The 19^{th} Problem: Part II

The Analytical Toolkit

A Leisurely Excursion

Introduction

In part I, for $n = 2$ case, we relied on a "happy accident of geometry." But as we step into dimensions greater than 2, this geometric safety net vanishes, for instance we saw in $n = 3$ case the case chopping argument fails as the energy integral converges even for singular functions without violating the global energy bounds. Furthermore, the "algebraic slack" in the linearized equation allows the gradient map to stretch space infinitely in one direction while remaining valid.

We are thus left in a position where the geometry is too flexible to force regularity. To solve Hilbert's problem in higher dimensions, we must abandon the "shape" of the gradient and instead look at the "mass" of the gradient. We need to stop asking "what does it look like?" and start asking "how much energy does it carry?".

This requires a new set of tools. Before we can appreciate the final solution, we must establish the prerequisites of some analytical toolkit: Weak Derivatives, Sobolev Spaces, and the Caccioppoli Inequality.

Weak Derivatives

Before going into what weak derivatives are, we need two ideas that extend what we learned in our introductory real analysis or honors calculus class:

Definition 1 (Locally Integrable)**.** *When we work with weak derivatives, we don't need the whole function to be integrable over an open subset, $\Omega \subseteq \mathbb{R}^n$. It's enough that for all compact interval K, the integral $\int_K |u(x)|\, dx$ exists. These functions are called locally integrable functions. We represent the set of such functions by $L^1_{loc}(\Omega)$.*

L^1 just means the absolute value of the function has a finite integral.

Definition 2 (Test Functions)**.** *A function $\phi : \mathbb{R}^n \to \mathbb{R}$ is a test function if:*

1. *Smoothness: ϕ is infinitely differentiable, i.e., $\phi \in C^\infty(\mathbb{R}^n)$.*
2. *Compact Support: There exists a compact set $K \subset \mathbb{R}^n$ such that*

$$\phi(x) = 0, \qquad \forall x \notin K.$$

Note that in $\mathbb{R}^n$, "compact" means the set is closed and bounded. We write the space of test functions as $C_c^\infty(\Omega)$, where Ω is an open subset of $\mathbb{R}^n$.

Before moving on, let us work with $u(x) = |x|$ and suppose for some $g(x)$ we want to find

$$\int_{-\infty}^{\infty} (|x|\; g(x))'\, dx.$$

Now, we have,

$$\int_{-\infty}^{\infty} (|x|\ g(x))'\,dx = \int_{-\infty}^{\infty} (|x|)' g(x)\,dx + \int_{-\infty}^{\infty} |x| g(x)'\,dx$$

The middle term seem to be problematic due to the presence of $(|x|)'$, and it definitely is. Does that mean problems of this kind are unsolvable? Fortunately, we are not completely out of luck; if $g(x)$ is any test function $\phi \in C_c^\infty(\mathbb{R})$, we immediately observe that

$$[|x|\phi(x)]_{-\infty}^{\infty} = \int_{-\infty}^{\infty} (|x|)'\phi\,dx + \int_{-\infty}^{\infty} |x|\phi'\,dx$$

[1]

$$\int_{-\infty}^{\infty} (|x|)'\phi(x)\,dx = -\int_{-\infty}^{\infty} |x|\phi'(x)\,dx \tag{1}$$

Now, let us try solving the right-hand side where our burden of having well-defined derivatives/smoothness transfers from $|x|$ to our test function $\phi(x)$, which, by definition, is smooth.

