Finite Element Methods for Elliptic Problems

Variational Formulation: The Poisson Problem

March 19 & 31, 2003

1 Motivation

SLIDE 1

- The Poisson problem has a strong formulation;
  a minimization formulation; and a weak formulation.

- The minimization/weak formulations are more general than the strong formulation in terms of regularity and admissible data.

SLIDE 2

- The minimization/weak formulations are defined by: a space X; a bilinear form a; a linear form ℓ.

- The minimization/weak formulations identify

  ESSENTIAL boundary conditions,
  Dirichlet — reflected in X;

  NATURAL boundary conditions,
  Neumann — reflected in a, ℓ.

SLIDE 3

- The *points of departure* for the *finite element method* are:

  the weak formulation (more generally);
  or
  the minimization statement (if a is SPD).

2 The Dirichlet Problem

2.1 Strong Formulation

SLIDE 4

Find u such that

$$\boxed{\begin{array}{rcll} -\nabla^2 u & = & f & \text{in } \Omega \\ u & = & 0 & \text{on } \Gamma \end{array}}$$

The boundary condition $u = 0$ is denoted "homogeneous Dirichlet." We consider Neumann boundary conditions ($\frac{\partial u}{\partial n}$ imposed) in Section 3, and inhomogeneous Dirichlet boundary conditions in Section 4.

where

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

and Ω is a domain in $\mathbb{R}^2$ with boundary Γ.

In general, we require that Ω be "Lipschitzian." We recall that a function of (say) one variable, w, satisfies a Lipschitz condition if there exists a constant K such that $|w(x) - w(y)| \leq K|x - y|$ for all x, y of interest. A domain Ω is Lipschitzian if the boundary Γ at any point admits a locally Lipschitzian representation — it can't be too wiggly or singular. Note also that, unless otherwise indicated, we will be speaking of open *domains Ω (e.g., $\Omega = (0, 1)$, which does not include 0 and 1); the closure of such a domain will be denoted $\overline{\Omega}$ (e.g., $\overline{\Omega} = [0, 1]$).*

2.2 Minimization Principle

The finite element method is not based on the strong form, but rather a minimization statement or, more generally, a weak formulation. We must thus develop and understand these formulations before proceeding with the finite element method.

2.2.1 Statement

SLIDE 5

Find

$$\boxed{u = \arg \min_{w \in X} J(w)}$$

where N1

$$X = \{v \ \text{ sufficiently smooth} \mid v|_\Gamma = 0\},$$

X here is a linear space, the precise definition of which will be given shortly; we shall also make "sufficiently smooth" precise during the course of this lecture.

and

$$J(w) = \frac{1}{2} \int_\Omega \underbrace{\nabla w \cdot \nabla w}_{w_x^2 + w_y^2} \, dA - \int_\Omega f \, w \, dA \, . \qquad \boxed{\text{N2}}$$

Note 1 ***Notation***

We explain here some of the notation that we will be using. First arg min means "the argument that minimizes," that is, the minimiz*er* (as opposed to the minim*um*). The symbol $\in$ means "in the set (or space) of"; $\subset$ means "a subset or subspace of"; $\forall$ means "for all"; $\exists$ means "there exists"; $|$ (and s.t.)

means "such that." Also, $\bigcup$ and $\bigcap$ indicate "union" and "intersection," and $\setminus$ means "set minus" (i.e., $A \setminus B$ is A with B removed).

Note 2 **Functionals**

A functional takes as input a member of a set or space (here X), and returns a scalar. We summarize this in the case above as $J\colon\ X \to \mathbb{R}$, which means J takes as input a member of X, and yields as output a real number. More generally, the notation $W\colon X \to Y$ means that W is a function (or application) from X, the input (*domain*) space, to Y, the output (*range*) space; if Y is $\mathbb{R}$, W is a functional.

SLIDE 6

In words:

Over all functions w in X,
u that satisfies

$$\begin{aligned} -\nabla^2 u &= f && \text{in } \Omega \\ u &= 0 && \text{on } \Gamma \end{aligned}$$

makes $J(w)$ as small as possible. N3

We give a geometric picture in the next lecture — $J(w)$ is an infinite dimensional paraboloid, the bottom of which occurs at $w = u$ and takes on the value $J(u)$.

Note 3 **Physical interpretation**

There are many cases in which this minimization principle (also known as the Dirichlet principle) has a meaningful and intuitive significance — often an "energy statement." For example, if u is a velocity potential for incompressible flow, then (say for $f = 0$ and inhomogeneous Dirichlet conditions — see Section 4) $J(w)$ is the kinetic energy, and minimizing J thus corresponds to minimizing energy. However, there are also cases (e.g., if u is temperature) in which a physical interpretation is rather strained, more of an *a posteriori* justification than any particularly useful perspective. For our purposes here we need only the mathematical properties of the minization principle; the physical interpretation is not central.

2.2.2 Proof

SLIDE 7

Let $w = u + v$.
Then

$$J(\overbrace{\underbrace{u}_{\in X} + \underbrace{v}_{\in X}}^{w \in X}) = \frac{1}{2}\int_\Omega \nabla(u+v)\cdot\nabla(u+v)\,dA - \int_\Omega f(u+v)\,dA\,.$$

Note $u|_\Gamma = v|_\Gamma = 0$, which ensures that $w|_\Gamma = 0$, and hence is a member of X. Recall that $w|_\Gamma$ means w restricted to Γ, that is, evaluated on Γ.

SLIDE 8

$$\begin{aligned} J(u+v) &= \frac{1}{2}\int_\Omega \nabla u\cdot\nabla u\,dA - \int_\Omega f\,u\,dA && \boldsymbol{J(u)} \\ &\quad + \int_\Omega \nabla u\cdot\nabla v\,dA - \int_\Omega f\,v\,dA && \boldsymbol{\delta J_v(u)} \\ & && \textit{first variation} \\ &\quad + \frac{1}{2}\int_\Omega \nabla v\cdot\nabla v\,dA && > \mathbf{0} \text{ for } \boldsymbol{v \neq 0} \end{aligned}$$

We can think of $J(u+v)$ as a "Taylor" series about u. Since J is only quadratic, it is not surprising that $J(u+v)$ contains a constant term, a linear (in v) term (a "gradient"), and a quadratic (in v) term (a "Hessian") — and then terminates.

SLIDE 9

$$\begin{aligned} \delta J_v(u) &= \int_\Omega \nabla u\cdot\nabla v\,dA - \int_\Omega f\,v\,dA \\ &= \int_\Omega \nabla\cdot(v\nabla u)\,dA - \int_\Omega v\nabla^2 u\,dA - \int_\Omega f\,v\,dA \\ &= \int_\Gamma \overbrace{v}^{0}\,\nabla u\cdot\hat{n}\,dS + \int_\Omega v\{\underbrace{-\nabla^2 u - f}_{0}\}\,dA \\ &= 0\,, \qquad \forall\, v \in X \qquad \boxed{\text{N4}} \end{aligned}$$

We know the gradient of a function vanishes at its minimizer; it is thus not surprising that the first variation of a functional (the gradient times a test function) vanishes at its minimizer. Here $\hat{n}$ is the unit normal on Γ.

Note 4 **Gauss and Green's Theorems**

Much of our analysis here is based on humble integration by parts, which in higher space dimensions is essentially one of Green's Theorems. The necessary

result is demonstrated most easily in indicial notation. In particular, we note that

$$\begin{aligned}\int_\Omega \frac{\partial v}{\partial x_j}\frac{\partial u}{\partial x_j}\,dA &= \int_\Omega \left(\frac{\partial}{\partial x_j}\left[v\frac{\partial u}{\partial x_j}\right] - v\,\frac{\partial^2 u}{\partial x_j \partial x_j}\right) dA \\ &= \int_\Gamma v\,\frac{\partial u}{\partial x_j}\,\hat{n}_j\,dS - \int_\Omega v\,\frac{\partial^2 u}{\partial x_j \partial x_j}\,dA \\ &= \int_\Gamma v\nabla u\cdot\hat{n}\,dS - \int_\Omega v\nabla^2 u\,dA\ ,\end{aligned}$$

where we have used Gauss' Theorem to convert the volume integral into a surface term. Note we adopt the convention of summation over repeated indices, here from 1 to 2 since we are in $\mathbb{R}^2$.

