma)$.

Remark 21.2.2. In the heuristic above, we performed a "leap of faith" by assuming that the interface Σ is smooth enough to admit a tubular neighborhood and well-defined normal distances. In reality, sequences of bounded energy may converge to characteristic functions $1_E - 1_{E^c}$ where the set E is highly irregular.

To formalize this in Calculus of Variations, we require the geometric measure theory of **Sets of Finite Perimeter** (Caccioppoli sets). The perimeter of a measurable set E in Ω is defined via distributional derivatives as:

$$\mathrm{Per}(E,\Omega) = \sup \left\{ \int_E \mathrm{div}\varphi\, dx : \varphi \in C_c^1(\Omega;\mathbb{R}^d), \|\varphi\|_\infty \le 1 \right\}.$$

This definition does not require ∂E to be a manifold. If E has a smooth boundary, $\text{Per}(E, \Omega) = \mathcal{H}^{d-1}(\partial E \cap \Omega)$. By De Giorgi's structure theorem, sets of finite perimeter have a well-defined "reduced boundary" $\partial^* E$, over which the Coarea formula and integration by parts rigorously hold.

This framework leads directly to the seminal result for phase transitions:

Theorem 21.2.3 (Modica-Mortola)
Let $W(u) = (1-u^2)^2/4$, and for any open set Ω with smooth boundary, consider the functional:

$$\mathcal{F}_\varepsilon(u) = \begin{cases} \displaystyle\int_\Omega \left(\frac{1}{\varepsilon} W(u) + \varepsilon |Du|^2 \right) dx & \text{if } u \in W^{1,2}(\Omega; \mathbb{R}), \\ +\infty & \text{otherwise.} \end{cases}$$

Then $\mathcal{F}_\varepsilon \xrightarrow{\Gamma} \mathcal{F}$ in the $L^1_{loc}(\Omega)$ topology, where the effective functional is:

$$\mathcal{F}(u) = \begin{cases} C_W \cdot \text{Per}(\{u = 1\}, \Omega) & \text{if } u = 1_E - 1_{E^c} \text{ for some } E \subseteq \Omega, \\ +\infty & \text{otherwise.} \end{cases}$$

Furthermore, if $\{u_\varepsilon\}$ is a sequence of uniformly bounded energy, i.e., $\sup_\varepsilon \mathcal{F}_\varepsilon(u_\varepsilon) \leq C$, then there exists $E \subseteq \Omega$ such that, up to a subsequence, $u_\varepsilon \xrightarrow{a.e.} 1_E - 1_{E^c}$ and $\text{Per}(E, \Omega) < \infty$.

§21.3 1D Case for Modica-Mortola

Let us focus on the one-dimensional case over an interval $I = (a, b)$. The functional simplifies to $\mathcal{F}_\varepsilon(u) = \displaystyle\int_I \left(\frac{1}{\varepsilon} W(u) + \varepsilon |u'|^2 \right) dx$. The target Γ-limit restricts u to piecewise constant functions:

$$\mathcal{F}(u) = \begin{cases} C_W \cdot \#J(u) & \text{if } u \in P.C.(I; \{-1, 1\}), \\ +\infty & \text{otherwise,} \end{cases}$$

where $P.C.(I; \{-1, 1\})$ is the space of functions for which there exists a finite minimal partition $a = x_0 < x_1 < \cdots < x_n = b$ such that u is constant on (x_i, x_{i+1}), and $\#J(u) = n$ counts the number of jumps.

Exercise 21.3.1
Let $u_k \in P.C.(I; \{-1, 1\})$ and assume the number of jumps is uniformly bounded, $\sup_k \#J(u_k) \leq C$. Prove that, up to a subsequence, there exists $u \in P.C.(I; \{-1, 1\})$ such that $u_k \xrightarrow{a.e.} u$ and the lower semicontinuity holds:

$$\#J(u) \leq \liminf_{k \to \infty} \#J(u_k).$$

Try to prove this constructively without explicitly invoking Helly's Selection Theorem.