Now, the right-hand side of **??** can be written as,

$$= -\left(\int_{-\infty}^{0} (-x)\phi'(x)\,dx + \int_{0}^{\infty} x\phi'(x)\,dx\right).$$

Now for $x < 0$,

$$\int_{-\infty}^{0} (-x)\phi'(x)\,dx = \underbrace{[-x\phi(x)]_{-\infty}^{0}}_{=0} - \int_{-\infty}^{0} (-1)\phi(x)\,dx = \int_{-\infty}^{0} \phi(x)\,dx.$$

Similarly, for $x > 0$,

$$\int_{0}^{\infty} x\phi'(x)\,dx = \underbrace{[x\phi(x)]_{0}^{\infty}}_{=0} - \int_{0}^{\infty} 1\cdot\phi(x)\,dx = -\int_{0}^{\infty} \phi(x)\,dx.$$

Now,

$$\int_{-\infty}^{0} \phi(x)\,dx - \int_{0}^{\infty} \phi(x)\,dx = -\left(\int_{-\infty}^{0} (-1)\phi(x)\,dx + \int_{0}^{\infty} (1)\phi(x)\,dx\right) = -\int_{-\infty}^{\infty} \operatorname{sgn}(x)\phi(x)\,dx.$$

The single point $x = 0$ contributes nothing to the integral, so we don't need to define $\operatorname{sgn}(0)$.

So, we found that

$$\int_{-\infty}^{\infty} (|x|)'\phi(x)\,dx = -\int_{-\infty}^{\infty} |x|\phi'(x)\,dx = \int_{-\infty}^{\infty} \operatorname{sgn}(x)\phi(x)\,dx.$$

Just looking at the left-most and right-most integrals we can see something interesting; it looks like $(|x|)'$ is equal to $\operatorname{sgn}(x)$. Well, it is, but in a *weak* sense. As we already know that $(|x|)'$ does not exist everywhere so we cannot really say that $(|x|)' = \operatorname{sgn}(x),\ \ \forall x \in \mathbb{R}$. So, what do we do? We name $\operatorname{sgn}(x)$ as the *weak derivative.* These so-called *Weak Derivatives* are introduced

[1]By definition, $\exists$ a compact set $K \subset \mathbb{R}$ such that $\phi(x) = 0\ \ \forall x \notin K$. Since K is compact in $\mathbb{R}$, it is bounded, so $\exists R > 0$ such that $K \subseteq [-R, R]$ and $\phi(x) = 0$ for all $|x| > R$. Now the boundary term is

$$[|x|\phi(x)]_{-\infty}^{\infty} = \lim_{a\to\infty} |a|\phi(a) - \lim_{b\to-\infty} |b|\phi(b).$$

For $a > R$, $a \notin K$, so $|a|\phi(a) = a \cdot 0 = 0$, hence $\lim_{a\to\infty} |a|\phi(a) = 0$. For $b < -R$, $b \notin K$, so $|b|\phi(b) = -b \cdot 0 = 0$, hence $\lim_{b\to-\infty} |b|\phi(b) = 0$. Thus, the boundary term vanishes.

so that there is a way to talk about "slope" as in our above example without breaking the rules of calculus. In physics and in variational calculus, we often deal with functions similar to $|x|$ and the idea here is to shift the burden of smoothness from the function u (which might be rough) onto a "test function" ϕ. This motivates our definition for weak derivatives:

Definition 3 (Weak Derivative in $\mathbb{R}$)**.** *Let u be a locally integrable function. We say a function v is the weak derivative of u if, for every test function $\phi \in C_c^\infty(\Omega)$ where $\Omega \subseteq \mathbb{R}$, the following equality holds:*

$$\int_\Omega u(x)\phi'(x)\,dx = -\int_\Omega v(x)\phi(x)\,dx$$

If such a v exists, it is essentially unique, and we write $u' = v$ in the weak sense.

Definition 4 (Weak Derivative in $\mathbb{R}^n$)**.** *Let $u \in L^1_{loc}(\Omega)$, where $\Omega \subseteq \mathbb{R}^n$ is an open set and α be a multi index[2]. We say a locally integrable function $v \in L^1_{loc}(\Omega)$ is the weak α-derivative of u if, for every test function $\phi \in C_c^\infty(\Omega)$, the following holds:*

$$\int_\Omega u(x)D^\alpha\phi(x)\,dx = (-1)^{|\alpha|}\int_\Omega v(x)\phi(x)\,dx$$

If such a v exists, it is essentially unique, and we write $D^\alpha u = v$ in the weak sense.

