Lecture Notes for MATH 592A

Vladislav Panferov

February 25, 2008

1. Introduction

As a motivation for developing the theory let's consider the following boundary-value problem,

$$\begin{aligned} -u'' + u &= f(x) \quad x \in \Omega = (0,1) \\ u(0) = 0, &\quad u(1) = 0, \end{aligned}$$

where f is a given (smooth) function. We know that the solution is unique, satisfies the stability estimate following from the maximum principle, and it can be expressed explicitly through Green's function. However, there is another way "to say something about the solution", quite independent of what we've done before.

Let's multuply the differential equation by u and integrate by parts. We get

$$-\Big[u'\,u\Big]_{x=0}^{x=1} + \int_0^1 (u'^{\,2} + u^2)\,dx = \int_0^1 f\,u\,dx.$$

The boundary terms vanish because of the boundary conditions. We introduce the following notations

$$\|u\|^2 = \int_0^1 u^2\,dx \quad \text{(the norm)}, \qquad \text{and} \quad (f,u) = \int_0^1 f\,u\,dx \quad \text{(the inner product)}.$$

Then the integral identity above can be written in the short form as

$$\|u'\|^2 + \|u\|^2 = (f,u).$$

Now we notice the following inequality

$$|(f,u)| \leqslant \|f\|\,\|u\|. \qquad \text{(Cauchy-Schwarz)}$$

This may be familiar from linear algebra. For a quick proof notice that

$$\|f + \lambda u\|^2 = \|f\|^2 + 2\lambda(f,u) + \lambda^2\|u\|^2 \geq 0$$

for any λ. This expression is a quadratic function in λ which has a minimum for $\lambda = -(f,u)/\|u\|^2$. Using this value of λ and rearranging the terms we get

$$(f,u)^2 \leqslant \|f\|^2 \, \|u\|^2.$$

Taking the square roots we obtain the Cauchy-Schwarz.

We use Cauchy-Schwarz to obtain

$$\|u'\|^2 + \|u\|^2 = (f,u) \leqslant \|f\| \|u\|.$$

Feeling generous we can increase the right-hand side to get

$$\|u'\|^2 + \|u\|^2 \leqslant \|f\| \, (\|u'\|^2 + \|u\|^2)^{1/2},$$

so

$$\|u'\|^2 + \|u\|^2 \leqslant \|f\|^2. \tag{1}$$

This is an example of an apriori estimate: by some formal manipulations with the equation we obtained quantitative information about the solution in terms of the "data" (the function f).

Estimate (1) is also called an energy estimate as the expression $\|u'\|^2 + \|u\|^2$ in some physical context can be viewed as the energy.

Using the energy estimate we can give another proof of uniqueness.

Theorem. (Uniqueness)

Assume that u_1, $u_2 \in C^2[0,1]$ are two solutions of the boundary-value problem. Then $u_1(x) = u_2(x)$, $x \in [0,1]$.

Proof. Take $w = u_1 - u_2$. By the apriori estimate $\|w'\|^2 + \|w\|^2 \leqslant 0$, so $w(x) = 0$, $x \in [0,1]$ since w is smooth. □

The formal calculation that led to estimate (1) can be extended (by adding a little sophistication) to the general case of a differential operator

$$\mathcal{A}\, u = -(au')' + bu' + cu$$

if a, b, c are smooth functions, $a(x) \geq a_0 > 0$ and $c(x) - b'(x)/2 \geq 0$.

In fact, this approach has many other useful consequences, but before we get there we want to learn some facts about functions for which all terms in the apriori estimate (1) are finite.

2. The function space $L^2(0,1)$

Definition. We say that a function $f : (0,1) \to \mathbb{R}$ belongs to $L^2(0,1)$, or that it is square integrable, if the integral

$$\|f\|^2 = \int_0^1 f(x)^2\,dx$$

is finite.

More generally, if $\Omega \subseteq \mathbb{R}^n$ we may define

$$\|f\|^2 = \int_\Omega f(x)^2\,dx$$

(the n-dimensional integral) and use the notation $L^2(\Omega)$ for the class of all functions for which $\|f\|$ is finite.

Examples: If $\alpha > -\frac{1}{2}$ then $f(x) = |x|^\alpha$, or more generally $f(x) = |x-a|^\alpha$ are in $L^2(0,1)$. The limiting case $f(x) = \frac{1}{\sqrt{|x|}}$ is not in $L^2(0,1)$. However, if $x = (x_1, x_2) \in \mathbb{R}^2$ and $|x|$ denotes the Euclidean length of the vector, then $f(x) = \frac{1}{\sqrt{|x|}}$ is in $L^2((-1,1)^2)$

Exercise: For which α does the function$f(x) = |x|^\alpha$, where $\mathbb{R}^n \ni x = (x_1, \ldots, x_n)$, satisfy $f \in L^2((-1,1)^n)$?

More examples: If $f(x) = 0$, $x < \frac{1}{2}$, $f(x) = 1$, $x \geq \frac{1}{2}$ then $f \in L^2(0,1)$. Notice that $f(x)$ is discontinuous.
If $\alpha = -\frac{1}{3}$ (say) then $f(x) = \sum_{i=1}^{n} c_i\,|x - a_i|^\alpha$, where $a_i \in (0,1)$, and c_i any, satisfies $f \in L^2(0,1)$. Again, f is discontinuous. For a more extreme example take $f(x) = \sum_{i=1}^{\infty} 2^{-i}\,|x - a_i|^\alpha$, where $\{a_i\}$ is a countable set in $(0,1)$, for example, the set of all rational numbers between 0 and 1. Then again, $f \in L^2(0,1)$.
Bottom line, $L^2(\Omega)$ is a good place for discontinuous functions.