To compute the intrinsic cost of a single jump C_W, we note that $C_W(T) = \inf \int_{-T}^{T} (W(f) + |f'|^2)\, dx \geq C_W(+\infty)$ naturally decreases as the interval expands. By the AM-GM inequality (or Young's inequality):

$$\int_{-T}^{T} \left(W(f) + |f'|^2\right)\, dx \geq 2 \int_{-T}^{T} \sqrt{W(f)}|f'|\, dx.$$

Assuming f is monotone increasing across the jump ($f(-T) = -1, f(T) = 1$), we can apply the change of variables $t = f(x)$ to obtain:

$$\int_{-T}^{T} \left(W(f) + |f'|^2\right)\, dx \geq \int_{-1}^{1} 2\sqrt{W(t)}\, dt = C_W.$$

Equality holds strictly when the profile satisfies the ODE $f'(x) = \sqrt{W(f(x))}$.

For $W(t) = \dfrac{(1-t^2)^2}{4}$, solving $f' = \dfrac{1-f^2}{2}$ with $f(0) = 0$ yields the optimal transition profile $f(z) = \tanh(z/2)$.

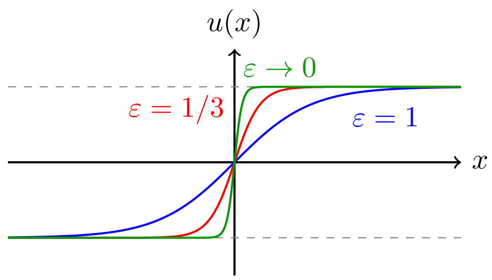


Figure 21.4: Optimal 1D transition profiles $f_\varepsilon(x) = \tanh(x/2\varepsilon)$ sharpening into a discrete phase jump as $\varepsilon \to 0$. The geometric energy cost of this layer is uniformly C_W.

22 Proof of Modica-Mortola in 1D

In this lecture, we rigorously establish the compactness, the Γ-liminf inequality, and the Γ-limsup inequality for the 1D Modica-Mortola functional over an interval $I = (a, b)$:

$$\mathcal{F}_\varepsilon(u) = \int_a^b \left(\frac{1}{\varepsilon} W(u) + \varepsilon |u'|^2 \right) dx.$$

§22.1 Equicoercivity and Γ-liminf Inequality

Theorem 22.1.1 (Compactness and Lower Bound)
Let $\{u_\varepsilon\}$ be a sequence such that $\sup_{\varepsilon>0} \mathcal{F}_\varepsilon(u_\varepsilon) \leq C$. Then, up to a subsequence, there exists a piecewise constant function $u \in P.C.((a, b); \{-1, 1\})$ such that $u_\varepsilon \to u$ in measure. Furthermore, the following Γ-liminf inequality holds:

$$C_W \# J(u) \leq \liminf_{\varepsilon \to 0} \mathcal{F}_\varepsilon(u_\varepsilon).$$

Proof. Let $\{u_\varepsilon\}$ be a sequence with uniformly bounded energy. We extract a subsequence $\{u_{\varepsilon_k}\}$ such that $\lim_{k\to\infty} \mathcal{F}_{\varepsilon_k}(u_{\varepsilon_k}) = \liminf_{\varepsilon\to 0} \mathcal{F}_\varepsilon(u_\varepsilon)$. Without loss of generality, via truncation, we may assume $|u_{\varepsilon_k}| \leq 1$ almost everywhere (otherwise, considering $\max\{\min\{u_{\varepsilon_k}, 1\}, -1\}$ strictly decreases the energy).

Step 1: Convergence in Measure.
The uniform energy bound implies:

$$\int_a^b W(u_{\varepsilon_k})\, dx \leq \mathcal{F}_{\varepsilon_k}(u_{\varepsilon_k}) \cdot \varepsilon_k \leq C\varepsilon_k \xrightarrow{k\to\infty} 0.$$

Because $W(t)$ vanishes only at $t \in \{-1, 1\}$, this guarantees that the sequence is driven towards the pure phases. Specifically, for any threshold $\eta > 0$:

$$|\{x \in I : \operatorname{dist}(u_{\varepsilon_k}(x), \{-1, 1\}) \geq \eta\}| \xrightarrow{k\to\infty} 0.$$

Step 2: Dyadic Partitioning and Bad Intervals.
Fix $N \in \mathbb{N}$ and divide (a, b) into 2^N subintervals of length $\dfrac{b-a}{2^N}$, denoted as $\{I_j^N\}_{j=1}^{2^N}$. For a fixed small $\eta > 0$, we classify these subintervals into two sets based on the oscillation of u_{ε_k}. We define the set of "Bad" intervals $\mathcal{B}_k^N$ as:

$$\mathcal{B}_k^N = \left\{ \tilde{I} \in \{I_j^N\} : \exists x, y \in \tilde{I} \text{ such that } u_{\varepsilon_k}(x) \geq 1 - 2\eta \text{ and } u_{\varepsilon_k}(y) \leq -1 + 2\eta \right\}.$$

The "Good" intervals $\mathcal{G}_k^N$ consist of all remaining intervals in the partition.

