LECTURE NOTES
TMA4305 PARTIAL DIFFERENTIAL EQUATIONS, SPRING 2007

SIGMUND SELBERG

Last week we proved existence of weak solutions of the Dirichlet problem for a class of elliptic operators (in particular, for the Laplace operator), by using the Riesz representation theorem. This method, however, is limited to linear PDEs. It is therefore of interest to study more robust methods, which can be applied also to nonlinear PDEs. One such method is to obtain the solution to the PDE as a minimizer of a functional (typically some kind of energy functional). We will encounter this method in chapter 7 of McOwen, but for this we shall require some further tools from functional analysis, namely some "compactness" results.

To motivate this, let us first recall that in a finite-dimensional vector space such as $\mathbb{R}^n$, every bounded sequence has a convergent subsequence. This property fails in infinite-dimensional spaces, but there are ways to fix this: For example, we can look for additional conditions on a sequence which guarantee that it *does* have a convergent subsequence, or we can settle for a weaker type of convergence than norm convergence.

Let us start with a discussion of weak convergence.

1. Weak convergence

Definition 1. A sequence $\{x_j\}$ in a normed space X is said to be *weakly convergent* if there exists $x \in X$ such that for every $F \in X^*$, we have $F(x_j) \to F(x)$ as $j \to \infty$. We then call x the *weak limit* of the sequence.

To distinguish weak convergence from the standard convergence in norm, i.e., $\|x_j - x\| \to 0$ as $j \to \infty$, the latter is sometimes called *strong convergence*.

Remark 1. Let us note that

(i) strong convergence $\implies$ weak convergence;

(ii) in a finite-dimensional space, weak convergence $\iff$ strong convergence;

(iii) the weak limit, if it exists, is unique;

(iv) if $\{x_j\}$ converges weakly, then $\{\|x_j\|\}$ is a bounded sequence.

We skip the proofs of these assertions; see any book on functional analysis.

Let us now specialize to a Hilbert space H (infinite-dimensional, of course), which is what we need for our applications. Recalling that H^* consists precisely of the functionals of the form $\langle \cdot, y \rangle$, where $y \in H$, we conclude that

$$x_j \to x \text{ weakly in } H \quad \iff \quad \langle x_j, y \rangle \to \langle x, y \rangle \text{ for all } y \in H.$$

Here is a classic example, showing that weak convergence does not imply strong convergence, in an infinite-dimensional space.

Example 1. Consider the space $l^2(\mathbb{N})$ of sequences $\{\alpha_n\} \subset \mathbb{R}$ such that $\sum_{n=1}^{\infty} \alpha_n^2 < \infty$. This is a Hilbert space when equipped with the inner product $\langle \{\alpha_n\}, \{\beta_n\} \rangle = \sum_{n=1}^{\infty} \alpha_n \beta_n$; the associated norm is $\|\{\alpha_n\}\| = \left(\sum_{n=1}^{\infty} \alpha_n^2\right)^{1/2}$. Let $e_n \in l^2(\mathbb{N})$ be the sequence consisting of all zeros except in the n-place, where the value is 1. Then $\{e_n\}_{n=1}^{\infty}$ is an orthonormal basis for $l^2(\mathbb{N})$. We claim that e_n converges weakly to zero. Indeed, if $y = \{\alpha_n\} \in l^2(\mathbb{N})$, then $\sum_{n=1}^{\infty} \alpha_n^2 < \infty$, so we must have $\alpha_n \to 0$ as $n \to \infty$. But

$$\langle e_n, y \rangle = \alpha_n,$$

so we conclude that $e_n \to 0$ weakly in $l^2(\mathbb{N})$. On the other hand, e_n certainly does not converge in norm; if it did, it would have to be Cauchy, but using the orthonormality we have $\|\alpha_n - \alpha_m\|^2 = 2$ for all $n \neq m$, so it is not Cauchy.

We shall need the following fundamental result:

Theorem 1. (Weak compactness in a Hilbert space.) *In a Hilbert space H, every bounded sequence has a weakly convergent subsequence.*

The proof is left as one of the exercises for this week (see the end of these notes).