In the above we assume that we know how to define integrals. In general this is done using Lebesgue's theory of integration, of which we will have to know very little (reading, say, the article on `wikipedia.org` would suffice). The most frequently used concept from there would be the one of a set of measure zero which comes into play as follows.

Definition. We say that two functions f_1 and $f_2 \in L^2(\Omega)$ are identical, or equal in the sense of $L^2(\Omega)$, if $\|f_1 - f_2\| = 0$.

Now, we saw that functions in L^2 do not have to be continuous. From the condition

$$\|f_1 - f_2\|^2 = \int_0^1 (f_1(x) - f_2(x))^2 \, dx = 0$$

we cannot conclude that $f_1(x) = f_2(x)$ for all $x \in (0,1)$. For example, if

$$f_1(x) = \begin{cases} 0, & x < \frac{1}{2} \\ 1, & x \geq \frac{1}{2} \end{cases} \quad \text{and} \quad f_2(x) = \begin{cases} 0, & x \leqslant \frac{1}{2} \\ 1, & x > \frac{1}{2} \end{cases}$$

then $f_1 = f_2$ in $L^2(0,1)$. More generally, the value of the integral is not changed if a function is modified for x on a set of measure zero. Examples of such sets in $\mathbb{R}$ are isolated points, finite and countable sets, and even some uncountable ones (a famous example is the Cantor set). In $\mathbb{R}^2$ we could add lines, or more general curves to that list, and in $\mathbb{R}^3$ (smooth) two-dimensional surfaces have measure zero.

Example: If $f(x) = 0$, $x \in (0,1)$ and $g(x)$ is the Dirichlet function

$$g(x) = \begin{cases} 1, & x \in \mathbb{Q} \\ 0, & \text{otherwise} \end{cases}$$

then $f = g$ in $L^2(0,1)$. Notice that g is not Riemann-integrable on $(0,1)$ (it is Lebesgue-integrable) but we will still say $\int_0^1 g(x)\, dx = 0$.

If a relation (such as equality) holds for all x except on a set of measure zero we say it holds almost everywhere or for almost all x (sometimes abbreviated as *a. e.* or *a. a.*). Another name for the same thing (borrowed from Probability) is "almost surely".

It is an easy observation that $L^2(\Omega)$ is a vector space. Indeed, if $f \in L^2$ and $\lambda \in \mathbb{R}$ then $\lambda f \in L^2$. Also, if $f \in L^2$ and $g \in L^2$ then

$$\int (f+g)^2 \, dx = \int (f^2 + 2fg + g^2) \, dx \leqslant \left(\sqrt{\int f^2 \, dx} + \sqrt{\int g^2 \, dx} \right)^2,$$

so $f + g \in L^2$ (we used the Cauchy-Schwarz).

We define the inner product (also called scalar product) on $L^2(0,1)$ as

$$(f,g) = \int_0^1 f(x)\, g(x) \, dx.$$

More generally, an inner product on a real vector space is defined by the properties

$$(f + \lambda g, h) = (f, h) + \lambda(g, h), \quad (f, g + \lambda h) = (f, h) + \lambda(f, h) \qquad \text{(bilinear)}$$

$$(f, g) = (g, f) \qquad \text{(symmetric)}$$

$$(f, f) \geq 0, \quad \text{and} \quad (f, f) = 0 \;\Leftrightarrow\; f = 0. \qquad \text{(positive definite)}$$

(For short, an inner product is a bilinear form that is symmetric and positive definite).

We also see that the norm $\|f\|$ satisfies $\|f\| = \sqrt{(f, f)}$ and

$$\|f\| \geq 0, \quad \text{and} \quad \|f\| = 0 \;\Leftrightarrow\; f = 0, \qquad \text{(positive definite)}$$

$$\|\lambda f\| = |\lambda| \, \|f\|, \qquad \text{(positive homogeneous)}$$

and

$$\|f + g\| \leqslant \|f\| + \|g\| \qquad \text{(triangle inequality)}$$

These must be familiar from linear algebra. The proof of the trinagle inequality follows from the Cauchy-Schwarz as above.

Definition. We say that a sequence $f_n \in L^2(\Omega)$ <u>converges</u> (as $n \to \infty$) to $f \in L^2(\Omega)$, denoted $f_n \to f$ in L^2, if $\|f_n - f\| \to 0$.

Other terms used frequently are "convergence in the mean square" or "convergence in the L^2-norm", or "convergence in L^2".

<u>Examples.</u> Set

$$f_n(x) = \begin{cases} 0, & x < \frac{1}{n} \\ \frac{1+nx}{2}, & -\frac{1}{n} \leqslant x \leqslant \frac{1}{n} \\ 1, & x > \frac{1}{n} \end{cases} \qquad \text{and} \quad f(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

Then $f_n \to f$ in $L^2(-1, 1)$. Indeed,

$$\int_{-1}^{1} (f_n(x) - f(x))^2 \, dx = \int_{-\frac{1}{n}}^{0} \frac{(1 + nx)^2}{4} \, dx + \int_{0}^{\frac{1}{n}} \frac{(1 - nx)^2}{4} \, dx = \frac{1}{6n} \to 0.$$

In this case $f(x)$ is discontinuous and f_n do not converge to f uniformly.