Now, let us do a specific example: Let $u(x_1, x_2) = |x_1|x_2$ in $\mathbb{R}^2$. Let us find its weak partial derivative with respect to x_1 which is $\alpha = (1, 0)$, so $D^\alpha\phi = \frac{\partial\phi}{\partial x_1}$ and $|\alpha| = 1$.

Now by using Fubini's Theorem, we have:

$$\int_{\mathbb{R}^2} u(x)D^\alpha\phi(x)\,dx = \int_{-\infty}^{\infty} x_2\left(\int_{-\infty}^{\infty} |x_1|\frac{\partial\phi}{\partial x_1}\,dx_1\right)dx_2$$

The inner integral is exactly the 1D case we already did. Thus:

$$\int_{-\infty}^{\infty} x_2\left(-\int_{-\infty}^{\infty} \operatorname{sgn}(x_1)\phi(x_1, x_2)\,dx_1\right)dx_2$$

$$= -\int_{\mathbb{R}^2}\big(\operatorname{sgn}(x_1)x_2\big)\phi(x)\,dx_1dx_2$$

By comparing this to the right-hand side of our definition, $(-1)^{|\alpha|}\int v\phi\,dx$, we immediately see that the weak derivative is $v(x_1, x_2) = \operatorname{sgn}(x_1)x_2$.

In our above examples, both $|x_1|x_2$ and $|x|$ have a weak derivative as it fails to be classically differentiable at a single point for $|x|$, and the line $x_1 = 0$ for $|x_1|x_2$. Sets like these, where classical differentiation is undefined, have *measure zero*[3], so the integral isn't affected and thus our definition stays intact. You might think that if a function is differentiable everywhere apart from a set of measure zero, then its classical derivative (where defined) is automatically its weak derivative. However, this is not the case, a counterexample is the Cantor function $f : [0, 1] \to [0, 1]$, which is continuous, non-decreasing, and differentiable almost everywhere with $f' = 0$ a.e. Yet, it does not have a weak derivative in L^1_{loc}. If $f' = 0$ were the weak derivative, then $\int_0^1 f\phi' dx$ would have to equal 0 for every test function ϕ, but this is false since f increases from 0 to 1.

Overall the concept of weak derivatives allows us to do calculus on functions that would otherwise be "illegal."

[2]An $n-$dimensional multi-index is a $n-$tuple of non-negative integers,$\alpha = (\alpha_1, \ldots, \alpha_n)$ and $|\alpha| = \alpha_1 + \ldots + \alpha_n$. For example if $\alpha = (2, 1)$, then $|\alpha| = 3$ and we have $D^\alpha u = \frac{\partial^3 u}{\partial x_1^2 \partial x_2^1}$.

[3]Roughly speaking, a set of measure zero means the set is so small that it does not affect the integral value. For instance, any finite set of points in $\mathbb{R}$ has measure zero and in $\mathbb{R}^2$ any line has measure zero.

The Sobolev Spaces

Now that we can differentiate rough functions (in the weak sense), we need a space to put them. We need a vector space of functions that have "finite energy", a space where we can measure the size of both a function and its derivatives.

Measurable Functions

Let us take a small digression to measurable functions as it is pretty fundamental to understand what we are doing.

As we already discussed we do not measure points because they have 0 measure, so we only measure subsets. We therefore define a family $\mathcal{B}$ of subsets of Ω that we can measure. *Measurable sets* are just any sets which belongs to $\mathcal{B}$. These are the sets we can consistently assign a size to, built from rectangles (products of intervals) using countable unions, intersections, and complements

If we have function $u : \Omega \to \mathbb{R}$, we want to transport this measure to $\mathbb{R}$. To find measure of outcome set $A \subset \mathbb{R}$, we must look at inverse image $u^{-1}(A)$ in Ω. For this to work, $u^{-1}(A)$ must be in $\mathcal{B}$. Since we care about intervals $I \subset \mathbb{R}$, we want $u^{-1}(I)$ to be measurable. This is same as saying $u^{-1}([-\infty, a))$[4] is measurable $\forall a \in A$.