Note that the above theorem applies to the Sobolev spaces $H_0^1(\Omega)$ and $H^1(\Omega)$, which are Hilbert spaces.

2. Compactness in the uniform norm: The Arzelà-Ascoli theorem

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with smooth boundary, and let us denote the uniform norm by

$$\|f\|_\infty = \sup_{x \in \Omega} |f(x)|.$$

We shall need the following special case of the classic Arzelà-Ascoli theorem:

Theorem 2. *Assume $\{f_j\} \subset C^1\left(\overline{\Omega}\right)$ is uniformly bounded, i.e., there exists $M > 0$ such that*

$$\|f_j\|_\infty + \|\nabla f_j\|_\infty \leq M \qquad \text{for all } j.$$

Then $\{f_j\}$ has a subsequence which converges uniformly on $\overline{\Omega}$.

Proof. To avoid some subtle topological issues, we shall assume that Ω is convex (this is enough for our applications of this theorem). Thus, if $x, y \in \Omega$, then the line segment between x and y is assumed to be completely contained in Ω, and therefore

$$|f_j(x) - f_j(y)| = \left| \int_0^1 \nabla f_j(x + t[x-y]) \cdot (x-y)\, dt \right| \leq M|x-y|, \tag{1}$$

for all $j \in \mathbb{N}$. By continuity, this inequality also holds for all x, y in the closure $\overline{\Omega}$.

Now pick a dense countable subset $\{x_j\}$ of Ω. Since $\{f_j(x_1)\}$ is bounded in $\mathbb{R}$, this sequence has a convergent subsequence, which we denote $f_j^1(x_1)$, converging to a limit which we denote $f(x_1)$. Next, we repeat the procedure on the sequence $f_j^1(x_2)$, extracting a subsequence $f_j^2(x_2)$ converging to a number $f(x_2)$. Continuing in this way, we get a double sequence of functions $\{f_j^k\}_{j,k=1}^\infty$, which we can organize in an infinite matrix

$$\begin{matrix} f_1^1 & f_2^1 & f_3^1 & \cdots \\ f_1^2 & f_2^2 & f_3^2 & \cdots \\ f_1^3 & f_2^3 & f_3^3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{matrix}$$

By construction, this has the properties that (i) the k-th row converges when evaluated on x_k, i.e., $\lim_{j\to\infty} f_j^k(x_k) = f(x_k)$, and (ii) each row is a subsequence of the previous row.

Now we use Cantor's diagonal trick: we consider the diagonal sequence f_j^j, which has the nice property that it is eventually a subsequence of every one of the rows in the above matrix. To be precise, $\{f_j^j\}_{j=k}^\infty$ is a subsequence of the k-th row. By the properties (i) and (ii) above, we can therefore conclude that

$$\lim_{j\to\infty} f_j^j(x_k) = f(x_k) \qquad \text{for all } k. \tag{2}$$

Next, we prove that $f_j^j(x)$ is Cauchy (and hence converges) for *all* $x \in \overline{\Omega}$. So fix $x \in \overline{\Omega}$, and let $\varepsilon > 0$. By density, there exists $m \in \mathbb{N}$ such that

$$|x - x_m| \le \frac{\varepsilon}{3M}. \tag{3}$$

Since $\{f_j^j(x_m)\}$ is Cauchy in $\mathbb{R}$, by (2), we can find $N(\varepsilon, x_m) \in \mathbb{N}$ such that

$$\left| f_j^j(x_m) - f_k^k(x_m) \right| \le \frac{\varepsilon}{3} \qquad \text{for all } j,k \ge N(\varepsilon, x_m). \tag{4}$$

Using (1), (3) and (4), we now find

$$\begin{aligned} \left| f_j^j(x) - f_k^k(x) \right| &\le \left| f_j^j(x) - f_j^j(x_m) \right| + \left| f_j^j(x_m) - f_k^k(x_m) \right| + \left| f_k^k(x_m) - f_k^k(x) \right| \\ &\le M|x - x_m| + \frac{\varepsilon}{3} + M|x - x_m| \\ &\le \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon, \end{aligned}$$

for all $j,k \ge N(\varepsilon, x_m)$.