Geometrically, it is pehaps easier to see what happens when the square is replaced by the absolute value,

$$\int_0^1 |f_n - f|\, dx \to 0$$

(in this case we say $f_n \to f$ in L^1; this is a different, but related type of convergence). The above simply means that the area between the graphs of f_n and f goes to zero. You can use this as a somewhat crude analogy for the mean square convergence.

The space $L^2(0,1)$ is complete with respect to the mean square convergence. This means that any Cauchy sequence,

$$\|f_n - f_m\| \to 0, \quad n,\, m \to \infty$$

has a limit in L^2, i. e. $\exists\, f$ such that $f_n \to f$.

Definition: A complete vector space with an inner product is called a Hilbert space.

The space $L^2(0,1)$ is therefore an example of a Hilbert space.

As in any space with an inner product orthogonality of two functions $f, g \in L^2$ is defined by the relation $(f,g) = 0$. The notation $f \perp g$ will be used for orthogonal functions. A family of functions $f_1,\ f_2, \ldots$ is orthogonal if for any $i \neq j$ we have $(f_i, f_j) = 0$. You should check that if two or more functions are orthogonal then they are linearly independent.

Example: The infinite system of functions 1, $\cos 2\pi kx$, $\sin 2\pi kx$, $k = 1, 2, \ldots$, with $x \in (0,1)$ is orthogonal. This shows that $L^2(0,1)$ has infinitely many linearly independent functions, so it is infinite-dimensional (there is no finite basis).

Historically, the spaces L^2 appeared out of study of Fourier series. An important theorem states that a Fourier series

$$\sum_{k=-\infty}^{\infty} c_k\, e^{2\pi ikx} \;\text{ converges in } L^2(0,1) \quad \text{iff} \quad \sum_{k=-\infty}^{\infty} |c_k|^2 \quad \text{converges.}$$

3. Subspaces of a Hilbert space. Orthogonal projection

In the following V will denote a vector space.

Definition: A subset $V_0 \subseteq V$ is called a (linear) subspace if it itself is a vector space.

Examples: 1) Straight lines and planes passing through the origin are subspaces of the three-dimensional space $\mathbb{R}^3$.

2) Pick any N functions in $L^2(0,1)$, $f_1, \dots, f_N$. The set of all linear combinations $c_1 f_1 + \cdots + c_N f_N$ is a subspace of $L^2(0,1)$, called the linear span of $\{f_1, \dots, f_N\}$.

If V is a Hilbert space then $V_0 \subseteq V$ is called closed if the limit of each convergent sequence remains in V_0.

More generally, if V_0 is any subspace then the set $\overline{V_0}$, called the closure of V_0, is obtained by adding to V_0 all its limiting points. The closure of V_0 is itself a linear subspace. (Verify this!) If the closure of V_0 is all of V then we say that V_0 is dense in V.

Examples: 1) All subspaces in the previous example are closed. 2) Take the infinite system of functions 1, $\cos 2\pi kx$, $\sin 2\pi kx$, $k = 1, 2, \dots$ Taking all possible linear combinations of functions from this system we obtain a subspace $V_0 \subseteq L^2(0,1)$, which we may call the linear span of this infinite system. From the theory of Fourier series, any function in $f \in L^2(0,1)$ can be represented by its Fourier series which converges to f in $L^2(0,1)$. Thus, V_0 is dense in $L^2(0,1)$.

Lemma (Orthogonal projection)

Let V be Hilbert space and let V_0 be a closed subspace.

(a) For any $v \in V$ there exists a unique $v_0 \in V$ such that $\|v - v_0\| = \inf_{w \in V_0} \|v - w\|$.

(b) Any $v \in V$ is represented uniquely as $v = v_0 + v_1$, where $v_0 \in V_0$, $v_1 \perp V_0$.

Proof. (a) Take any $v \in V$. If $v \in V_0$ then we can take $v_0 = v$ and we are done. The interesting case is when $v \notin V_0$. We set $d = \inf_{w \in V_0} \|v - w\|$. By definition of infimium we can pick a sequence w_n such that $\|v - w_n\| \to d$. Using the parallelogram identity,

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2),$$

we obtain, for any n, m,

$$\|v - w_n\|^2 + \|v - w_m\|^2 = 2\big(\|v - \tfrac{w_n + w_m}{2}\|^2 + \|\tfrac{w_n - w_m}{2}\|^2\big).$$

(Draw a figure!) Since $\|v - \frac{w_n + w_m}{2}\|^2 \geq d^2$ and $\|v - w_n\|^2 \to d^2$, $\|v - w_m\|^2 \to d^2$, as $n, m \to \infty$, we obtain

$$\|w_n - w_m\|^2 \to 0, \quad \text{as} \quad n, m \to \infty.$$

Thus, the sequence w_n is Cauchy, and since V is complete, there exists $v_0 \in V$ such that $w_n \to v_0$. Since V_0 is closed, $v_0 \in V_0$. Also,

$$\|v - v_0\| \leqslant \|v - w_n\| + \|w_n - v_0\|.$$

Passing to the limit $n \to \infty$ we obtain

$$\|v - v_0\| \leqslant d,$$

and since also $\|v - v_0\| \geq d$ we obtain $\|v - v_0\| = d$. The uniqueness will follow from part (b).

