LECTURE NOTES
TMA4305 PARTIAL DIFFERENTIAL EQUATIONS,
SPRING 2007

SIGMUND SELBERG

We have proved the existence of a weak solution of the Dirichlet problem for a class of linear second order elliptic differential operators by means of the Riesz representation theorem. This method is limited to linear PDEs. Now we shall look at another way of proving weak existence, which has the advantage that it can be applied also to non-linear PDEs. The idea is to obtain the solution as a minimizer of a functional, which typically represents the energy of the system. To implement this idea, we shall need the compactness results proved last week. Also, for simplicity we shall limit our attention to the Poisson equation, but the method has a wide applicability.

1. Directional derivatives and the Euler-Lagrange equation

Let X be a normed vector space, with norm $\|\cdot\|$, and assume that $F : X \to \mathbb{R}$. Such a map F is called a *functional*, but it should be emphasized that we do *not* assume it is linear.

Definition 1. The *directional derivative* of $F : X \to \mathbb{R}$ at a point $x \in X$ in the direction $v \in X$, is the number

$$D_v F(x) = \lim_{\varepsilon \to 0} \frac{F(x + \varepsilon v) - F(x)}{\varepsilon},$$

provided this limit exists.

Assume $D_v F(x)$ exists for all directions $v \in X$, for a given $x \in X$. Then x is called a *critical point* for F if

$$D_v F(x) = 0 \qquad \text{for all } v \in X. \tag{1}$$

This equation is called the *Euler-Lagrange equation* for the functional $F : X \to \mathbb{R}$.

Lemma 1. *Assume that $F : X \to \mathbb{R}$, and that at a point $x \in X$, the derivative $D_v F(x)$ exists for all directions $v \in X$. If in addition x is a local maximum or minimum point for F, then x is a critical point, i.e., $D_v F(x) = 0$ for all $v \in X$.*

Proof. Fix v, and define $\phi(t) = F(x + tv)$. Then by the definition of $D_v F(x)$,

$$\phi'(0) = \lim_{t \to 0} \frac{F(x + tv) - F(x)}{t} = D_v F(x).$$

Moreover, ϕ has a local maximum or minimum at $t = 0$, hence $\phi'(0) = 0$. □

2. The energy functional for the Poisson equation

From now on we specialize the discussion to the following functional that we shall use in our applications.

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain, set

$$X = H_0^1(\Omega),$$

let $f \in L^2(\Omega)$ be given, and define $F : X \to \mathbb{R}$ by

$$F(u) = \int_\Omega \frac{1}{2} |\nabla u|^2 + f u \, dx \qquad (u \in X).$$

Alternatively, we can write this as

$$F(u) = \frac{1}{2}(u,u)_1 + \langle f, u \rangle,$$

where $(\cdot,\cdot)_1$ is the (modified) inner product on $H_0^1(\Omega)$ (see the lecture notes from two weeks ago, §2.3), and $\langle \cdot,\cdot \rangle$ is the inner product on L^2. From this last formulation of F, we see that F is certainly well-defined. Note also that F is nonlinear. We can interpret F physically as an energy functional.

Fix an arbitrary $u \in X$. We want to calculate $D_v F(u)$, for any $v \in X$. First observe that, by properties of the inner products,

$$\begin{aligned} F(u+v) - F(u) &= \frac{1}{2}(u+v, u+v)_1 + \langle f, u+v \rangle - \frac{1}{2}(u,u)_1 - \langle f, u \rangle \\ &= \frac{1}{2}(u,u)_1 + (u,v)_1 + \frac{1}{2}(v,v)_1 + \langle f, u \rangle + \langle f, v \rangle - \frac{1}{2}(u,u)_1 - \langle f, u \rangle \\ &= (u,v)_1 + \frac{1}{2}(v,v)_1 + \langle f, v \rangle. \end{aligned} \tag{2}$$

From this we get

$$\frac{F(u+\varepsilon v) - F(u)}{\varepsilon} = (u,v)_1 + \frac{1}{2}\varepsilon(v,v)_1 + \langle f, v \rangle,$$

and hence, letting $\varepsilon \to 0$,

$$D_v F(u) = (u,v)_1 + \langle f, v \rangle. \tag{3}$$

The Euler-Lagrange equation (1) therefore becomes, in the present situation,

$$(u,v)_1 + \langle f, v \rangle = 0 \qquad \text{for all } v \in X = H_0^1(\Omega).$$

By density, it is enough to have this for all $v \in C_0^\infty(\Omega)$. Thus, we see that the Euler-Lagrange equation is nothing else than the weak formulation of the Poisson equation $\Delta u = f$ (see the lecture notes from two weeks ago, §2.3).