So now we know that $f_j^j(x)$ converges for all $x \in \overline{\Omega}$, and it only remains to show that the convergence is *uniform*. For this, we use the fact that $\overline{\Omega}$ is compact. Therefore, given $\varepsilon > 0$, we can cover $\overline{\Omega}$ by finitely many balls of the type (3), centered at $x_1, \dots, x_m$, say. Then the argument above shows that

$$\left| f_j^j(x) - f_k^k(x) \right| \le \varepsilon$$

for all $x \in \overline{\Omega}$ and for all

$$j,k \ge \max\bigl(N(\varepsilon, x_1), \dots, N(\varepsilon, x_m)\bigr).$$

This proves the uniform convergence. □

The following general terminology is quite convenient: Suppose X and Y are normed spaces with $X \subset Y$, and that the inclusion map is bounded, i.e., $\|x\|_Y \le C\|x\|_X$ for all $x \in X$. If every bounded sequence in X has a subsequence which converges in Y, then we say that the *inclusion $X \subset Y$ is compact*. Note carefully that the subsequence need not converge in the X-norm, only in the Y-norm.

A quick way to state the above theorem is then: *The inclusion $C^1(\overline{\Omega}) \subset C(\overline{\Omega})$ is compact.*

3. Compactness of the inclusion $H_0^1(\Omega) \subset L^2(\Omega)$

Recall that we defined both $L^2(\Omega)$ and $H_0^1(\Omega)$ as completions of $C_0^\infty(\Omega)$, with respect to the norms, respectively,

$$\|f\|_2 = \left(\int_\Omega |f(x)|^2 \, dx\right)^{1/2},$$

$$\|f\|_{1,2} = \left(\int_\Omega |f(x)|^2 + |\nabla f(x)|^2 \, dx\right)^{1/2}.$$

Since $\|f\|_2 \le \|f\|_{1,2}$, if follows by our construction of the completion (see the notes from last week) that $H_0^1(\Omega) \subset L^2(\Omega)$, and that the inclusion map is bounded. Our aim now is to prove that this inclusion is in fact compact.

Rellich's Theorem. *Assume Ω is a bounded domain in $\mathbb{R}^n$. Then every bounded sequence in $H_0^1(\Omega)$ has a subsequence which converges in $L^2(\Omega)$.*

To prove this, we shall need to use a *mollifier*. Let $\rho : \mathbb{R}^n \to [0,\infty)$ be a function which is C^∞, supported in the unit ball in $\mathbb{R}^n$, and satisfies $\int_{\mathbb{R}^n} \rho(x)\, dx = 1$. For example, the function

$$\rho(x) = \begin{cases} Ce^{-\frac{1}{1-|x|^2}} & \text{for } |x| < 1, \\ 0 & \text{for } |x| \ge 1, \end{cases}$$

is C^∞, and by choosing C appropriately we can make $\int_{\mathbb{R}^n} \rho(x)\, dx = 1$.

For $h > 0$ we then define

$$\rho_h(x) = \frac{1}{h^n}\rho\left(\frac{x}{h}\right).$$

Note that ρ_h is supported in the ball $|x| \le h$, and that (by a change of variables)

$$\|\rho_h\|_1 = \int_{\mathbb{R}^n} \rho_h(x)\, dx = \int_{\mathbb{R}^n} \rho(x)\, dx = 1, \tag{5}$$

for all $h > 0$.

If $f : \mathbb{R}^n \to \mathbb{R}$, we define, for $h > 0$,

$$f^h(x) = \rho_h * f(x) = \int_{\mathbb{R}^n} f(x-y)\rho_h(y)\, dy = \int_{\mathbb{R}^n} f(y)\rho_h(x-y)\, dy,$$

provided this integral exists.