(b) For the uniqueness we notice that if $v = v_0 + v_1 = v_0' + v_1'$ and $v_0, v_0' \in V_0$ and $v_1, v_1' \perp V_0$ then

$$0 = (v_0 - v_0') + (v_1 - v_1').$$

Multiplying the above equality scalarly by $v_0 - v_0'$ and $v_1 - v_1'$ we obtain, respectively,

$$0 = \|v_0 - v_0'\|^2 \quad \text{and} \quad 0 = \|v_1 - v_1'\|^2.$$

Thus, $v_0 = v_0'$ and $v_1 = v_1'$.

We found v_0 in part (a). It remains to prove then that $v_1 = v - v_0$ is orthogonal to V_0. For this take any $0 \neq y \in V_0$ and assume that $(v_1, y) \neq 0$. We will show that, informally, in such case it is possible to "decrease" the distance between v and V_0 by taking the vector $v_0 + \lambda y$, for λ suitably chosen. More precisely, we compute

$$\|v - v_0 - \lambda y\|^2 = \|v_1 - \lambda y\|^2 = \|v_1\|^2 - 2\lambda(v_1, y) + \lambda^2\|y\|^2.$$

Minimizing the right-hand side in λ we see that if we take $\lambda = \frac{(v_1,y)}{\|y\|^2}$ then

$$\|v - v_0 - \lambda y\|^2 = \|v_1\|^2 - \frac{(v_1, y)^2}{\|y\|^2} < \|v_1\|^2.$$

However, since $v_0 + \lambda y \in V_0$, it follows from part (a) that $\|v - v_0 - \lambda y\|^2 \geq \|v_1\|^2$, a contradiction. $\square$

4. Linear functionals on a Hilbert space

Definition. Let V be a vector space with a norm. A linear mapping $\ell : V \to \mathbb{R}$ is called a linear functional. A linear functional ℓ is bounded if

$$|\ell(v)| \leqslant C\|v\|, \quad \text{for all} \quad v \in V.$$

The smallest possible constant C is denoted by

$$\|\ell\| = \sup_{v \in V} \frac{|\ell(v)|}{\|v\|}$$

and is called the norm of the linear functional.
A bounded linear functional is continuous as a mapping from V to $\mathbb{R}$, simply because

$$|\ell(v_1) - \ell(v_2)| = |\ell(v_1 - v_2)| \leqslant C\|v_1 - v_2\|.$$

Conversely, if a linear functional on V is continuous then it is bounded. (Exercise: prove this.)

Examples. 1) For $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ define

$$\ell(x) = b_1 x_1 + b_2 x_2 + \ldots b_n x_n,$$

where $b_1, \ldots, b_n \in \mathbb{R}$ are constants. Then ℓ is a linear functional on $\mathbb{R}^n$.
2) Take $V = L^2(0,1)$. The integral $I(f) = \int_0^1 f(x)\,dx$ is a linear functional. More generally, $(w, f) = \int_0^1 w(x)\,f(x)\,dx$ where $w \in L^2(0,1)$ is a fixed function is a linear functional.

In the above examples all linear functionals were given by an inner product with a certain fixed element of V. In fact, as the next theorem shows, on a Hilbert space there are no other examples.

Theorem (F. Riesz)
Let V be a Hilbert space and let $\ell : V \to \mathbb{R}$ be a bounded linear functional. Then there exists a unique $u \in V$ such that

$$\ell(v) = (u, v) \quad \text{for all} \quad v \in V.$$

Moreover, $\|u\| = \|\ell\|$.

Proof. Take $V_0 = \ker \ell := \{v \in V : \ell(v) = 0\}$. Then V_0 is a closed subspace of V. (Verify this!) If $V_0 = V$ then $\ell = 0$ and we can take $u = 0$ and be done with the proof. The more interesting case is when $\ell \neq 0$. Then take $\overline{v}$ such that $\ell(\overline{v}) \neq 0$. By the Projection Lemma, $\overline{v} = \overline{v}_0 + \overline{v}_1$, $\overline{v}_0 \in V_0$, $\overline{v}_1 \perp V_0$.
Conjecture: $\ell(v) = k(\overline{v}_1, v)$, where k is a constant, i. e. $u = k\overline{v}_1$. (Motivation: they have the same kernel).

We can determine k by plugging in $v = \overline{v}_1$:

$$\ell(\overline{v}_1) = k(\overline{v}_1, \overline{v}_1) = k\|\overline{v}_1\|^2,$$

so $k = \ell(\overline{v}_1)/\|\overline{v}_1\|^2$. (Notice that $\|\overline{v}_1\| \neq 0$. Why?) It remains to show that $\ell(v) = k(\overline{v}_1, v)$ for all $v \in V$.

Take $v \in V$. We know that $v = v_0 + v_1$, where $v_0 \in V_0$ and $v_1 \perp V_0$. Believing that v_1 is parallel to $\overline{v}_1$ we write

$$v = \left(v - \tfrac{\ell(v)}{\ell(\overline{v}_1)}\,\overline{v}_1\right) + \tfrac{\ell(v)}{\ell(\overline{v}_1)}\,\overline{v}_1,$$

then the expression in the parentheses is in V_0 (check it!), so by the uniqueness in the Projection Lemma it must be v_0. But v_0 is orthogonal to $\overline{v}_1$, so

$$\left(v - \tfrac{\ell(v)}{\ell(\overline{v}_1)}\,\overline{v}_1, \overline{v}_1\right) = 0,$$

which rewrites itself as $\ell(v) = \frac{\ell(\overline{v}_1)}{\|\overline{v}_1\|^2}(\overline{v}_1, v) = k(\overline{v}_1, v)$. This completes the proof. $\square$

5. Approximation by smooth functions

Notation: Ω will be used to denote an open set in $\mathbb{R}^n$, with smooth boundary Γ.