Therefore, if we can prove that F attains its minimum at some $u \in X$, then u must be a weak solution of the Dirichlet problem

$$\begin{cases} \Delta u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{4}$$

Our main task will be to implement this strategy. Note that this method of proving existence of a weak solution to a PDE is very general, and can be applied also to nonlinear problems (unlike the approach based on the Riesz representation theorem).

3. A lower bound for F, and a minimizing sequence

Let F be the functional defined in §2. To implement our strategy of obtaining a weak solution as a minimizer for F, we first need to know that F is bounded below.

To obtain a lower bound, we first note that by Cauchy-Schwarz,

$$F(u) = \frac{1}{2}(u,u)_1 + \langle f, u \rangle \geq \frac{1}{2}(u,u)_1 - \|f\|_2 \|u\|_2 . \tag{5}$$

Now recall the Poincaré inequality (see the lecture notes from two weeks ago, §2.3), which states that

$$\|u\|_2^2 \leq C \|\nabla u\|_2^2 = C(u,u)_1 \qquad \text{for all } u \in X = H_0^1(\Omega), \tag{6}$$

where C only depends on Ω. To apply this in (5), we use the inequality (Exercise 1 below)

$$ab \leq \varepsilon a^2 + \frac{1}{4\varepsilon} b^2 \qquad \text{for all } \varepsilon > 0 \text{ and } a, b \in \mathbb{R}. \tag{7}$$

Applying this to the last term in the right hand side of (5), and then using (6), we get

$$\begin{aligned} F(u) &\geq \frac{1}{2}(u,u)_1 - \frac{1}{4\varepsilon}\|f\|_2^2 - \varepsilon \|u\|_2^2 \\ &\geq \frac{1}{2}(u,u)_1 - \frac{1}{4\varepsilon}\|f\|_2^2 - C\varepsilon (u,u)_1 \\ &= \left(\frac{1}{2} - C\varepsilon\right)(u,u)_1 - \frac{1}{4\varepsilon}\|f\|_2^2 . \end{aligned}$$

Now choose $\varepsilon > 0$ so small that $1/2 - C\varepsilon = 1/4$, i.e., $\varepsilon = 1/4C$. Then we conclude that

$$F(u) \geq \frac{1}{4}(u,u)_1 - C\|f\|_2^2 . \tag{8}$$

In particular, $F(u) \geq -C\|f\|_2^2$, so F is bounded below. With this information in hand, we can define $I \in \mathbb{R}$ by

$$I = \inf_{u \in X = H_0^1(\Omega)} F(u). \tag{9}$$

We can then choose a sequence $\{u_j\} \subset X$ such that

$$F(u_j) \to I \qquad \text{as } j \to \infty. \tag{10}$$

We may also assume

$$F(u_j) \leq I + 1 \qquad \text{for all } j. \tag{11}$$

Let us now show that the sequence $\{u_j\}$ is bounded in $X = H_0^1(\Omega)$. Indeed,

$$\begin{aligned} \frac{1}{2}(u_j,u_j)_1 &= F(u_j) - \langle f, u_j \rangle && \text{by definition of } F \\ &\leq I + 1 + \|f\|_2 \|u_j\|_2 && \text{by (11) and Cauchy-Schwarz} \\ &\leq I + 1 + \frac{1}{4\varepsilon}\|f\|_2^2 + \varepsilon \|u_j\|_2^2 && \text{by (7)} \\ &\leq I + 1 + \frac{1}{4\varepsilon}\|f\|_2^2 + C\varepsilon (u_j,u_j)_1 && \text{by (6),} \end{aligned}$$

so choosing again $\varepsilon = 1/4C$, we conclude that

$$|u_j|_{1,2} \leq 4(I+1) + 4C\|f\|_2^2 \qquad \text{for all } j, \tag{12}$$

which proves boundedness in X of the minimizing sequence $\{u_j\}$.

In the following discussion, whenever we extract some subsequence of our minimizing sequence $\{u_j\}$, we shall always denote the subsequence by $\{u_j\}$. In other words, we simply replace the original sequence by the subsequence. We can do this because the properties (10) and (11) are obviously true for every subsequence.

Now, if it should happen that some subsequence of $\{u_j\}$ has a limit u in X, then by the continuity of F we would have $F(u) = \lim_{j\to\infty} F(u_j) = I$. So in this case, u would be a minimizer.