The family $\{\rho_h\}_{h>0}$ is called a *mollifier*, and f^h is called the *mollification* of f. The reason for this terminology is that convolution with ρ_h tends to smooth out the function f. Think of $f^h(x)$ as a kind of average of the values of f in a small neighborhood of radius h around x, with ρ_h as weight function. This average tends to be smoother than f (if f is not smooth already), but it should also get closer and closer to f as $h \to 0$, since ρ_h then converges to the Dirac distribution δ, which we know has the property that $\delta * f = f$. See Theorem 3 in Section 6.5 of McOwen for some precise statements along these lines. Here, however, we shall only need two very simple properties. The first one is:

Lemma 1. *If $f \in C_0^1(\mathbb{R}^n)$, then*

$$\left\| f^h - f \right\|_2 \le \|\nabla f\|_2 \, h.$$

Here $\|\cdot\|_2$ denotes the norm on $L^2(\mathbb{R}^n)$.

Proof. Using (5) we can write

$$\begin{aligned}
\left| f^h(x) - f(x) \right| &= \left| \int_{\mathbb{R}^n} \left[f(x-y) - f(x) \right] \rho_h(y)\, dy \right| \\
&= \left| \int_{|y|\le h} \left(\int_0^1 \nabla f(x-ty)\cdot y \, dt \right) \rho_h(y)\, dy \right| \\
&\le \int_{|y|\le h} \int_0^1 \left| \nabla f(x-ty) \right| \left| y \right| \rho_h(y)\, dt\, dy \\
&\le h \int_0^1 \int_{|y|\le h} \left| \nabla f(x-ty) \right| \rho_h(y)\, dy\, dt.
\end{aligned}$$

Therefore, by Minkowski's integral inequality,[1]

$$\begin{aligned}
\left\| f^h - f \right\|_2 &\le h \int_0^1 \int_{|y|\le h} \|\nabla f(\cdot - ty)\|_2 \, \rho_h(y)\, dy\, dt \\
&= h \left\| \nabla f \right\|_2 \int_0^1 \int_{|y|\le h} \rho_h(y)\, dy\, dt \\
&= h \left\| \nabla f \right\|_2,
\end{aligned}$$

where we again used (5), and also the fact that $\left\| \nabla f(\cdot - z) \right\|_2 = \left\| \nabla f \right\|_2$ for all $z \in \mathbb{R}^n$, by a change of variables. □

The second property of mollifiers that we need is:

Lemma 2. *There exists $C = C(\rho)$ such that for all $f \in L^2(\mathbb{R}^n)$,*

$$\left\| f^h \right\|_\infty \le C h^{-n/2} \|f\|_2 .$$

Proof. Using the Cauchy-Schwarz inequality, we get for any $x \in \mathbb{R}^n$,

$$\left| f^h(x) \right| \le \int_{\mathbb{R}^n} \left| f(x-y) \right| \rho_h(y)\, dy \le \|f\|_2 \, \|\rho_h\|_2 ,$$

and a simple change of variables shows that $\left\| \rho_h \right\|_2 = C h^{-n/2}$, where $C = \left\| \rho \right\|_2$. □

We now have the tools we need to prove the main theorem of this section.

[1] We use the following special case of this inequality:

$$\left\| \int F(\cdot, y)\, dy \right\|_2 \le \int \|F(\cdot, y)\|_2 \, dy.$$

Proof of Rellich's Theorem. We split the proof into two steps.

Step 1. Let $\{f_j\} \subset C_0^\infty(\Omega)$, and assume there exists $M > 0$ such that

$$\|f_j\|_{1,2} \le M \qquad \text{for all } j. \tag{6}$$

It will be convenient to extend f_j by zero outside Ω. For $h > 0$, define $f_j^h = \rho_h * f_j$. Observe that if we assume also $h \le 1$, then the f_j^h are all supported in a fixed compact subset of $\mathbb{R}^n$. Thus, in the remainder of the proof, all L^2-norms and all uniform norms can be taken over the whole space $\mathbb{R}^n$.