$C^k(\Omega)$, where $k = 1, 2, \ldots$ will denote the class of functions that have continuous (partial) derivatives up to the order k in Ω.

By "smooth" we will generally mean C^k, with k sufficiently high (sometimes $k = 1$ will be enough).

The notation C^∞ will be used for the intersection of all C^k, $k = 1, 2 \ldots$ and will therefore mean that a function has continuous derivatives of any order in Ω.

We write $f \in C^k(\overline{\Omega})$ to indicate that $f \in C^k(\Omega)$ and all its partial derivatives have well-defined (finite) limits as $x \longrightarrow x_0 \in \Gamma$.

Example. The function

$$f_a(x) = \begin{cases} 0 & x < -a \\ \frac{(x+a)^2}{2a^2} & x \in [-a, 0) \\ 1 - \frac{(x-a)^2}{2a^2} & x \in [0, a) \\ 1 & x \geq a \end{cases}$$

is in $C^1(\mathbb{R})$ but not in $C^2(\mathbb{R})$.

The notation $C_0^1(\Omega)$ is used to denote the set of $C^1(\Omega)$ functions with compact support in Ω. Recall that the support (denoted $\operatorname{supp} f$) is the closure of the set of x

for which $f(x) \neq 0$. For instance, for a function to have a compact support in $(0,1)$ is has to vanish identically outside of a certain interval $[\varepsilon, 1-\varepsilon] \subseteq (0,1)$.

Exercise: give an example of a function in $C_0^1(0,1)$.

The following theorem states that smooth functions (and even smooth functions with compact support) are dense in $L^2(\Omega)$.

Theorem. (Approximation by smooth functions)

Let $f \in L^2(\Omega)$. Then there exists a sequence $\varphi_n \in C_0^\infty(\Omega)$ such that $\varphi_n \to f$ in $L^2(\Omega)$.

We will not give a proof, but we will use this theorem often. For a plausible explanation for why this result may be true we may say that while functions in L^2 can be extremely discontinuous, one can always find a sequence of smooth (but possibly extremely wiggly) functions so that the "area" between the graphs of f and φ_n goes to zero.

Example. The functions $f_a(x)$ from the previous example approximate the discontinuous function $f(x) = 1$, if $x \in (0,1)$, and $f(x) = 0$ if $x \in (-1,0)$ as $a \to 0$.

6. Sobolev spaces

Definition. The Sobolev space $H^1(0,1) = \{f \in L^2(0,1) : f' \in L^2(0,1)\}$.

This definition assumes that we are somehow able to define the derivative for $f \in L^2$. The classical pointwise definition wouldn't work so easily, since for instance the Dirichlet function $f(x) = 1$, $x \in (0,1) \cap \mathbb{Q}$, and $f(x) = 0$, $x \in (0,1) \setminus \mathbb{Q}$ is nowhere differentiable, while $\tilde{f}(x) \equiv 0$ is. However, they represent the same element in $L^2(0,1)$. We will therefore introduce a new concept of derivative.

Definition. Let $f \in L^2(0,1)$. We say that f is weakly differentiable, and $g \in L^2(0,1)$ is its weak derivative if

$$\int_0^1 g(x)\,\varphi(x)\,dx = -\int_0^1 f(x)\,\varphi'(x)\,dx, \quad \text{for all} \quad \varphi \in C_0^1(0,1).$$

In this case we write $g = f'$ (in the weak sense).

The motivation for the above definition is the following. If $f \in C^1[0,1]$ (or with piecewise continuous derivative) and $\varphi \in C^1[0,1]$ then we are allowed to integrate by parts to obtain

$$\int_0^1 f'(x)\,\varphi(x)\,dx = \Big[f(x)\varphi(x)\Big]_{x=0}^{x=1} - \int_0^1 f(x)\,\varphi'(x)\,dx.$$

If $\varphi \in C_0^1(0,1)$ then the boundary terms vanish, so f' satisfies the integral identity that appears in the defintion for all $\varphi \in C_0^1(0,1)$. Conversely, if $f \in C^1[0,1]$ has a weak derivative in the sense of the above definition then the weak derivative must coincide with f'. (Verify this!)

We next check the consistency of the definition of the weak derivative.

Lemma. (Uniqueness of weak derivative)
If $g_1, g_2 \in L^2(0,1)$ are both weak derivatives of $f \in L^2(0,1)$ then $g_1(x) = g_2(x)$, except for x in a set of measure zero.

Proof. We have $(g_1, \varphi) = -(f, \varphi')$, $(g_2, \varphi) = -(f, \varphi')$, for all $\varphi \in C_0^1$. Then $(g_1 - g_2, \varphi) = 0$. Functions in C_0^1 are dense in $L^2(0,1)$. Therefore $(g_1 - g_2, \varphi) = 0$ holds for all $\varphi \in L^2(0,1)$, in particular,

$$(g_1 - g_2, g_1 - g_2) = \|g_1 - g_2\|^2 = 0.$$

Therefore, $g_1(x) = g_2(x)$ except for x in a set of measure zero. □

We can interpret this Lemma by saying that a Sobolev function, as well as its derivative, may be modified arbitrarily on a set of measure zero, without changing the "identity" of the function. ("The values on a set of measure zero do not matter".) In fact, to determine a Sobolev function uniquely it would be sufficient to specify its values eveywhere except a "thin set" of measure zero.