However, even though $\{u_j\}$ is a bounded sequence in X, we cannot in general deduce that it has a convergent subsequence, because X is an infinite-dimensional space (see the discussion in last week's notes). To fix this problem, we can use weak compactness. Recall the theorem from last week, stating that every bounded sequence in a Hilbert space has a weakly convergent subsequence. We may therefore assume (passing to a subsequence—see the above remark on notation) that our minimizing sequence $\{u_j\}$ converges weakly in X, to a limit u. It remains to prove that u is actually a minimizer, i.e., that $F(u) = I$. For this, we need some kind of "weak continuity" of F. In the following section we make this vague idea precise.

Exercise 1. Prove (7).

4. Weak lower semicontinuity of F

As in the previous section, assume

$$u_j \to u \qquad \text{weakly in } X = H_0^1(\Omega). \tag{13}$$

We then claim that

$$F(u) \le \liminf_{j\to\infty} F(u_j), \tag{14}$$

with $F(u) = (1/2)(u,u)_1 + \langle f, u \rangle$, as before. (If you are not familiar with the limit inferior, see the appendix to these notes.)

The property (14) is called *weak lower semicontinuity*. We can restate it as follows. Given any $\varepsilon > 0$, there exists $N = N(\varepsilon) \in \mathbb{N}$ such that

$$F(u) - \varepsilon \le F(u_j) \qquad \text{for all } j \ge N. \tag{15}$$

Before proving this property, let us assume it is true, and use it to prove that the weak limit u is a minimizer, i.e., that $F(u) = I := \inf_{v\in X} F(v)$. The inequality $F(u) \ge I$ holds by definition of I, so we need only prove $F(u) \le I$. It suffices to show that $F(u) \le I + 2\varepsilon$ for every $\varepsilon > 0$. Fix $\varepsilon > 0$, and let $N = N(\varepsilon) \in \mathbb{N}$ be as in (15). By (10) we can also find $N' \in \mathbb{N}$ such that $F(u_j) \le I + \varepsilon$ for all $j \ge N'$. From this and (15) we then conclude that

$$F(u) - \varepsilon \le F(u_j) \le I + \varepsilon \qquad \text{for all } j \ge \max(N, N'),$$

hence $F(u) \le I + 2\varepsilon$, as desired.

It remains to prove (14). We need the following fact:

Lemma 2. *In every Hilbert space H, the norm is weakly lower semicontinuous. I.e., if $x_j \to x$ weakly in H, then*

$$\|x\| \le \liminf_{j\to\infty} \|x_j\|.$$

Also,

$$\|x\|^2 \le \liminf_{j\to\infty} \left(\|x_j\|^2\right).$$

The proof of the first inequality is Exercise 7.1.3 in McOwen (assigned for this week). The second inequality follows from the first one, by the following general fact:

Exercise 2. Suppose $a \ge 0$ and $\{a_j\} \subset [0,\infty)$. Show that if

$$a \le \liminf_{j\to\infty} a_j,$$

then

$$a^2 \le \liminf_{j\to\infty} (a_j^2).$$

Applying Lemma 2 to the sequence $\{u_j\}$, and using (13), we conclude that

$$\frac{1}{2}(u,u)_1 \le \liminf_{j\to\infty} \frac{1}{2}(u_j,u_j)_1. \tag{16}$$

This takes care of the first term in $F(u)$, but we also need control the second term, $\langle f,u\rangle$. To handle this, we can use a theorem proved last week:

Theorem 1. *The inclusion $H_0^1(\Omega) \subset L^2(\Omega)$ is compact.*

In view of this, we may assume that $\{u_j\}$ converges (strongly) in $L^2(\Omega)$ to a limit u'. We claim that $u' = u$. To see this, note that $\{u_j\}$ converges weakly to u in $H_0^1(\Omega)$, hence it also converges weakly to u in $L^2(\Omega)$. On the other hand, $\{u_j\}$ converges strongly, hence also weakly, to u' in $L^2(\Omega)$. By uniqueness of the weak limit we conclude that $u' = u$.