So intuitively, measurable function just guarantees inverse image of intervals are measurable, which allows us to transport measure and compute integral.

The L^p Spaces

First, we need a way to measure how big a function is. It is convenient to measure size using integrals rather than maximum values (which are too delicate; changing a function at a single point can ruin its supremum, but doesn't affect integrals because points have measure zero).

For $1 \leq p < \infty$, we define the space $L^p(\Omega)$, where Ω is an open subset in $\mathbb{R}^n$, as the set of measurable functions such that

$$\|u\|_{L^p} = \left(\int_\Omega |u|^p \, dx \right)^{1/p} < \infty.$$

(When $p = 2$, this is like an infinite-dimensional version of the Euclidean length $\sqrt{\sum u_i^2}$.)

Remark 1. *Strictly speaking, L^p consists of equivalence classes of functions where two functions are considered the same if they differ only on a set of measure zero.*

Definition 5 (Sobolev Space)**.** *The Sobolev space $W^{1,p}(\Omega)$ contains every function $u \in L^p(\Omega)$ whose first weak derivatives $\frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_n}$ also belong to $L^p(\Omega)$. We measure functions in this space with the norm*

$$\|u\|_{W^{1,p}} = \left(\|u\|_{L^p}^p + \sum_{i=1}^n \left\| \frac{\partial u}{\partial x_i} \right\|_{L^p}^p \right)^{1/p}.$$

More generally, one can define $W^{k,p}(\Omega)$ by requiring weak derivatives up to order k to be in L^p, but for our purposes $k = 1$ suffices. In-fact, $W^{1,2}(\Omega)$ is important to us. It is often

[4] When we deal with stuff related to measure theory, we usually work in the extended real number system, $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ as sometimes it is useful to allow sets that have infinite measure and integrals whose value is infinite.

written as H^1, which is a *Hilbert space*. The minimizer in Hilbert's 19^{th} problem we are trying to study *initially* lives in H^1. This space is also *complete*[5], therefore once we prove the sequence of approximate minimizers is a Cauchy sequence, we can be sure that their limit, our desired minimizer, exists in the space itself.

A nice thing about *Hilbert space* is that the norm comes from an inner product:

$$\langle u, v\rangle_{H^1} = \int_\Omega uv\,dx + \int_\Omega \nabla u \cdot \nabla v\,dx.$$

Think of it like the dot product between *vectors* that you already know, we are doing the same thing here. As we are in a vector space and these functions are the elements of that vector space, so functions are vectors.

Example: Are you sure?

Consider $\Omega = (0,1) \subset \mathbb{R}$. Let:

$$u(x) = x \quad \text{and} \quad w(x) = \left|x - \frac{1}{2}\right|.$$

Notice that the function w is not differentiable in the classical sense at $x = \frac{1}{2}$ but we do have the weak derivative which is $\operatorname{sgn}(x - \frac{1}{2})$. This weak derivative is a step function which is discontinuous, but certainly in $L^2(0,1)$ since $\int_0^1 (\operatorname{sgn}(x - \frac{1}{2}))^2 dx = 1$.

We now verify that both u and w live in $H^1(0,1) = W^{1,2}(0,1)$:
For u:

$$\|u\|_{H^1}^2 = \int_0^1 u(x)^2\,dx + \int_0^1 u'(x)^2\,dx = \frac{4}{3}.$$

For w:

$$\|w\|_{H^1}^2 = \int_0^1 w(x)^2\,dx + \int_0^1 w'(x)^2\,dx = \int_0^1 \left|x - \frac{1}{2}\right|^2 dx + \int_0^1 (\operatorname{sgn}(x - \frac{1}{2}))^2\,dx.$$

The first integral is symmetric about $x = \frac{1}{2}$, so

$$\int_0^1 \left|x - \frac{1}{2}\right|^2 dx = 2\int_0^{1/2} \left(\frac{1}{2} - x\right)^2 dx = \frac{1}{12}.$$

The second integral is $\int_0^1 (\operatorname{sgn}(x - \frac{1}{2}))^2 dx = 1$. Hence

$$\|w\|_{H^1}^2 = \frac{13}{12}.$$

Both norms are finite, so $u, w \in H^1(0,1)$.