SLIDE 10

$$J(\underbrace{u+v}_{w}) = J(u) + \frac{1}{2}\underbrace{\int_\Omega \nabla v\cdot\nabla v\,dA}_{>\,0 \text{ unless } v\,=\,0},\ \forall\, v\in X$$

$\Rightarrow$

$$\boxed{\begin{array}{c} J(w) > J(u)\,, \qquad \forall\, w\in X\,,\ w\neq u \\ \Updownarrow \\ u \text{ is } \textit{the} \text{ minimizer of } J(w)\end{array}}$$

E1

What PDEs admit such a minimization statement? PDEs associated with operators that are SPD (symmetric positive definite). We define this more precisely, and indicate how the FEM (finite element method) proceeds in the absence of this property, in a future lecture. For now, we focus on the simplest case — almost all of which turns out to be directly relevant to the more general case.

We could also derive the result above by applying the general machinery of the calculus of variations. In this sense, we may view $-\nabla u = f$ as the Euler or Euler-Lagrange equations associated with minimization of the functional J.

▷ ***Exercise 1*** Consider the problem $-u_{xx} = 1$, $0 < x < 1$, $u(0) = u(1) = 0$, with solution $u = \frac{1}{2}x(1-x)$. Show by explicit calculation that $\delta J_v(u) = \int_0^1 u_x v_x - v\,dx = 0$ for all (smooth) v such that $v(0) = v(1) = 0$. ■

2.3 Weak Formulation

2.3.1 Statement

SLIDE 11

Find $u \in X$ such that

$$\boxed{\delta J_v(u) = 0\,, \quad \forall\, v \in X}$$

$$\Updownarrow$$

$$\boxed{\int_\Omega \nabla u \cdot \nabla v \, dA = \int_\Omega f\, v \, dA\,, \qquad \forall\, v \in X} \;;$$

see Slide 9 for proof. N5

This equation has a great deal of structure which we cannot obviously see in this explicit statement. We thus digress to some more general mathematical definitions so that we can present a more succinct restatement. Note that the weak formulation of a PDE, in which we introduce a test function v to "absorb" some of the derivatives, will always exist (indeed is more general than the strong statement) even when no minimization principle is available — that is, even when the problem is not SPD. The weak formulation is thus the most general point of departure for the finite element method.

Note 5 ***Du Bois-Reymond lemma***

In fact, we have already derived the weak statement: we know from Slide 9 that if u satisfies $-\nabla^2 u = f$ in Ω, $u|_\Gamma = 0$, then $\delta J_v(u) = 0$, $\forall\, v \in X$; the latter is simply (defined to be) the weak statement.

We might ask whether we can go "the other way," that is, show that if $u \in X$ satisfies $\delta J_v(u) = 0$, $\forall\, v \in X$, then u satisfies $-\nabla^2 u = f$ in Ω. Yes: By integration by parts we know that

$$\int_\Omega \nabla u \cdot \nabla v \, dA = \int_\Gamma \cancelto{0}{v} \, \nabla u \cdot \hat{n} \, dS - \int_\Omega v \nabla^2 u \, dA \;,$$

and thus

$$\int_\Omega \nabla u \cdot \nabla v - fv \, dA = \int_\Omega v\{-\nabla^2 u - f\} \, dA = 0, \qquad \forall\, v \in X \;.$$

Now assume that $-\nabla^2 u - f$ does not equal zero at some point; we can then take v nonzero localized about this point, which contradicts $\delta J_v(u) = 0$, $\forall\, v \in X$. We thus conclude that $-\nabla^2 u = f$ in Ω; this is known (in certain circles) as the Du Bois-Reymond lemma.

2.3.2 Definitions

SLIDE 12

Linear space, Y:

A set Y is a linear (or vector) space
if

$$\boxed{\begin{array}{lll} & \forall\, v_1, v_2 \in Y, & v_1 + v_2 \in Y \\ \forall\, \alpha \in \mathbb{R}, & \forall\, v \in Y, & \alpha v \in Y \end{array}} \; .$$

SLIDE 13

Linear forms, $L(v)$:

L: $\underbrace{Y}_{\text{input}} \to \underbrace{\mathbb{R}}_{\text{output}}$ (*form* or *functional*)

$L(\alpha v_1 + v_2) = \alpha L(v_1) + L(v_2)$ (*linear*)
$\forall\, \alpha \in \mathbb{R}\,, \quad \forall\, v_1, v_2 \in Y$.

SLIDE 14

Bilinear forms, $B(w, v)$:

B: $Y \times Z \to \mathbb{R}$ (*form*) ;

$B(w, \overline{v})$ linear form in w for fixed $\overline{v}$,
$B(\overline{w}, v)$ linear form in v for fixed $\overline{w}$ (*bilinear*) .

Note that B: $Y \times Z \to \mathbb{R}$ indicates that B has two inputs (arguments), the first from the space Y, the second from the space Z; the output is a real number.

SLIDE 15

SPD bilinear forms, $B(w, v)$:

B: $Y \times Y \to \mathbb{R}$ is *bilinear* ;

$B(w, v) = B(v, w)$ *S*PD ;

$B(w, w) > 0\,, \qquad \forall\, w \in Y\,,\ w \neq 0$ S*PD* .

2.3.3 Restatement

SLIDE 16

Let

$$a(w, v) = \int_\Omega \nabla w \cdot \nabla v \, dA\,, \quad \forall\, w, v \in X$$

an *SPD bilinear form* E2

and

$$\ell(v) = \int_\Omega f\, v\, dA, \quad \forall\, v \in X$$

a *linear form* .

▷ ***Exercise 2*** Prove that a is indeed an SPD bilinear form over X. *Hint*: you must use the boundary conditions. (Note a is SPD because the underlying operator is SPD.) ■

SLIDE 17

Minimization Principle:

$$u = \arg \min_{w \in X} \underbrace{\frac{1}{2}\, a(w,w) - \ell(w)}_{J(w)} \ .$$

Weak Statement: $u \in X$,

$$\underbrace{a(u,v) = \ell(v)}_{\Leftrightarrow\ \delta J_v(u)=0}, \qquad \forall\, v \in X \ . \qquad \text{E3}$$

▷ ***Exercise 3***

(a) Show that if J: $Y \to \mathbb{R}$ is defined by $J(w) = \frac{1}{2}a(w,w) - \ell(w)$ for *any* SPD bilinear form a and linear form ℓ over Y, then the minimizer u satisfies $a(u,v) = \ell(v)$, $\forall\, v \in Y$. (In this way, given a weak statement, one can "anti-variation" to find J.)

(b) Take $Y = \mathbb{R}^n$, and thus show, by appropriate choice of a and ℓ, that the minimizer $u \in Y$ of $J(w) = \frac{1}{2}w^T G w - w^T F$ — for any SPD matrix $G \in \mathbb{R}^{n\times n}$ and $F \in \mathbb{R}^n$ — satisfies $Gu = F$.