Examples. 1) Take $f(x) = 0$, $x \in (-1,0)$, $f(x) = x$, $x \in (0,1)$. It is defined alsmost eveywhere on $(-1,1)$. In is easy to see that $f \in L^2(-1,1)$:

$$\int_{-1}^{1} f(x)^2\,dx = \int_0^1 x^2\,dx = \frac{1}{3} < +\infty.$$

Guess: $f'(x) = 0$, $x \in (-1,0)$, and $f'(x) = 1$, $x \in (0,1)$ (defined everywhere on $(-1,1)$ except $x = 0$). To verify take $\varphi \in C_0^1(-1,1)$, then

$$\int_{-1}^{1} f'(x)\,\varphi(x)\,dx = \int_0^1 \varphi(x)\,dx = -\int_0^1 x\,\varphi'(x)\,dx = -\int_{-1}^{1} f(x)\,\varphi'(x)\,dx$$

holds for all test functions $\varphi \in C_0^1(-1,1)$.

2) Take $f(x) = 0$, $x \in (-1,0]$, $f(x) = 1$, $x \in (0,1)$. Then obviously $f \in L^2(-1,1)$. Guess: $f'(x) = 0$; not good because $(f', \varphi) = 0$ for all test functions φ, while (f, φ') does not have to be zero.

It turns out that in this case the weak derivative does not exist. Indeed assume that for some $g \in L^2(-1,1)$ and for all $\varphi \in C_0^1(-1,1)$

$$\int_{-1}^{1} g(x)\,\varphi(x)\,dx = -\int_{-1}^{1} f(x)\,\varphi'(x)\,dx = -\int_{0}^{1} \varphi'(x)\,dx = \varphi(0).$$

Take a sequence φ_n such that $\varphi_n(0) = 1$ and $\varphi_n(x) \to 0$ for all $x \neq 0$. Then the limit of the left-hand side is zero, while the right-hand side is identically one, a contradiction.

We list the important properties of H^1.

Theorem. (Properties of $H^1(0,1)$.)

(a) $H^1(0,1)$ is a vector space.

(b) The quantity

$$\langle f, g\rangle = (f,g) + (f',g')$$

is an inner product, and

$$\|f\|_{H^1} = \sqrt{(f,f)+(f',f')}$$

is a norm.

(c) Using the norm we define the convergence in H^1: $f_n \to f$ in H^1 means $\|f_n - f\|_{H^1} \to 0$. Then $f_n \to f$ in $H^1 \Leftrightarrow f_n \to f$ in L^2 and $f_n' \to f'$ in L^2.

(d) The space $H^1(0,1)$ is complete with respect to the convergence in H^1.

(e) $\forall\, f \in H^1(0,1)\ \ \exists\, \varphi_n \in C^1[0,1]$ such that $\|f - \varphi_n\|_{H^1} \to 0$. (Any H^1 function can be approximated by C^1 functions.)

(f) If $f \in H^1(0,1)$ then for x_1 and x_2 outside a set of measure zero in $(0,1)$ we have

$$f(x_2) - f(x_1) = \int_{x_1}^{x_2} f'(y)\,dy$$

(the fundamental theorem of calculus).

Parts (a), (b) and (c) are easy to prove; part (d) is one of the homework problems. Parts (e), (f) will be left without proof.

Notice that H^1 is <u>not</u> complete with respect to the L^2 norm, that's the reason why we need to define the norm differently. To see this we may take the step function $f(x)$

from the second part of the last example, and approximate it in L^2 by a sequence of smooth functions φ_n. Then $\varphi_n \in H^1$ but $f = \lim_{n\to\infty} \varphi_n$ is not in H^1.
Also, unlike the case of L^2 we cannot approximate any function in H^1 by functions with compact support; only some of them will enjoy this property.

Definition. The Sobolev space $H_0^1(0,1)$ is composed of all functions $f \in H^1(0,1)$ that can be approximated in H^1 by smooth functions with compact support. Using the set notation,

$$H_0^1(0,1) = \{f \in H^1(0,1) : \exists\, \varphi_n \in C_0^1(0,1),\ \|\varphi_n - f\|_{H^1} \to 0\}.$$

The next theorem will give a useful desciption of Sobolev functions on $(0,1)$.

Theorem. Let $f \in H^1(0,1)$. Then

(a) There exists a function $\tilde{f} \in C[0,1]$ such that $f(x) = \tilde{f}(x)$ for almost all $x \in (0,1)$.

(b) If $f_n \in H^1(0,1)$ is such that $f_n \to f$ in H^1 then $\tilde{f}_n \to \tilde{f}$ uniformly on $[0,1]$.

(c) If $f \in H_0^1(0,1)$ then $\tilde{f}(0) = \tilde{f}(1) = 0$.

Another way to phrase part (a) would be to say that if $f \in H^1(0,1)$ then it is continuous up to a modification on a set of measure zero ("we just need to comb it a little bit"). Part (c) means that H_0^1 is the subspace of H^1 for which we have homogeneous boundary conditions.