This example is mainly to show that the "rough" functions can live in H^1 so we do not have to worry if the approximate minimizers themselves are rough or if their limit is rough.

A natural follow up is, how can we be sure that the minimizers are not rough functions which live outside H^1? as there exists rough functions which don't live in H^1 for example consider the step function on $\Omega = (0,1)$:

$$h(x) = \begin{cases} 0 & \text{if } x < \frac{1}{2}, \\ 1 & \text{if } x > \frac{1}{2}. \end{cases}$$

[5]When we say that a space is complete we mean that for a given cauchy sequence in the space, the limit is also in the space.

For any test function $\phi \in C_c^\infty(0,1)$,

$$\int_0^1 h(x)\phi'(x)\,dx = \int_{\frac{1}{2}}^1 \phi'(x)\,dx = -\phi(\frac{1}{2}) = -\int_{-\infty}^{\infty} v(x)\phi(x)\,dx$$

$$\int_0^1 \delta_{\frac{1}{2}}(x)\phi(x)\,dx = \phi(\frac{1}{2}) = \int_{-\infty}^{\infty} v(x)\phi(x)\,dx$$

where $v(x)$ is the weak derivative. Thus, the weak derivative in this case is the Dirac delta distribution, $h' = \delta_{\frac{1}{2}}(x)$.

Suppose for contradiction, that $h \in H^1(0,1)$ then $\delta_{\frac{1}{2}} \in L^2(0,1)$.But $\int_0^1 |\delta_{\frac{1}{2}}(x)|^2\,dx$ is infinite, which is a contradiction and hence $h \notin H^1(0,1)$.

In this case, notice that $J(u)$ blows up which is not something we are looking at (if it blows up or is undefined, the problem of finding minimizer itself becomes meaningless). In general, if a rough function is outside the H^1 space, then $J(u)$ blows up or is undefined.

So to answer that question: the problem (the functional $J(u)$) itself ensures us that the rough functions which may encounter will certainly live in H^1.

Sobolev Inequality

Another natural follow up question is: could an H^1 function be really wild? Like, discontinuous everywhere? Unbounded? Oscillating infinitely many times? To answer that we utilize Sobolev inequality.

Definition 6 (Sobolev Conjugate)**.** *For $1 \leq p < n$, the number*

$$p^* = \frac{np}{n-p}$$

is called the Sobolev conjugate of p. Notice $p^* > p$.

Theorem 1 (Sobolev Inequality)**.** *Let $\Omega \subset \mathbb{R}^n$ be bounded and open. For $1 \leq p < n$, there exists a constant $C = C(p,n,\Omega)$ such that for every $u \in W^{1,p}(\Omega)$,*

$$\|u\|_{L^{p^*}(\Omega)} \leq C\|u\|_{W^{1,p}(\Omega)}.$$

To get a sense of this inequality, let us first start with a familiar concept - The Fundamental Theorem of Calculus.

$$u(x) = u(a) + \int_a^x u'(t)\,dt.$$

If $u' \in L^p(0,1)$, then from Hölder's inequality,

$$|u(x) - u(a)| \leq \left|\int_a^x u'(t)\,dt\right| \leq \|u'\|_{L^p}|x-a|^{1/q}, \quad \text{where} \quad \frac{1}{p} + \frac{1}{q} = 1.$$

What we have here is, we started with the information $u' \in L^p(0,1)$ and then we ended with the fact that $u \in L^\infty$ since u is Hölder continuous. Similarly, Theorem **??** says that if we start with $\nabla u \in L^p$, then we end up with $u \in L^{p^*}$ where $p^* > p$.