■

2.3.4 Proper Spaces: $u \in X$

SLIDE 18

Since a involves *only first derivatives* ,

$$X = \{v \in H^1(\Omega) \mid v|_\Gamma = 0\} \equiv H_0^1(\Omega)\text{:}$$

$$H^1(\Omega) \equiv \{v \mid \int_\Omega v^2\, dA\,,\ \int_\Omega v_x^2\, dA\,,\ \int_\Omega v_y^2\, dA \text{ finite}\}\ ;$$

$$\underbrace{(w,v)_{H^1(\Omega)}}_{\text{inner product}} = \int_\Omega \nabla w \cdot \nabla v + wv \, dA \; ;$$

$$\underbrace{\|w\|_{H^1(\Omega)}}_{\text{norm}} = \left(\int_\Omega |\nabla w|^2 + w^2 \, dA \right)^{1/2} . \qquad \boxed{\text{N6}} \; \boxed{\text{E4}}$$

Important theoretical *and* numerical *implications.*

Note 6 **Important spaces, inner products, and norms**

Hilbert and Banach Spaces

A Hilbert space is a *linear space* Y with which we associate an *inner product* $(\cdot,\cdot)_Y$ — this is simply an SPD bilinear form — which then induces a *norm*, $\|w\|_Y \equiv (w,w)^{1/2}$. In fact, what we have just described is an inner product space: a Hilbert space is a *complete* inner product space; by completeness we mean that any Cauchy sequence ($y_n \in Y$ such that $\|y_n - y_m\|_Y \to 0$ as $n, m \to \infty$) converges to a member of Y.

A Hilbert space is a special case of a Banach space Z, which is a (complete) *normed* linear space. The norm $\|\cdot\|_Z$ associated with a Banach space is not, in general, induced from any bilinear form, but must still satisfy certain conditions (the conditions we intuitively associate with any measure of "length"):

$$\begin{aligned} \|w\|_Z &> 0 && \forall\, w \in Z,\ w \neq 0 \, , \\ \|\alpha w\|_Z &= |\alpha| \, \|w\|_Z, && \forall\, \alpha \in \mathbb{R},\ \forall\, w \in Z \, , \\ \|w+v\|_Z &\leq \|w\|_Z + \|v\|_Z && \forall\, w \in Z,\ \forall\, v \in Z \, , \end{aligned}$$

the last being the triangle inequality (the shortest distance between two points ...).

It can readily be shown that a norm induced by an inner product automatically satisfies the above conditions. The triangle inequality is proven with the help of the Cauchy-Schwarz inequality, which states that for an inner product $(\cdot,\cdot)_Y$,

$$(w,v)_Y \leq \|w\|_Y \, \|v\|_Y \, .$$

We give the proof here:

$$\begin{aligned} 0 \leq \left\| w - \frac{(w,v)_Y}{\|v\|_Y^2} \, v \right\|_Y^2 &= \left(w - \frac{(w,v)_Y}{\|v\|_Y^2} \, v, w - \frac{(w,v)_Y}{\|v\|_Y^2} \, v \right)_Y \\ &= \|w\|_Y^2 - 2\frac{(w,v)_Y^2}{\|v\|_Y^2} + \frac{(w,v)_Y^2}{\|v\|_Y^2} \\ &= \|w\|_Y^2 - \frac{(w,v)_Y^2}{\|v\|_Y^2} \; ; \end{aligned}$$

so $(w,v)_Y^2 \leq \|w\|_Y^2 \, \|v\|_Y^2$ upon multiplying by $\|v\|_Y^2$. We obtain strict equality if w is proportional to v.

Spaces $H_0^1(\Omega), H^1(\Omega), H^m(\Omega)$

It is a simple matter to show that $H_0^1(\Omega)$ and $H^1(\Omega)$ are inner product spaces. It is decidedly less simple to show that these spaces are complete, though they do indeed have this property, and hence are indeed Hilbert spaces. (We note that an alternative view of $H_0^1(\Omega)$ in fact *defines* this space as the completion of a class of infinitely smooth functions with respect to the H^1 norm.) Completion (and relatedly, closure, which ensures that the limit of a sequence of members of a subset is itself a member of the subset) is a rather theoretical notion that we will simply take for granted; but it *is* important, making sure that we do not find ourselves in a situation in which the limit of a sequence of functions has very different properties than each member of the sequence. We will briefly discuss the latter again below.

We can easily generalize the spaces $H^1(\Omega)$ to $H^m(\Omega)$ for any non-negative integer m. We do so in $\mathbb{R}^1$, say $\Omega = (0,1)$, to avoid multi-indices for mixed derivatives in higher dimensions. We thus have that

$$H^m(\Omega) = \left\{ v \mid \int_0^1 v^2 \, dx < \infty, \ \int_0^1 v_x^2 \, dx < \infty, \ldots, \int_0^1 \left(\frac{d^m v}{dx^m} \right)^2 dx < \infty \right\} ,$$

with associated inner product

$$(w,v)_{H^m(\Omega)} = \sum_{j=0}^{m} \int_0^1 \frac{d^j w}{dx^j} \, \frac{d^j v}{dx^j} \, dx ,$$

and norm

$$\|w\|_{H^m(\Omega)} = \left(\sum_{j=0}^{m} \int_0^1 \left(\frac{d^j w}{dx^j} \right)^2 dx \right)^{1/2} .$$

These spaces are important not only in understanding well-posedness of weak statements, but also in expressing the convergence rate of the finite element method. We shall have most need for $H_0^1(\Omega)$, $H^1(\Omega)$, and $H^2(\Omega)$; the latter requires that, in addition to the function and the derivative, the second derivative must also have finite energy.

We also introduce the $H^m(\Omega)$ *semi-norm* as

$$|w|_{H^m(\Omega)} = \left(\int_0^1 \left(\frac{d^m w}{dx^m} \right)^2 dx \right)^{1/2} ,$$

which includes *only* the m^{th} derivative. The H^1 semi-norm is simply $|w|_{H^1(\Omega)} = \left(\int_0^1 w_x^2 \, dx \right)^{1/2}$. In general, a semi-norm must satisfy all the properties of a norm

except that it is permitted to vanish for $w \neq 0$. However, for the case of $H_0^1(\Omega)$, the H^1 semi-norm is *equivalent* to the full H^1 norm, by which we mean that

$$C_{\rm PF} \, \|w\|_{H^1(\Omega)} \leq |w|_{H^1(\Omega)} \leq \|w\|_{H^1(\Omega)}, \qquad \forall \, w \in H_0^1(\Omega) \; ;$$

the condition that $w \in H_0^1(\Omega)$ must vanish on Γ ensures that w is not free to "float." The left-hand inequality is known as the Poincaré-Friedrichs inequality, and can also be related to the minimum eigenvalue of the Dirichlet Laplacian problem through the Rayleigh quotient. We shall prove this result for a particular problem in a later lecture.

The Lebesgue spaces, $L^p(\Omega)$

Another set of spaces that are very important in finite element analysis are the Lebesgue spaces, $L^p(\Omega)$, $p \geq 1$, which are *not* Hilbert spaces except in the particular (and perhaps most important) case $p = 2$: for $p = 2$, $L^2(\Omega) \equiv H^0(\Omega)$, with inner product

$$(w, v)_{L^2(\Omega)} = \int_\Omega w \, v \, dA$$

and norm

$$\|v\|_{L^2(\Omega)} = \left(\int_\Omega v^2 \, dA \right)^{1/2} .$$

More generally,

$$L^p(\Omega) = \left\{ v \mid \int_\Omega |v|^p \, dA < \infty \right\} \; ,$$

with norm

$$\|v\|_{L^p(\Omega)} = \left(\int_\Omega |v|^p \, dA \right)^{1/p} \; ;$$

it can be shown by the Holder inequality that $L^q(\Omega) \subset L^p(\Omega)$ for $q \geq p$.

The sense of integration here is (appropriately enough) Lebesgue integration. We do not enter into the theory of integration here except to note that Lebesgue integration is very forgiving of (very) occasional omissions — the Lebesgue integral will not change if we change the value of the integrand only on a set of "zero measure" (e.g., a point, or a finite set of points, or a countably infinite set of points).