Proof. (a) We use part (f) of the previous theorem which tells us that removing a set Y that has measure zero in $(0,1)$ we have

$$f(x_2) - f(x_1) = \int_{x_1}^{x_2} f'(y)\, dy, \quad \forall\, x_1, x_2 \in (0,1) \setminus Y.$$

We fix $x_1 \in (0,1) \setminus Y$ and define

$$\tilde{f}(x) = f(x_1) + \int_{x_1}^{x} f'(y)\, dy, \quad x \in [0,1].$$

Then $f(x) = \tilde{f}$ outside the set Y that has measure zero. We next show that $\tilde{f} \in C[0,1]$. Indeed, if $x \in (0,1)$ then for any h small enough

$$|\tilde{f}(x+h) - \tilde{f}(x)| \leqslant \Big| \int_x^{x+h} f'(y)\, dx \Big| \leqslant \Big(\Big| \int_x^{x+h} f'(y)^2\, dx \Big| \Big)^{1/2} \Big(\Big| \int_x^{x+h} 1\, dx \Big| \Big)^{1/2}.$$

Here we used the Cauchy-Schwarz; the absolute value is needed in the case $h < 0$. Since $f'(y)^2 \geq 0$ and $f' \in L^2(0,1)$ we have

$$|\tilde{f}(x+h) - \tilde{f}(x)| \leqslant \Big(\int_0^1 f'(y)^2 \, dx \Big)^{1/2} \sqrt{|h|} = \|f'\| \sqrt{|h|} \to 0, \text{ as } h \to 0$$

If $x \in \{0,1\}$ then the same calculation goes through if we take h of good sign ($h > 0$ for $x = 0$ and $h < 0$ if $x = 1$).

(b) Assume that $f_n \in H^1(0,1)$ is such that $f_n \to f$ in H^1. Then also $\tilde{f}_n \to \tilde{f}$ in H^1. Since $\tilde{f}_n - \tilde{f} \to 0$ in H^1 we can assume without loss of generality that $\tilde{f} = 0$, so $\tilde{f}_n \to 0$. For any $x, x_0 \in (0,1)$ we have

$$\tilde{f}_n(x) = \tilde{f}_n(x_0) + \int_{x_0}^{x} f_n'(y) \, dy.$$

By the computation in part (a), then

$$|\tilde{f}_n(x)| \leqslant |\tilde{f}_n(x_0)| + \|f_n'\| \sqrt{|x - x_0|}.$$

We can take the supremum over x and then integrate over $x_0 \in (0,1)$, so

$$\sup_{x \in (0,1)} |\tilde{f}_n(x)| \leqslant \int_0^1 |\tilde{f}_n(x_0)| \, dx_0 + \|f_n'\| \leqslant \|f_n\| + \|f_n'\| \to 0, \text{ as } n \to \infty.$$

Since $\tilde{f}_n$ are continuous the supremum is actually the maximum, and it follows that $\tilde{f}_n \to 0$ uniformly on $[0,1]$.

(c) Let $f \in H_0^1(0,1)$. By definition of $H_0^1(0,1)$ there exists $\varphi_n \in C_0^1(0,1)$ such that $\varphi_n \to f$ in H^1. By part (b) $\tilde{\varphi}_n \to \tilde{f}$ uniformly on $[0,1]$, so in particular $\tilde{f}(0) = \lim_{n\to\infty} \tilde{\varphi}_n(0) = 0$ and $\tilde{f}(1) = \lim_{n\to\infty} \tilde{\varphi}_n(1) = 0$. □

7. An application to the boundary-value problem

We return to the bondary-value problem

$$\begin{aligned} \mathcal{A}u = -(au')' + bu' + cu = f \quad \text{in} \quad \Omega = (0,1) \\ u(0) = 0, \quad u(1) = 0. \end{aligned} \tag{2}$$

We assume that a, c are smooth functions, bounded on $[0,1]$, $a(x) \geq a_0 > 0$, and $c(x) \geq 0$. We take for simplicity $b = 0$; our approach would work in a more general case when b is smooth, bounded, and $c(x) - b'(x)/2 \geq 0$ holds instead of $c(x) \geq 0$.

For the right-hand side we only assume that $f \in L^2(0,1)$ (we could in fact get away with less than that...).

We look for a way to extend the approach of the energy estimate and to obtain an existence theorem.

We multiply the differential equation by a test function $\varphi \in C_0^1(0,1)$ and integrate by parts. Then

$$-\Big[u'\varphi\Big]_{x=0}^{x=1} + \int_0^1 (au'\varphi' + cu\varphi)\,dx = \int_0^1 f\varphi\,dx.$$

The terms with the end points vanish since φ has compact support in $(0,1)$. We define the bilinear form $a(u,v)$ and the linear functional $\ell(v)$ by setting

$$a(u,v) = \int_0^1 (au'v' + cuv)\,dx \quad \text{and} \quad \ell(\varphi) = \int_0^1 fv\,dx.$$

The integral identity

$$a(u,\varphi) = \ell(\varphi), \quad \forall\,\varphi \in C_0^1(0,1) \tag{3}$$

is then a formal consequence of (2). It will be called the variational formulation.

The variational formulation is equivalent to (2) for smooth solutions. Indeed, if $u \in C^2[0,1]$, $u(0) = u(1) = 0$, and (3) holds, then we are allowed to integrate by parts to obtain

$$\int_0^1 (-(au')' + cu - f)\varphi\,dx = 0, \quad \forall\,\varphi \in C_0^1(0,1).$$

But then $-(au')' + cu - f = 0$ identically on $(0,1)$ since the the integrand is continuous. We then obtain (2).

The bilinear form $a(u,v)$ and the linear functional $\ell(\varphi)$ are defined (at least) for all functions $u, \varphi \in C^1[0,1]$. In fact, we have the following.

Lemma. The bilinear form a is bounded on $H^1(0,1) \times H^1(0,1)$. The linear functional ℓ is bounded on $H^1(0,1)$.