For any bad interval $\tilde{I} \in \mathcal{B}_k^N$, we can apply the AM-GM inequality to bound the local energy. Assuming without loss of generality that $x < y$:

$$\int_{\tilde{I}} \left(\frac{W(u_{\varepsilon_k})}{\varepsilon_k} + \varepsilon_k |u'_{\varepsilon_k}|^2 \right) dx \geq 2 \left| \int_x^y \sqrt{W(u_{\varepsilon_k})} |u'_{\varepsilon_k}|\, dt \right| = 2 \left| \int_{u_{\varepsilon_k}(x)}^{u_{\varepsilon_k}(y)} \sqrt{W(t)}\, dt \right|.$$

Given the bounds on x and y, the transition must cross $[-1+2\eta, 1-2\eta]$. Hence:

$$\int_{\tilde{I}} \left(\frac{W(u_{\varepsilon_k})}{\varepsilon_k} + \varepsilon_k |u'_{\varepsilon_k}|^2 \right) dx \geq 2 \int_{-1+2\eta}^{1-2\eta} \sqrt{W(s)}\, ds \geq C_W + \mathcal{O}(\eta).$$

For sufficiently small η, we have $C_W + \mathcal{O}(\eta) > \dfrac{C_W}{2}$. Since the intervals are disjoint, summing over all bad intervals yields:

$$\frac{C_W}{2} \cdot \#\mathcal{B}_k^N \leq \sum_{\tilde{I} \in \mathcal{B}_k^N} \mathcal{F}_{\varepsilon_k}(u_{\varepsilon_k}, \tilde{I}) \leq \mathcal{F}_{\varepsilon_k}(u_{\varepsilon_k}, I) \leq C.$$

This implies that the number of bad intervals $\#\mathcal{B}_k^N$ is uniformly bounded independently of N. Consequently, the total measure of the bad region $C_k^N = \bigcup_{\tilde{I} \in \mathcal{B}_k^N} \tilde{I}$ satisfies $|C_k^N| = \mathcal{O}(2^{-N})$.

Step 3: Piecewise Constant Approximation.
We construct an approximating sequence $U_{N,k}(x)$ on the grid:

$$U_{N,k}(x) = \begin{cases} 1 & \text{if } x \in \tilde{I} \in \mathcal{G}_k^N \text{ and } \exists y \in \tilde{I} \text{ s.t. } u_{\varepsilon_k}(y) \geq 1 - 2\eta, \\ -1 & \text{if } x \in \tilde{I} \in \mathcal{G}_k^N \text{ and } \exists y \in \tilde{I} \text{ s.t. } u_{\varepsilon_k}(y) \leq -1 + 2\eta, \\ 0 & \text{otherwise (in bad intervals).} \end{cases}$$

By construction, outside the bad intervals of measure $\mathcal{O}(2^{-N})$, the deviation $|u_{\varepsilon_k} - U_{N,k}|$ is large only if u_{ε_k} strays from the pure phases, which we know vanishes in measure. For fixed N, there exists $K(N)$ such that for all $k \geq K(N)$, the total deviation measure is bounded by $1/N$.

Defining the diagonal sequence $V_N = U_{N,K(N)}$, we observe that the sets $\{V_N = 1\}$ and $\{V_N = -1\}$ stabilize, and $|\{V_N = 0\}| = \mathcal{O}(2^{-N}) \to 0$. Thus, $V_N \to u$ in measure. By a standard diagonalization argument, we extract a subsequence such that $u_{\varepsilon_k} \to u$ almost everywhere. Because the number of transitions (bad intervals) is uniformly bounded, $u \in P.C.((a,b); \{-1,1\})$.