We can see the practical import of the Lebesgue definition of integration by considering $L^\infty(\Omega)$, that is, $L^p(\Omega)$ as $p \to \infty$. It is clear that as $p \to \infty$ our norm will pick up only the largest value of $|v|$, so we might be tempted to write

$$\|v\|_{L^\infty(\Omega)} = \sup_{x \in \Omega} |v| \; ,$$

where we recall that sup means we look for the least upper bound, that is, the smallest value C such that $|v(x)| \leq C$ for all $x \in \Omega$. However the above would say that, for a function v which is zero everywhere but equal to, say, 10 at one

point, $\|v\|_{L^\infty(\Omega)} = 10$; however, since the Lebesgue integral does not "see" the 10, the correct answer is zero. We thus need to write

$$\|v\|_{L^\infty(\Omega)} = \text{ess } \sup_{x\in\Omega} |v|$$

where ess sup (essential supremum) means the smallest supremum over $\Omega \setminus B$ (Ω excluding B) for all sets B of zero measure.

Although we are used to thinking of the laws of physics (and the equations we use to describe them) as being satisfied at each and every point, in fact we know that this is not the case — the (say) continuum of solid mechanics and fluid mechanics is an idealization only meaningful at a certain (supra-molecular) scale. The weak form and Lebesgue integration is, in some sense, a mathematical description of this "local averaging."

Sobolev spaces

We describe Sobolev spaces for the simple case of $\Omega = (0,1) \in \mathbb{R}^1$. Then $W^{m,p}(\Omega)$ for $m \geq 0$ integer and $p \geq 1$ is given by

$$W^{m,p}(\Omega) = \left\{ v \mid \frac{d^j v}{dx^j} \in L^p(\Omega), \quad j = 0, \ldots, m \right\} ,$$

with norm

$$\|w\|_{W^{m,p}(\Omega)} = \left(\sum_{j=0}^{m} \int_0^1 \left| \frac{d^j w}{dx^j} \right|^p \, dx \right)^{1/p} .$$

Essentially, the $W^{m,p}(\Omega)$ norm measures the first m derivatives of w in the L^p norm.

We note that $W^{m,2}(\Omega) = H^m(\Omega)$, our earlier Hilbert spaces. For $p \neq 2$, the Sobolev spaces are not Hilbert spaces. For $m = 0$, the $W^{0,p}(\Omega)$ spaces are simply the Lebesgue spaces $L^p(\Omega)$.

$C^m(\Omega)$ spaces

We have already encountered these spaces in the finite-difference context. For the case (say) of $\Omega = (0,1) \subset \mathbb{R}^1$,

$$C^m(\Omega) = \left\{ v \mid v, \frac{dv}{dx}, \ldots, \frac{d^m v}{dx^m} \text{ continuous in } \Omega \right\} ,$$

defined for any integer m. Note $C^0(\Omega)$ is the space of continuous functions; $C^{-1}(\Omega)$ is the space of functions whose antiderivative is continuous over Ω; and $C^\infty(\Omega)$ is the set of functions in which all derivatives exist and are continuous over Ω.

There is an important relationship between $C^0(\Omega)$ and Sobolev spaces, known as the Sobolev embedding theorem. This theorem tells us, for example, that for any regular domain $\Omega \subset \mathbb{R}^d$, if $v \in H^m(\Omega)$, $m > \frac{d}{2}$, then $v \in C^0(\Omega)$, and

$$\|v\|_{L^\infty(\Omega)} \leq C \, \|v\|_{H^m(\Omega)} ,$$

where C does not depend on v. For example, if $d = 1$, $u \in H^1(\Omega)$ implies u is continuous, since $m = 1 > \frac{d}{2} = \frac{1}{2}$; however for $d = 2$, one can find functions $u \in H^1(\Omega)$ which are not continuous — unbounded — as $m = 1 \not> \frac{d}{2} = 1$. This has important practical ramifications.

Apology

This section contains much material, the relevance of all of which is probably not clear. The student should read it once now, and then on several occasions during this sequence of finite element lectures. The student is only "responsible" for understanding those bits that enter into results of this and later lectures. The rest is included only to make sure that when these various entities are encountered in the finite element literature — *which they surely will be* — they will look less intimidating.

▷ ***Exercise 4*** True or false?

(a) The set $S = \{v \in C^0(0,1)) \mid v(\frac{1}{2}) = 1\}$ is a linear space.

(b) For $X = H_0^1((0,1))$, $L(v) = \int_0^1 xv\, dx$ is a linear functional.

(c) For $Z \equiv \mathbb{R}$, $(x,y)_Z = |x|\,|y|$ is a valid inner product (SPD bilinear form).

(d) The only w in $H^1(\Omega)$ for which $|w|_{H^1(\Omega)}$ (the H^1 semi-norm) is zero is $w = 0$.

(e) The function $x^{3/4}$ is in $L^2((0,1))$; in $H^1((0,1))$; in $H^2((0,1))$.

(f) For $w = e^{-10x}$, $|w|_{H^2((0,1))} = |w|_{H^1((0,1))}$.

■

2.3.5 Proper Spaces: $\ell \in X'$

SLIDE 19

The "data" ℓ: $H_0^1(\Omega) \to \mathbb{R}$ must satisfy

$$|\ell(v)| \le C\, \|v\|_{H^1(\Omega)},\ \forall\, v \in H_0^1(\Omega)\ (\textit{bounded}).$$

$\ell \in$ *dual space* $X' = (H_0^1(\Omega))' \equiv H^{-1}(\Omega)$:

all linear functionals bounded for $v \in H_0^1(\Omega)$.

Dual norm: $$\|\ell\|_{(H_0^1(\Omega))'} = \sup_{v \in H_0^1(\Omega)} \frac{\ell(v)}{\|v\|_{H^1(\Omega)}}\ .$$ N7 N8

Again, this result looks abstract, but it has important practical implications — what "loads" (heat sources, ...) are we allowed to consider? It will also tell us what outputs we can accurately measure (numerically). We will see that the space $H^{-1}(\Omega)$ is quite large, admitting rather "irregular" functions; $u \in H^1(\Omega)$

is, however, relatively smooth. This is not surprising since $-\nabla^2 u = f$ — u is "two integrals" smoother than f.

Note 7 **Dual spaces (Optional)**

In general, given a Hilbert space Y, we can define the dual space Y' as the space of all *bounded* linear functionals, $L(v)$, where $L(v)$ is bounded if $L(v) \leq C \|v\|_Y$, $\forall\, v \in Y$. The norm of $L(v)$ is given by

$$\|L\|_{Y'} = \sup_{v \in Y} \frac{L(v)}{\|v\|_Y} .$$

Clearly this space of functionals is a linear space, since the *sum* of two bounded linear functionals is also a bounded linear functional, as is a *scalar multiple* of any bounded functional. Note a bounded linear functional is *continuous*: $|\ell(v) - \ell(w)| = |\ell(v - w)| \leq C \|v - w\|_Y \Rightarrow \ell(v) \to \ell(w)$ as $v \to w$.

As an example of a bounded functional, take $\Omega = (0, 1)$, $Y = L^2(\Omega)$, and $L(v) = \int_0^1 v\, dx$. We now note that

$$L(v) = \int_0^1 1\, v\, dx \leq \left(\int_0^1 1^2\, dx \right)^{1/2} \left(\int_0^1 v^2\, dx \right)^{1/2} \leq \|v\|_{L^2(\Omega)} ,$$

where we have used the Cauchy-Schwarz inequality in the L^2 inner product, $(w, v)_{L^2(\Omega)} = (w, v)_{H^0(\Omega)} = \int_0^1 wv\, dx$, with $w = 1$. Thus $L(v)$ is in the dual space of $L^2(\Omega)$.