Proof. Since we assumed a and c bounded we have

$$\begin{aligned} a(u,v) &\leqslant C\int_0^1 (u'v' + uv)\,dx \\ &\leqslant C\big(\|u'\|\,\|v'\| + \|u\|\,\|v\|\big) \leqslant C\|u\|_{H^1}\|v\|_{H^1}, \end{aligned} \tag{4}$$

where $C = \max\{\sup|a|, \sup|c|\}$. Since $f \in L^2(0,1)$ we also have

$$\ell(v) = \int_0^1 f v \, dx \leqslant \|f\| \, \|v\| \leqslant \|f\| \, \|v\|_{H^1}.$$

(In fact, we obtained that the functional ℓ is bounded on $L^2(0,1)$ which is more than we need). $\square$

Now we claim that we have

$$a(u,v) = \ell(v), \quad \forall\, v \in H_0^1(0,1). \tag{5}$$

Indeed, for any $v \in H_0^1(0,1)$ just take a sequence $\varphi_n \to v$, $\varphi_n \in C_0^1(0,1)$. Then

$$|a(u,\varphi_n) - a(u,v)| = |a(u,\varphi_n - v)| \leqslant C\|u\|_{H^1} \, \|\varphi_n - v\|_{H^1} \to 0, \text{ as } n \to \infty,$$

and similarly,

$$|\ell(\varphi_n) - \ell(v)| = |\ell(\varphi_n - v)| \leqslant \|f\| \, \|\varphi_n - v\|_{H^1} \to 0, \text{ as } n \to \infty,$$

But then

$$a(u,v) = \lim_{n\to\infty} a(u,\varphi_n) = \lim_{n\to\infty} \ell(\varphi_n) = \ell(v).$$

(The above is an example of an argument "by continuity"; what we used was essentially that $a(u,v)$ and $l(v)$ are continuous on H^1, so we are allowed to pass to the limit in the identity $a(u,\varphi_n) = \ell(\varphi_n)$.)

Definition. We say that a function $u \in H_0^1(0,1)$ is a weak solution of (2) if (5) holds.

Weak solution is a more general concept than classical solution. If we only know that u is a weak solution we cannot automatically conclude that it satisfies (2) in the classical sense, since a function in H_0^1 need not have continuous (or even weak) second derivative. However, from what we discussed previously, if u is a weak solution and $u \in C^2[0,1]$ then it is indeed a classical solution of (2).

The advantage of going to the weak formulation is that now the boundary-value problem is formulated as a "linear equation in a Hilbert space", so speaking loosely, the problem is reduced to linear algebra. More precisely, we can now use Riesz' theorem to prove the following.

Theorem. (Existence and uniqueness of the weak solution)

For any $f \in L^2(0,1)$ the problem (5) has a unique solution $u \in H_0^1(0,1)$

Proof. Let $V = H_0^1(0,1)$; we know it is a Hilbert space with the inner product $\langle u, v\rangle = (u,v) + (u', v')$. However, this is not the only possible choice of the inner product. Let's set

$$[u,v] = a(u,v) = \int_0^1 (au'v' + cuv)\,dx.$$

Then $[u,v]$ is a bilinear form on V, and it is symmetric. We also have

$$[v,v] = \int_0^1 (av'^{\,2} + cv^2)\,dx \geq a_0 \int_0^1 v'^{\,2}\,dx = a_0\|v'\|^2 \geq 0.$$

Moreover, since $v(0) = 0$,

$$v(x)^2 = \Big(\int_0^x v'(y)\,dy\Big)^2 \leqslant \int_0^x v'^{\,2}\,dx \int_0^x 1\,dx \leqslant \int_0^1 v'^{\,2}\,dx = \|v'\|^2,$$

for almost all $x \in (0,1)$. Integrating over $x \in (0,1)$ we get

$$\|v\|^2 \leqslant \|v'\|^2.$$

(This is an example of a so-called Poincaré inequality, which has an analog for higher dimensions, and other boundary conditions.) Then

$$[v,v] \geq a_0\|v'\|^2 = \frac{a_0}{2}\left(\|v'\|^2 + \|v\|^2\right) \geq \frac{a_0}{2}\,\|v\|^2_{H^1} > 0,$$

unless $v = 0$. Thus, $[v,v]$ is positive definite and thus an inner product. The corresponding norm

$$[\![v]\!] = \sqrt{[v,v]}$$

is equivalent to $\|v\|_{H^1}$: we have

$$c_1\|v\|_{H^1} \leqslant [\![v]\!] \leqslant c_2\|v\|_{H^1}, \tag{6}$$

where $c_1 = \sqrt{a_0/2}$ and $c_2 = \sqrt{\max\{\sup|a|, \sup|c|\}}$ (recall estimate (4)). For equivalent norms the notions of convergence, completeness, etc. are equivalent: we have $[\![v - v_n]\!] \to 0$ if and only if $\|v - v_n\|_{H^1} \to 0$. Thus, the vector space V is complete with respect to the norm $[\![\cdot]\!]$ and therefore is a Hilbert space.

We know that the linear functional $\ell(v)$ is bounded with respect to the norm $\|\cdot\|_{H^1}$. By the equivalence of the norms (6) it is also bounded in the norm $[\![\cdot]\!]$. Riesz' theorem now states that there is a unique $u \in V$ such that

$$\ell(v) = [u, v], \quad \forall\, v \in V.$$

But then this is precisely the definition of the weak solution (5), so we are done. $\square$