Step 4: Liminf Bound on Jumps.
Given the piecewise constant limit u, let $\{x_1, \dots, x_n\}$ be its minimal partition of jump points. We can choose a grid sufficiently fine such that each jump is isolated within distinct intervals. Applying the exact same 1D local energy lower bound across each jump isolating interval confirms that each jump contributes at least C_W to the total energy. Thus:

$$C_W \cdot \#J(u) \leq \liminf_{k \to \infty} \mathcal{F}_{\varepsilon_k}(u_{\varepsilon_k}).$$

■

Remark 22.1.2. The underlying topological principle of the 1D liminf inequality relies on the fact that any continuous transition between the phases must pay the minimum energetic cost of the profile. Specifically:

$$\int_A^B \left(\frac{W(u)}{\varepsilon} + \varepsilon |u'|^2 \right) dx \geq 2 \int_{\min u}^{\max u} \sqrt{W(s)}\, ds \geq f(\operatorname{osc}(u, I)),$$

where f is a monotonically increasing function of the function's oscillation over the interval.

§22.2 Γ-limsup Inequality and Optimal Profile Construction

To prove the Γ-limsup, we must construct a recovery sequence $u_\varepsilon \xrightarrow{L^2} u$ that asymptotically matches the energy bound. We construct this sequence by carefully "gluing" rescaled optimal profiles into the jump discontinuities of the target piecewise constant function u.

Construction of the Recovery Sequence. Let $u \in P.C.((a,b);\{-1,1\})$. For simplicity, assume u has a single jump at $x = c$, transitioning from -1 to 1.

By definition of the optimal profile cost, $C_W = \inf\left\{\int_{-\infty}^{\infty} (W(v) + |v'|^2)\,dz : v(\pm\infty) = \pm 1\right\}$. For any arbitrary tolerance $\delta > 0$, we can find a large finite interval $[-T, T]$ and a smooth truncated profile $\bar{u} : [-T, T] \to \mathbb{R}$ such that:

1. $\bar{u}(0) = 0$,
2. $\bar{u}(T) \geq 1 - \delta$ and $\bar{u}(-T) \leq -1 + \delta$,
3. The total energy approximates the optimal cost: $\displaystyle\int_{-T}^{T} (W(\bar{u}) + |\bar{u}'|^2)\,dz \leq C_W + \delta$.

We now construct the recovery sequence u_ε by physically rescaling this optimal profile by a factor of ε and translating it to the jump point c:

$$u_\varepsilon(x) = \begin{cases} \bar{u}(T) & \text{if } x \in [c + \varepsilon T, b], \\ \bar{u}\left(\dfrac{x - c}{\varepsilon}\right) & \text{if } x \in [c - \varepsilon T, c + \varepsilon T], \\ \bar{u}(-T) & \text{if } x \in [a, c - \varepsilon T]. \end{cases}$$

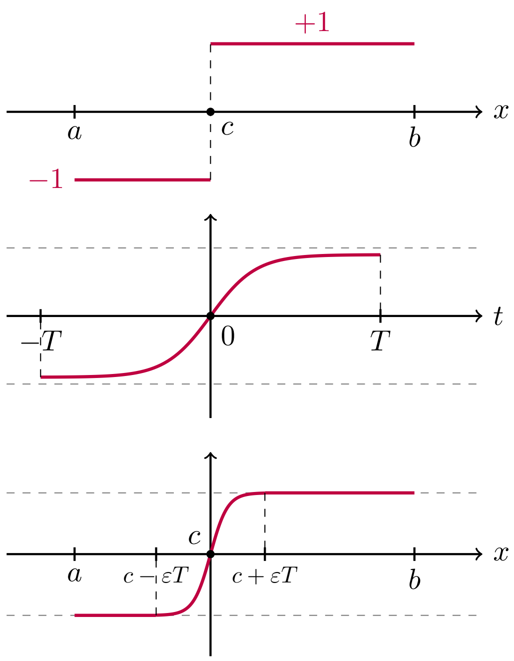


Figure 22.1: Construction of the Γ-limsup recovery sequence. Top: The target macroscopic step function u. Middle: The truncated optimal profile $\bar{u}$ constructed over $[-T, T]$. Bottom: The rescaled profile u_ε compressed into a phase transition layer of width $2\varepsilon T$ centered at c.