The dual space of $L^2(\Omega)$ turns out to be $L^2(\Omega)$, in the sense that any functional $L_\eta(v)$ of the form $\int_0^1 \eta v\, dx$ for $\eta \in L^2(\Omega)$ is bounded in $L^2(\Omega)$ (and conversely, any bounded functional in $L^2(\Omega)$ can be expressed in this form). Indeed, by the Cauchy-Schwarz inequality, $L_\eta(v) \leq \|\eta\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}$, $\forall v \in L^2(\Omega)$, with equality for $v = \eta$: the norm of $L_\eta(v)$ is thus $\|\eta\|_{L^2(\Omega)}$ — the smallest constant C for which $L_\eta(v) \leq C \|v\|_{L^2(\Omega)}$, $\forall\, v \in L^2(\Omega)$. We can also compute the norm of L_η in $(L^2(\Omega))'$ directly from the definition

$$\|L_\eta(v)\|_{(L^2(\Omega))'} = \sup_{v \in L^2(\Omega)} \frac{\int_0^1 \eta v\, dx}{\|v\|_{L^2(\Omega)}} ;$$

from the Cauchy-Schwarz inequality (again ...) the sup is attained for $v = \eta$, and thus $\|L_\eta(v)\|_{(L^2(\Omega))'} = \|\eta\|^2_{L^2(\Omega)} / \|\eta\|_{L^2(\Omega)} = \|\eta\|_{L^2(\Omega)}$. In this particular case, the sup expression above is just a fancy way of writing the usual L^2 norm, $(\int_0^1 \eta^2 dx)^{1/2}$. Note that since the weak formulation for the Poisson problem only requires (on the right-hand side of our equation) a linear functional, the distinction between the function η and its associated linear functional L_η is blurred — indeed, there really is no distinction, L_η and η are interchangeable.

We now consider a more difficult case: $\Omega = (0,1)$, $Y = H_0^1(\Omega)$, with $L(v) = v(\frac{1}{2})$. In fact, this is the delta distribution, since $v(\frac{1}{2}) =$"$\int_0^1 \delta(x - \frac{1}{2})\, v\, dx$," where $\delta(x - \frac{1}{2}) = 0$ for $x \neq \frac{1}{2}$, $\delta(x - \frac{1}{2}) = \infty$ at $x = \frac{1}{2}$, and $\delta(x - \frac{1}{2})$ is of unit mass, "$\int_0^1 \delta(x - \frac{1}{2})\, dx = 1$." (In fact, $\delta(x - \frac{1}{2})$ is not an integrable function, as we discuss below.) Now, since $v(0) = 0$ for $v \in H_0^1(\Omega)$,

$$v(\tfrac{1}{2}) = \int_0^{1/2} v'\, dx = \int_0^{1/2} 1\, v'\, dx \quad < \quad \left(\int_0^{1/2} 1^2\, dx \right)^{1/2} \left(\int_0^{1/2} (v')^2\, dx \right)^{1/2}$$

$$< \quad \frac{1}{\sqrt{2}} \|v\|_{H^1(\Omega)} \ ,$$

and thus $L \in (H_0^1(\Omega))'$ (though our constant $\frac{1}{\sqrt{2}}$ is not the norm); it is clearly *not* in $(L^2(\Omega))'$, however, since we can make the $L^2(\Omega)$ norm of v as small as we like while keeping $v(\frac{1}{2})$ fixed, and we can thus not bound $L(v)$ by $C\|v\|_{L^2(\Omega)}$, $\forall v \in L^2(\Omega)$, for any finite C. (Note the delta distribution is also not in $(H_0^1(\Omega))'$ for $\Omega \subset \mathbb{R}^2$ — this is a consequence of the Sobolev embedding theorem.)

We observe from the above example that $(H_0^1(\Omega))'$ (which is *not* equal to $H_0^1(\Omega)$) is larger than $(L^2(\Omega))'$. This is not surprising, since

$$\|L\|_{(H_0^1(\Omega))'} = \sup_{v \in H_0^1(\Omega)} \frac{L(v)}{\|v\|_{H^1(\Omega)}}$$

has *more derivatives* in the denominator than the corresponding L^2 expression, and hence *more* $L(v)$ will be bounded. This notion is consistent with the fact the $(H_0^1(\Omega))'$ is denoted $H^{-1}(\Omega)$, where $H^{-1}(\Omega)$ may be interpreted as the space of functions for which only the *anti*derivatives need be in $L^2(\Omega)$ (this is the case for the delta function, since the antiderivative is the Heaviside function). The dual spaces of $H^m(\Omega)$ — denoted $H^{-m}(\Omega)$ — are thus increasingly large as m increases (the $H^m(\Omega)$ of course become increasingly smaller). So we have $H_0^1(\Omega) \subset L^2(\Omega)$, but $(H_0^1(\Omega))' \supset (L^2(\Omega))' = L^2(\Omega)$.

Finally, the Hilbert spaces Y and Y' are related by the celebrated Riesz representation theorem. For each $L \in Y'$, there exists a unique u_L in Y such that

$$(u_L, v)_Y = L(v), \qquad \forall\, v \in Y \ .$$

It follows that

$$\|L\|_{Y'} = \sup_{v \in Y} \frac{(u_L, v)_Y}{\|v\|_Y} = \|u_L\|_Y \ ,$$

by the Cauchy-Schwarz inequality applied in the Y inner product. Note for $Y = Y' = L^2(\Omega)$, u_L is simply η of L_η introduced above, and we recover directly $\|L\|_{L^2(\Omega)'} = \|\eta\|_{L^2(\Omega)}$. This also proves our claim that any L^2 functional can be expressed as $L_\eta = \int_0^1 \eta\, v\, dx$ for some (unique) $\eta \in L^2(\Omega)$.

Dual spaces also play an important role in another context, which we illustrate in $\mathbb{R}^1$ for $\Omega = (0,1)$. The space $C_0^\infty(\Omega)$ is the space of $C^\infty(\Omega)$ functions v such that v and *all* of its derivatives vanish at $x = 0$ and $x = 1$. The space of *distributions* is the dual space to $C_0^\infty(\Omega)$ — all continuous linear functionals $L(v)$ for $v \in C_0^\infty(\Omega)$. (Here continuity is defined as $L(v_n) \to L(v)$ for all sequences v_n that converge uniformly in all derivatives to v.) Since $C_0^\infty(\Omega)$ is a very "small" space (with many derivatives), we expect from our earlier arguments that our dual space will be very large — and indeed it is.

As one example, if η is any function in $L^2(\Omega)$, then $L_\eta(v) = \int_0^1 \eta v \, dx$ is a distribution. We know that L_η is also a member of $(L^2(\Omega))' = L^2(\Omega)$. But we may also consider other linear functionals that do not correspond to any L^2 function: for example, the delta distribution $L_{\delta_{x_0}}$ may be defined as

$$L_{\delta_{x_0}}(v) = v(x_0) \ .$$

It is not strictly correct to write $L_{\delta_{x_0}}(v) = \int_0^1 \delta(x - x_0)\, v \, dx$, since $\delta(x - x_0)$ is not an integrable function; and we know from Note 7 that $L_{\delta_{x_0}}$ is not in $(L^2(\Omega))' = L^2(\Omega)$. However, we can write

$$L_{\delta_{x_0}}(v) = \langle L_{\delta_{x_0}}, v \rangle$$

where $\langle \cdot, \cdot \rangle$, defined by the above expression, is an "integral-like" *duality pairing* — a proper replacement for the L^2 inner product — here between $(C_0^\infty)'$ (for L) and C_0^∞ (for v). The essential point is not what notation we use, $L(v)$ or $\langle L, v \rangle$, but rather that we appreciate that both represent a distribution, which is a particular (large) class of linear functionals. Given a v, we need only know how to evaluate $L(v)$ — quite simple in the case of the delta distribution. In our two examples here, our distributions L_η and $L_{\delta_{x_0}}$ are in $L^2(\Omega)$ and $H^{-1}(\Omega)$, respectively; but, in general, many distributions not only will not be in $L^2(\Omega)$, but *also* will not be in $H^{-1}(\Omega)$ — see Exercise 7 for an example.