Applying Lemma 2 to f_j^h and also to $\nabla(f_j^h) = \rho_h * (\nabla f_j)$, and making use of the bound (6), we obtain

$$\left\|f_j^h\right\|_{1,\infty} \le Ch^{-n/2} \qquad \text{for all } j \text{ and } h > 0.$$

for some constant C independent of j and h. Therefore, for any fixed $h > 0$, it follows from the Arzelà-Ascoli theorem (Theorem 2) that $\{f_j^h\}$ has a subsequence which converges uniformly on $\mathbb{R}^n$ (recall that our functions are all supported in a fixed compact set).

Applying the above argument with $h = 1/m$ for $m = 1, 2, \dots$, taking successive subsequences, and applying the diagonal argument, we can then find a *single* subsequence f_{j_k} of f_j which works for all $h = 1/m$. To be precise, we can find a sequence $j_1 < j_2 < \cdots < j_k < \dots$ in $\mathbb{N}$ with the property that

$$\{f_{j_k}^h\}_{k=1}^\infty \qquad \text{is uniformly Cauchy, for every } h = \frac{1}{m},\ m \in \mathbb{N}. \tag{7}$$

We now claim that f_{j_k} converges in L^2. By completeness of L^2, it suffices to prove that f_{j_k} is Cauchy in the L^2-norm. To this end, we write

$$\|f_{j_k} - f_{j_l}\|_2 \le \left\|f_{j_k} - f_{j_k}^h\right\|_2 + \left\|f_{j_k}^h - f_{j_l}^h\right\|_2 + \left\|f_{j_l}^h - f_{j_l}\right\|_2 \tag{8}$$

Let $\varepsilon > 0$. From Lemma 1 and the bound (6), we obtain

$$\left\|f_{j_k}^h - f_{j_k}\right\|_2 \le Mh \qquad \text{for all } k.$$

Choose $h = 1/m$ with $m \in \mathbb{N}$ so large that $Mh \le \varepsilon/3$. This guarantees that the first and third terms on the right hand side of (8) are no larger than $\varepsilon/3$. As for the middle term, we can estimate it by

$$\left\|f_{j_k}^h - f_{j_l}^h\right\|_2 \le C\left\|f_{j_k}^h - f_{j_l}^h\right\|_\infty,$$

where C is the square root of the volume of the (fixed) support of our functions. Now use (7) (with the h chosen above) to conclude that there exists $N \in \mathbb{N}$ such that

$$\left\|f_{j_k}^h - f_{j_l}^h\right\|_\infty \le \frac{\varepsilon}{3C} \qquad \text{for all } k, l \ge N.$$

Thus, the middle term of the right hand side of (8) is no larger than $\varepsilon/3$ if $k, l \ge N$, and we conclude that

$$\|f_{j_k} - f_{j_l}\|_2 \le \varepsilon \qquad \text{for all } k, l \ge N.$$

This concludes Step 1.

Step 2. Assume now that $\{f_j\} \subset H_0^1(\Omega)$ is a bounded sequence, so $\|f_j\|_{1,2} \le M$ for some M independent of j. By density, we can choose $g_j \in C_0^\infty(\Omega)$ such that $\|f_j - g_j\|_{1,2} \le 1/j$. Applying Step 1 to the sequence $\{g_j\}$, for which we have the bound $\|g_j\|_{1,2} \le M + 1$, we then obtain a subsequence g_{j_k} which converges in L^2 to a function f. But then f_{j_k} clearly also converges to f in L^2, and the proof of Rellich's theorem is complete. □

4. Compactness of the inclusion $H^1(\Omega) \subset L^2(\Omega)$

We state the following without proof (see McOwen Section 6.5):

Theorem 3. *Suppose $\Omega \subset \mathbb{R}^n$ is a bounded domain with smooth boundary. Then every bounded sequence in $H^1(\Omega)$ has a subsequence which converges in $L^2(\Omega)$.*

Exercises for this week: (from McOwen)

6.1: 5.

6.2: 4.

6.3: 3, 7.