Thus, we may assume that $\{u_j\}$ converges to u in two senses: (i) weakly in $X = H_0^1(\Omega)$, and (ii) strongly in $L^2(\Omega)$. By (ii) we have, of course,

$$\lim_{j\to\infty} \langle f,u_j\rangle = \langle f,u\rangle. \tag{17}$$

Finally, combining (16) and (17), and using also the general fact that

$$\liminf_{j\to\infty}(a_j+b_j) \ge \liminf_{j\to\infty} a_j + \liminf_{j\to\infty} b_j, \tag{18}$$

for sequences in $\mathbb{R}$, we get:

$$\begin{aligned}
\liminf_{j\to\infty} F(u_j) &= \liminf_{j\to\infty}\left(\frac{1}{2}(u_j,u_j)_1 + \langle f,u_j\rangle\right) && \text{by definition of } F\\
&\ge \liminf_{j\to\infty}\frac{1}{2}(u_j,u_j)_1 + \liminf_{j\to\infty}\langle f,u_j\rangle && \text{by (18)}\\
&\ge \frac{1}{2}(u,u)_1 + \langle f,u\rangle && \text{by (16) and (17)}\\
&= F(u), && \text{by definition of } F
\end{aligned}$$

which proves (14). Here we also used the fact (see the appendix) that if $\{a_j\} \subset \mathbb{R}$ is a convergent sequence, then $\liminf_{j\to\infty} a_j = \lim_{j\to\infty} a_j$.

Exercise 3. Prove (18).

5. Convexity and uniqueness

From (2) and (3) we see that

$$F(u+v) > F(u) + D_v F(u) \qquad \text{for all } u, v \in X,\ v \neq 0. \tag{19}$$

This property is called *strict convexity*. (To explain this terminology, think of the analogy with the case where $F:\mathbb{R}^2 \to \mathbb{R}$. Then $v \mapsto F(u)+D_v F(u)$ parametrizes the tangent plane to the graph of F at the point u, and (19) says that the graph of F lies strictly above the tangent plane, except at the point of tangency, of course.)

We now show that the strict convexity implies that F can have at most one critical point. Once we have proved this, it follows, of course, that the weak solution of (4) is unique.

So suppose there are two different critical points, u_1 and u_2. Thus, we assume $D_v F(u_1) = D_v F(u_2) = 0$ for all $v \in X$. Setting first $v = u_2 - u_1$, we see from (19) that

$$F(u_2) = F(u_1 + v) > F(u_1).$$

By symmetry, we also have $F(u_1) > F(u_2)$; we have arrived at a contradiction, hence there cannot exist two different critical points.

Finally, let us remark that (19) implies that any critical point is necessarily a strict global minimum. (This also implies uniqueness, of course.)

6. Appendix: The limit inferior and limit superior of a sequence

Suppose $\{a_j\} \subset \mathbb{R}$ is bounded below. The *limit inferior* of $\{a_j\}$ is defined by

$$\liminf_{j\to\infty} a_j = \lim_{N\to\infty}\left(\inf_{j\geq N} a_j\right).$$

Note that the limit on the right hand side does exist, since

$$b_N := \inf_{j\geq N} a_j$$

is monotone increasing as $N \to \infty$ (the inf is over a smaller and smaller set as N increases).

If $\{a_j\}$ is also bounded above, we can define the *limit superior* by

$$\limsup_{j\to\infty} a_j = \lim_{N\to\infty}\left(\sup_{j\geq N} a_j\right).$$

Again, this is well-defined.

Let us write $\alpha = \liminf_{j\to\infty} a_j$ and $\beta = \limsup_{j\to\infty} a_j$, assuming these exist.

For proving properties of these limits, the following facts are quite useful:

Since $\alpha = \lim_{N\to\infty} b_N$, we see that the following holds. Given $\varepsilon > 0$, there exists $N = N(\varepsilon) \in \mathbb{N}$ such that $\alpha - \varepsilon \leq b_N$, and hence

$$\alpha - \varepsilon \leq a_j \qquad \text{for all } j \geq N.$$

(We remark that $\liminf a_j$ can be characterized as the *largest* number α satisfying the above property. This provides an equivalent definition of the limit inferior.)

Similarly, if β exists, then given $\varepsilon > 0$, there exists $N' = N'(\varepsilon) \in \mathbb{N}$ such that

$$a_j \le \beta + \varepsilon \qquad \text{for all } j \ge N'.$$

(We remark that $\limsup a_j$ can be characterized as the *smallest* number β satisfying the above property. This provides an equivalent definition of the limit superior.)

The advantage of the limit inferior/superior is that these always exist (for a bounded sequence), whereas the ordinary limit does not always exist. Here are some simple relations between these various types of limits (here we assume $\{a_j\}$ is bounded, and we define α and β as above):

(i) $\alpha \le \beta$.

(ii) If $\{a_j\}$ converges to a number a, then $\alpha = \beta = a$.

(iii) Conversely, if $\alpha = \beta$, then $\{a_j\}$ converges, and $\lim_{j\to\infty} a_j = \alpha = \beta$.

Exercise 4. Prove the properties listed above.

Exercises for this week: (from McOwen) 7.1: 3, 6. Also: Exercises 1, 2, 3 and 4 from these notes.