We can now define the m^{th} distributional derivative of L, $D^m L$, as

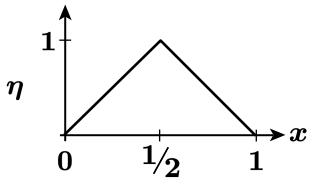
$$D^m L(v) = \langle D^m L, v \rangle \equiv (-1)^m \left\langle L, \frac{d^m v}{dx^m} \right\rangle \ ,$$

where the right-hand side clearly exists since $v \in C_0^\infty(\Omega)$ — we have put the derivatives on the *smooth* member. We shall say that a function η has m distributional derivatives in $L^2(\Omega)$ if

$$\langle D^m L_\eta, v \rangle = \int_0^1 \widetilde{D^m \eta}\, v \, dx$$

for some L^2 function $\widetilde{D^m \eta}$; recall that $\langle L_\eta, v \rangle = \int_0^1 \eta \, v \, dx$.

As a concrete example, take $\eta = 1 - 2|x - \frac{1}{2}|$, which looks like



Now, $L_\eta(v) = \int_0^1 \eta v \, dx$, so (taking $m = 1$)

$$\langle D^1 L_\eta, v\rangle = -1 \int_0^1 \eta \, \frac{dv}{dx} \, dx = -1 \lim_{\varepsilon \to 0} \left(\int_0^{\frac{1}{2}-\varepsilon} \eta \, \frac{dv}{dx} \, dx + \int_{\frac{1}{2}-\varepsilon}^1 \eta \, \frac{dv}{dx} \, dx \right) ,$$

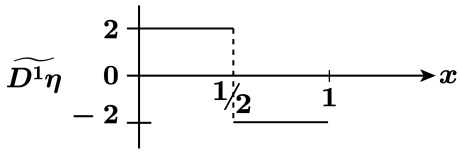
from the properties of Lebesgue integration (we can omit any point ...). But, by integration by parts,

$$\begin{aligned} \int_0^{\frac{1}{2}-\varepsilon} \eta \, \frac{dv}{dx} \, dx &= \eta(\tfrac{1}{2}-\varepsilon)\, v(\tfrac{1}{2}-\varepsilon) - \int_0^{\frac{1}{2}-\varepsilon} 2\, v \, dx \\ \int_{\frac{1}{2}-\varepsilon}^1 \eta \, \frac{dv}{dx} \, dx &= -\eta(\tfrac{1}{2}-\varepsilon)\, v(\tfrac{1}{2}-\varepsilon) - \int_{\frac{1}{2}-\varepsilon}^1 -2\, v \, dx , \end{aligned}$$

since $v(0) = v(1) = 0$. Thus, as $\varepsilon \to 0$, we find

$$\langle D^1 L_\eta, v\rangle = \int_0^1 \widetilde{D^1\eta}\, v \, dx$$

where $\widetilde{D^1\eta}$ is a Heaviside-like function



exactly as we would have expected. Note the value of $\widetilde{D^1\eta}$ at $x = \frac{1}{2}$ is irrelevant since it will not affect the integral.

We thus conclude that our function η has a first distributional derivative in $L^2(\Omega)$. It is in this sense that the $\int v_x^2 \, dA$ and $\int v_x^2 \, dA$ in Slide 18 must be interpreted: v is a member of $H^1(\Omega)$ if v has a first distributional derivative in $L^2(\Omega)$ (and, more generally, a member of $H^m(\Omega)$ if it has m distributional derivatives in $L^2(\Omega)$). In effect, this allows us to include many functions — including $\eta = 1 - 2|x - \frac{1}{2}|$ — for which the derivative is not defined in the classical sense at all points in Ω. These new functions we can include are precisely the ones of interest in the numerical context.

We can not go too far. If we take the second distributional derivative of our function η, we find by integration by parts that

$$\begin{aligned}
\langle D^2 L_\eta, v\rangle &= \int_0^1 \eta \, \frac{d^2 v}{dx^2}\, dx \\
&= -\int_0^{1/2} 2\, \frac{dv}{dx}\, dx + \int_{1/2}^1 2\, \frac{dv}{dx}\, dx \\
&= -2\, v(\tfrac{1}{2}) - 2v(\tfrac{1}{2}) \\
&= -4v(\tfrac{1}{2}) \\
&= -4\langle L_{\delta_{x_0=\frac{1}{2}}}, v\rangle \ ,
\end{aligned}$$

again as expected. But $L_{\delta_{x_0=\frac{1}{2}}}$ is certainly not square integrable, that is, in $L^2(\Omega)$ (see Note 7), so η is *not* in $H^2(\Omega)$ — it does not have $m = 2$ distributional derivatives in $L^2(\Omega)$. Similarly, the Heaviside-like function, $\widetilde{D^1\eta}$, is *not* in $H^1(\Omega)$.

2.3.6 Proper Spaces: Well-Posedness

SLIDE 20

Given $\ell \in H^{-1}(\Omega)$, find $u \in H_0^1(\Omega)$
such that

$$a(u,v) = \ell(v), \qquad \forall\, v \in H_0^1(\Omega)\ .$$

Well-posedness:

u *exists* and is *unique* ; E5 N9

$\|u\|_{H^1(\Omega)} \le C\, \|\ell\|_{H^{-1}(\Omega)}$ — *stability.*

N10 E6 E7

▷ ***Exercise 5*** Demonstrate that, assuming u exists, it is unique. (*Hint*: consider two solutions, u_1, and u_2, and show that they must be equal.) ■

Note 9 Lax-Milgram theorem: coercivity and continuity (Optional)

It follows from the Lax-Milgram theorem (related here to the Riesz representation theorem) that a problem of the form posed above has a unique solution if a is coercive and continuous. By *coercivity* we mean that there exists a positive α such that

$$\alpha\, \|v\|^2_{H^1(\Omega)} \le a(v,v), \qquad \forall\, v \in H_0^1(\Omega)\ ;$$

in our problem this property follows from the Poincaré-Friedriches inequality (see Note 6; indeed, since $a(v,v) = |v|^2_{H^1(\Omega)}$, $\alpha = C^2_{\rm PF}$). By *continuity* we mean that

$$a(w,v) \le \beta \, \|w\|_{H^1(\Omega)} \, \|v\|_{H^1(\Omega)} \qquad \forall \, w \in H^1(\Omega), \; \forall \, v \in H^1(\Omega) \; ;$$

this follows in our case (with $\beta = 1$) by the Cauchy-Schwarz inequality.

It is then a simple matter to prove stability: since $u \in H^1_0(\Omega)$, taking $v = u$ gives

$$\alpha \|u\|^2_{H^1(\Omega)} \le a(u,u) = \ell(u) \; ,$$

so

$$\alpha \|u\|_{H^1(\Omega)} \le \frac{\ell(u)}{\|u\|_{H^1(\Omega)}} \le \sup_{v \in H^1_0(\Omega)} \frac{\ell(v)}{\|v\|_{H^1(\Omega)}} = \|\ell\|_{H^{-1}(\Omega)} \; ,$$

and thus $C = \alpha^{-1}$.

Note 10 **Delta distribution data (Optional)**

We consider

$$-u_{xx} = f \qquad \text{in } \Omega = (0,1)$$

$$u(0) \; = \; u(1) = 0 \; .$$

The weak form is: find $u \in H^1_0(\Omega)$ such that

$$\int_0^1 u_x \, v_x \, dx = \int_0^1 f \, v \, dx, \qquad \forall \, v \in H^1_0(\Omega) \; ,$$

which can be written as

$$\int_0^1 u_x \, v_x = \langle f, v \rangle, \qquad \forall \, v \in H^1_0(\Omega) \; ,$$

where $\langle f, v \rangle \equiv f(v)$ $(= \int_0^1 f v dx$ if $f \in L^2(\Omega))$ is now the duality pairing between $(f \in)$ $H^{-1}(\Omega)$ and $(v \in)$ $H^1(\Omega)$. Note $\int_0^1 u_x \, v_x \, dx$ can be viewed as minus the second distributional derivative of u, integrated by parts once, and thus we see that the weak form is the original equation interpreted in the "distributional" sense.