Evaluating the energy of this constructed sequence, the contributions from the flat extensions are identically zero (since derivatives vanish and the potential is uniformly small away from the interface). The energy is completely localized to the transition layer. By applying the change of variables $z = \frac{x-c}{\varepsilon}$:

$$\begin{aligned}\mathcal{F}_\varepsilon(u_\varepsilon) &= \int_{c-\varepsilon T}^{c+\varepsilon T} \left(\frac{1}{\varepsilon} W\left(\bar{u}\left(\frac{x-c}{\varepsilon}\right)\right) + \varepsilon \left| \frac{1}{\varepsilon} \bar{u}'\left(\frac{x-c}{\varepsilon}\right)\right|^2 \right) dx \\ &= \int_{-T}^{T} \left(W(\bar{u}(z)) + |\bar{u}'(z)|^2 \right) dz.\end{aligned}$$

By our selection of $\bar{u}$, this yields $\mathcal{F}_\varepsilon(u_\varepsilon) \le C_W + \delta$.

To verify L^2 convergence, we bound the L^1 difference (which implies L^2 via dominated convergence since the functions are bounded):

$$\int_a^b |u_\varepsilon - u|\, dx \le 4\varepsilon T + \delta(b-a).$$

As $\varepsilon \to 0$, the L^1 error is dominated by δ. By employing a diagonal selection argument, mapping $\delta(\varepsilon) \to 0$ sufficiently slowly as $\varepsilon \to 0$, we select $u_\varepsilon = u_\varepsilon^{\delta(\varepsilon)}$ as the precise recovery sequence. This strictly guarantees $u_\varepsilon \to u$ and $\limsup_{\varepsilon\to 0} \mathcal{F}_\varepsilon(u_\varepsilon) \le C_W$, concluding the proof. ■

Bibliography

[ACM18] Luigi Ambrosio, Alessandro Carlotto, and Annalisa Massaccessi, *Lectures on elliptic partial differential equations*, Appunti. Scuola Normale Superiore di Pisa (Nuova Serie), vol. 18, Edizioni della Normale, Pisa, 2018.

[BGH98] Giuseppe Buttazzo, Mariano Giaquinta, and Stefan Hildebrandt, *One-dimensional variational problems: An introduction*, Oxford Lecture Series in Mathematics and Its Applications, vol. 15, Oxford University Press, 1998.

[Bra02] Andrea Braides, *Gamma-convergence for beginners*, Oxford Lecture Series in Mathematics and its Applications, vol. 22, Oxford University Press, Oxford, 2002.

[Bra24] Lorenzo Brasco, *Handbook of calculus of variations for absolute beginners*, Lecture Notes, 2024.

[Bre11] Haim Brezis, *Functional analysis, sobolev spaces and partial differential equations*, Universitext, Springer, New York, 2011.

[CV10] Luis A. Caffarelli and Alexis F. Vasseur, *The de giorgi method for regularity of solutions of elliptic equations and its applications to fluid dynamics*, Discrete Contin. Dyn. Syst. Ser. S **3** (2010), no. 3, 409–427.

[Dac07] Bernard Dacorogna, *Direct methods in the calculus of variations*, 2nd ed., Applied Mathematical Sciences, vol. 78, Springer, 2007.

[DG57] Ennio De Giorgi, *Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari*, Memorie della Accademia delle Scienze di Torino. Classe di Scienze Fisiche, Matematiche e Naturali **3** (1957), no. 3, 25–43.

[Eva10] Lawrence C. Evans, *Partial differential equations*, 2nd ed., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, Rhode Island, 2010.

[GH96] Mariano Giaquinta and Stefan Hildebrandt, *Calculus of variations i: The lagrangian formalism*, Grundlehren der mathematischen Wissenschaften, vol. 310, Springer, 1996.

[GMS98] Mariano Giaquinta, Giuseppe Modica, and Jiří Souček, *Cartesian currents in the calculus of variations i: Cartesian currents*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics, vol. 37, Springer, 1998, Chapter 1.

[Rin18] Filip Rindler, *Calculus of variations*, Universitext, Springer, 2018.

[San23] Filippo Santambrogio, *A course in the calculus of variations: Optimization, regularity, and modeling*, Universitext, Springer, 2023.