Take $p = 2$ (our H^1 case). Then the Sobolev conjugate is

$$2^* = \frac{2n}{n-2}.$$

Let's check what this means in different dimensions: For $n = 3$: $2^* = \frac{6}{1} = 6$. So $H^1 \hookrightarrow L^6$. So in 3-dimension, we start with $\nabla u \in L^2$, then we end up with $u \in L^6$. Similarly in 2-dimensions, H^1 functions belong to L^q for every finite q, but not necessarily L^∞.

To answer our original question: while an H^1 function need not be bounded or continuous (especially in dimensions $n \geq 2$), the Sobolev inequality prevents it from being arbitrarily wild. It forces the function into a strictly better integrability class L^{p^*}, which prevents situations like being discontinuous everywhere or oscillating without control.

Caccioppoli Inequality

We now know that our minimizer lives in H^1, and the Sobolev inequality told us it actually lives in a better space than we thought. Now what? For that let us pause for a moment and again remember our main goal which is to see if we can make the minimizer u smooth.

A natural follow up question would be: we have better control over u itself. Can we turn that into better control over ∇u? Because if ∇u gets upgraded to a better space, then Theorem **??** applies again and u gets even better space to live in, and then ∇u gets even better, and so on ultimately achieving u is in L^∞. This is also what we commonly call as "bootstrapping".

The Caccioppoli inequality is the first step to make this idea work.

Theorem 2 (Caccioppoli Inequality)**.** *Let v be a solution to an elliptic equation of the form*

$$\sum_{i,j=1}^{n} \int_\Omega a_{ij}(x) \partial_i v \partial_j \phi \, dx = 0 \quad \text{for all } \phi \in C_c^\infty(\Omega),$$

where the coefficients a_{ij} are bounded and elliptic (meaning $\sum a_{ij}\xi_i\xi_j \geq \lambda|\xi|^2$ for some $\lambda > 0$). Then for any ball $B_r \subset \Omega$,

$$\int_{B_{\frac{r}{2}}} |\nabla v|^2 \, dx \leq \frac{C}{r^2} \int_{B_r} v^2 \, dx,$$

where C depends only on the ellipticity constants and dimension.

For us, $v = \partial_k u$, meaning v is the k^{th} component of ∇u. Notice that we are bootstrapping $v = \partial_k u$, not u itself. By improving the integrability of the gradient's gradient, we are indirectly improving the regularity of u.

Now what Theorem **??** is basically saying is if we know how large v is on a ball B_r, then on a smaller ball $B_{\frac{r}{2}}$ we can control its gradient. More precisely, since $v \in L^2$ (because $\nabla u \in L^2$), the right-hand side is finite. Thus, the left-hand side is finite, which means ∇v exists as a weak derivative and is in L^2.

But the Caccioppoli inequality alone only gives us $\nabla v \in L^2$, which we already knew. To really get the bootstrap going, we need something more, a reverse Hölder inequality combined with Gehring's lemma (see [**?**]):

$$\left(\frac{1}{|B_{\frac{r}{2}}|} \int_{B_{\frac{r}{2}}} |\nabla v|^p \right)^{\frac{1}{p}} \leq C \left(\frac{1}{|B_r|} \int_{B_r} |\nabla v|^2 \right)^{\frac{1}{2}}$$

with $p = 2 + \epsilon$ for some $\epsilon > 0$ and v is as described above in Theorem **??**. Now, this says that ∇v is in L^p. Now from Theorem **??**, we get v is in L^{p^*}. As $v = \partial_k u$, so it is pretty clear as how this improvement is for u also all the way till L^∞. Once $\nabla u \in L^\infty$, the gradient is essentially bounded, but not yet continuous.

Next step

Now that we have all the required tools and know that $\nabla u \in L^{\infty}$, we will move on to analyzing Hilbert's 19^{th} problem for n-dimensional case, where we will be using De Giorgi's Theorem and Schauder estimates to see if we can prove the analyticity of the minimizer.