Now consider a sequence of heat sources f_n that became increasingly concentrated at $x = \frac{1}{2}$ with $\int_0^1 f_n \, dx$ (the total heat input) $= 4$. Thus, in the limit that $n \to \infty$ and $f_n \to 4\delta(x - \frac{1}{2})$,

$$\text{“ } -u_{xx} = 4 \, \delta(x - \tfrac{1}{2})\text{”}$$

in the distributional sense. The solution to this problem is our function $\eta = 1 - 2|x - \frac{1}{2}|$. Thus $u = \eta$ does not have second derivatives in the usual sense, and

is not even in $H^2(\Omega)$ from our arguments of Note 8. The strong form has lost its meaning — for example, how would we apply finite differences to this equation? Certainly setting f to infinity at one grid point will produce nonsense.[1] We would need to break the problem up into two subdomains (Left and Right)

$$\begin{aligned} -u^{\mathrm{L}}_{xx} &= 0 && 0 < x < \tfrac{1}{2} \\ -u^{\mathrm{R}}_{xx} &= 0 && \tfrac{1}{2} < x < 1 \\ u^{\mathrm{L}}(0) &= u^{\mathrm{R}}(1) = 0 \\ u^{\mathrm{L}}(\tfrac{1}{2}) &= u^{\mathrm{R}}(\tfrac{1}{2}) \\ -u^{\mathrm{R}}_x(\tfrac{1}{2}) + u^{\mathrm{L}}_x(\tfrac{1}{2}) &= 4\,, \end{aligned}$$

and apply the special conditions at the interface. In short, the f_n in $L^2(\Omega)$ (and corresponding solutions $u_n \in H^2(\Omega)$) and their limit $f \in H^{-1}(\Omega)$ (and corresponding limit $u \in H^1_0(\Omega)$) require different treatment in the strong context. (This also tells us that $H^2(\Omega)$ is not complete in the H^1 norm; $H^1(\Omega)$ is not complete in the L^2 norm; $L^2(\Omega)$ is not complete in the H^{-1} norm, ...)

However, as we know, u *is* in $H^1_0(\Omega)$, and the delta distribution *is* in $H^{-1}(\Omega)$, and thus our weak formulation,

$$\int_0^1 v_x\, u_x\, dx = 4\, v(\tfrac{1}{2}), \qquad \forall\, v \in H^1_0(\Omega)\ ,$$

requires no modification to handle this — and many other — important limiting cases. In terms of J, it means that often the minimizer will not be in $H^2(\Omega)$, but only in $H^1(\Omega)$; had we required $H^2(\Omega)$ (second derivatives), there would be many "Strangian pinpricks" in our paraboloid. By filling those holes we correctly treat a much larger class of problems — and with considerably greater numerical ease, as we shall see. (Note $H^2(\Omega)$ is *dense* in $H^1(\Omega)$, that is, $H^1(\Omega)$ is the closure or completion of $H^2(\Omega)$ in the H^1 norm; any function (e.g., $u = \eta$) in $H^1(\Omega)$ is the limit in the H^1 norm of a sequence of functions (e.g., the u_n) in $H^2(\Omega)$. Thus only pinpricks — not large holes — need be filled as we expand our space from $H^2(\Omega)$ to $H^1(\Omega)$.)

▷ ***Exercise 6*** Show by explicit calculation that $u = \eta = 1 - 2|x - \frac{1}{2}|$ does indeed satisfy the weak form given in the last paragraph of the preceding Note. *Hint*: break the integral into two pieces, $\int_0^{\frac{1}{2}-\varepsilon}\, dx$ and $\int_{\frac{1}{2}+\varepsilon}^1\, dx$, and integrate by parts. ■

[1] Note a finite volume approach — based on a control-volume integral conservation statement — could deal with $f_n \to 4\delta(x - \frac{1}{2})$ gracefully. Finite volumes share some aspects in common with strong form finite differences, and some aspects in common with weak form finite elements.

▷ ***Exercise 7*** Consider the fourth-order problem

$$u_{xxxx} = f \qquad \text{in } \Omega = (0,1) \ ,$$

$$u(0) = u_x(0) = u(1) = u_x(1) = 0 \ ;$$

this "biharmonic" equation is relevant to, amongst other applications, the bending of beams.

(a) Find an SPD bilinear form a over X and a linear form ℓ such that

$$u = \arg\min_{w \in X} \ J(w) = \tfrac{1}{2}\, a(w,w) - \ell(w)$$

$$\Updownarrow$$

$$a(u,v) = \ell(v), \qquad \forall\, v \in X \ ,$$

where $w \in X$ are sufficiently smooth and satisfy $w(0) = w_x(0) = w(1) = w_x(1) = 0$. (*Hint*: work backwards, multiplying the strong form by v, and integrating by parts and applying the boundary conditions until symmetry "appears"; then verify your result — prove a is SPD, u is a minimizer, ... — *a posteriori*.)

(b) How should X be defined — which Hilbert space $H^m(\Omega)$ do you think is appropriate?

(c) Do you think that $\ell(v) = v_x(\frac{1}{2})$ is an admissible linear functional, in the sense that $\ell \in X'$, that is, $|\ell(v)| \le C\|v\|_X$, $\forall\, v \in X$?

■

3 The Neumann Problem

3.1 Strong Formulation

SLIDE 21

Find u such that

$$\boxed{\begin{array}{rcll} -\nabla^2 u & = & f & \text{in } \Omega \\ u & = & 0 & \text{on } \Gamma^D \\ \dfrac{\partial u}{\partial n} & = & g & \text{on } \Gamma^N \end{array}}$$

where $\overline{\Gamma} = \overline{\Gamma}^D \cup \overline{\Gamma}^N$, Γ^D non-empty. N11

Note 11 **Solvability for pure Neumann problem**

Note if Γ^D is empty, the Neumann problem may not have a solution. In particular, the equation tells us that

$$-\int_\Omega \nabla^2 u \, dA = \int_\Gamma -\nabla u \cdot \hat{n} \, dS \quad = \quad \int_\Omega f \, dA \ ,$$

while the boundary condition tells us that

$$\int_\Gamma \frac{\partial u}{\partial n} \, dS = \int_\Gamma \nabla u \cdot \hat{n} \, dS \quad = \quad \int_\Gamma g \, dS \ ;$$

clearly, $\int_\Omega f \, dA + \int_\Gamma g \, dS$ must vanish — the heat generated must balance the heat in through the boundaries.

If the *solvability* condition relating f and g is satisfied, our problem will have a solution — unfortunately, it will have an infinity of solutions that all differ by a constant (in some sense, this is the right nullspace condition associated with the left nullspace solvability condition). In particular, it is clear that if u is a solution, then so is u + any constant. To pin the solution down we must specify (say)

$$\int_\Omega u \, dA = 0$$

(or some other value, perhaps depending on an initial condition to the corresponding temporal problem). Note in practice (numerically) we can also ask that u at a particular point be specified, though strictly speaking this may be mathematically suspect for the continuous problem.

The problem in which Γ^D is non-empty (a so-called "mixed" problem) raises no such solvability issues; the flux over Γ^D adjusts itself to ensure global balance.

3.2 Minimization Principle

3.2.1 Statement

SLIDE 22

Find

$$\boxed{u = \arg \min_{w \in X} J(w)}$$

where

$$X = \{v \in H^1(\Omega) \mid v|_{\Gamma^D} = 0\}$$

$$J(w) = \frac{1}{2} \int_\Omega \nabla w \cdot \nabla w \, dA - \int_\Omega f \, w \, dA - \int_{\Gamma^N} g \, w \, dS \ .$$

Note we do not require that $v \in X$ satisfy $\frac{\partial v}{\partial n}|_{\Gamma^N} = 0$ (or g), the reasons for which will become clear shortly.

3.2.2 Proof

SLIDE 23

Let $w = u + v$.
Then

$$J(\overbrace{\underbrace{u}_{\in X} + \underbrace{v}_{\in X}}^{w \in X}) = \frac{1}{2} \int_\Omega \nabla(u+v) \cdot \nabla(u+v)\, dA$$
$$- \int_\Omega f(u+v)\, dA - \int_{\Gamma^N} g(u+v)\, dS\ .$$

SLIDE 24

$$J(u+v) =$$
$$\frac{1}{2} \int_\Omega \nabla u \cdot \nabla u\, dA - \int_\Omega f\, u\, dA - \int_{\Gamma^N} g\, u\, dS$$
$$+ \int_\Omega \nabla u \cdot \nabla v\, dA - \int_\Omega f\, v\, dA - \int_{\Gamma^N} g\, v\, dS$$
$$+ \frac{1}{2} \int_\Omega \nabla v \cdot \nabla v\, dA$$

SLIDE 25

$$\delta J_v(u) = \int_\Omega \nabla u \cdot \nabla u\, dA - \int_\Omega f\, v\, dA - \int_{\Gamma^N} g\, v\, dS$$
$$= \int_\Omega \nabla \cdot (v \nabla u)\, dA - \int_\Omega v \nabla^2 u\, dA - \int_\Omega f\, v\, dA - \int_{\Gamma^N} g\, v\, dS$$
$$= \int_{\Gamma^D} \not{v}^{0}\ \nabla u \cdot \hat{n}\, dS + \int_\Omega v \{\underbrace{-\nabla^2 u - f}_{0}\}\, dA$$
$$+ \textstyle\int_{\Gamma^N} v \{\underbrace{\nabla u \cdot \hat{n} - g}_{0}\}\, dS \quad = \quad 0\,, \qquad \forall\, v \in X$$

SLIDE 26

$$J(u+v) = J(u) + \underbrace{\frac{1}{2} \int_\Omega \nabla v \cdot \nabla v\, dA}_{> 0 \text{ unless } v = 0}\,,\ \forall\, v \in X$$

Note if Γ^D is empty, then v = constant also renders our last term zero, and thus u is not a unique minimizer — u + any constant will do just as well. We already observed this in Note 11.

$\Rightarrow$

$$\boxed{\begin{array}{c} J(w) \geq J(u)\,, \qquad \forall\, w \in X\,; \\ \Updownarrow \\ u \text{ is } \textit{the} \text{ minimizer of } J(w)\,. \end{array}}$$

E8

▷ ***Exercise 8*** Consider

$$-u_{xx} = 1$$

$$u(0) = 0, \qquad u_x(1) = 1\,.$$

Find the analytical solution to this problem, and show by explicit computation that $\delta J_v(u) = 0, \forall\, v \in X$. Recall $X = \{v \in H^1(\Omega) \mid v(0) = 0\}$. ■

3.3 Weak Formulation

3.3.1 Statement

SLIDE 27

Find $u \in X$ such that

$$\boxed{\delta J_v(u) = 0\,, \quad \forall\, v \in X}$$

$$\Updownarrow$$

$$\boxed{\int_\Omega \nabla u \cdot \nabla v \, dA = \int_\Omega f\, v \, dA + \int_{\Gamma^N} g\, v \, dS\,, \quad \forall\, v \in X\,;}$$

see Slide 25 for proof.

SLIDE 28

Let:

$$a(w, v) = \int_\Omega \nabla w \cdot \nabla v \, dA\,, \qquad \forall\, w, v \in X$$

bilinear, SPD form ;

and

$$\ell(v) = \int_\Omega f\, v \, dA + \int_{\Gamma^N} g\, v \, dS$$

linear, bounded form (in $H^{-1}(\Omega)$) .

In order to ensure that $\ell \in H^{-1}(\Omega)$ we must require that g is sufficiently smooth on Γ^N. In order to avoid fractional-derivative norms we say that $g \in L^2(\Gamma^N)$; in fact, slightly less regularity is required. Note also that a is SPD only if Γ^D is non-empty.

SLIDE 29

Minimization Principle:

$$u = \arg \min_{w \in X} \underbrace{\frac{1}{2}\, a(w,w) - \ell(w)}_{J(w)} \ .$$

Weak Statement: $u \in X$,

$$\underbrace{a(u,v) = \ell(v)}_{\Leftrightarrow\ \delta J_v(u)=0}, \qquad \forall\, v \in X \ .$$

3.3.2 Essential vs. Natural

SLIDE 30

Essential boundary conditions: Imposed by X.
Natural boundary conditions: Imposed by J (or a, ℓ).
Here:

Essential $\Leftrightarrow$ Dirichlet ($v|_{\Gamma^D} = 0$) ,
Natural $\Leftrightarrow$ Neumann ($v|_{\Gamma^N}$ unrestricted) . N12

Important theoretical and numerical ramifications.

E9 E10 E11

Note 12 ***Natural boundary conditions***

Note that since ∇u is only in $L^2(\Omega)$, we can *not* really impose $\nabla u \cdot \hat{n} = g$ in a strong way — the *trace* (boundary limit) of a function in $L^2(\Omega)$ makes little sense, since jumps are permitted and individual points are "ignored." It thus must be the case that these boundary conditions are imposed in some other fashion. In fact, this other fashion — though a and ℓ — is much more convenient, since we do not need to compute any normals. (It is also very instructive to go "the other way" — show how certain test functions v concentrated on the boundary weakly "impose" the natural conditions.) This is another example (see our delta distribution of Note 10, and Exercise 9 below) in which the weak formulation *greatly facilitates* subsequent numerical treatment.

Note that essential is not always Dirichlet, and natural is not always Neumann. There are mixed formulations (related to complementary energy) in which we approximate directly ∇u in $H^1(\Omega)$ and u in $L^2(\Omega)$; as might be expected from the above arguments, in this case we can impose ∇u strongly — Neumann is essential — but we can not impose u (only in $L^2(\Omega)$) strongly — Dirichlet is natural. We shall not consider these dual formulations further in this (rather short) series of lectures on the finite element method.

▷ ***Exercise 9*** Consider a problem with a discontinuous jump in conductivities

$$\begin{aligned} -\kappa^{\mathrm{L}} u^{\mathrm{L}}_{xx} &= f^{\mathrm{L}} \qquad 0 < x < \tfrac{1}{2} \ , \\ -\kappa^{\mathrm{R}} u^{\mathrm{R}}_{xx} &= f^{\mathrm{R}} \qquad \tfrac{1}{2} < x < 1 \ , \end{aligned}$$

with boundary conditions

$$u^{\mathrm{L}}(0) = 0, \qquad u^{\mathrm{R}}(1) = 0 \ ,$$

$$u^{\mathrm{L}}(\tfrac{1}{2}) = u^{\mathrm{R}}(\tfrac{1}{2}) \ ,$$

$$-\kappa^{\mathrm{L}} u^{\mathrm{L}}_x(\tfrac{1}{2}) = -\kappa^{\mathrm{R}} u^{\mathrm{R}}_x(\tfrac{1}{2}) \qquad \text{(continuity of flux)} \ ;$$

here κ^{L} and κ^{R} are strictly positive.

(a) For $X = \{v \in H^1((0,1)) \mid v(0) = 0, \ v(1) = 0\}$, show that

$$u = \arg\min_{w \in X} \ \tfrac{1}{2}\, a(w,w) - \ell(w) \ ,$$

and

$$a(u,v) = \ell(v), \qquad \forall\, v \in X \ ,$$

where

$$\begin{aligned} a(w,v) &= \int_0^{1/2} \kappa^{\mathrm{L}}\, w_x\, v_x \, dx + \int_{1/2}^1 \kappa^{\mathrm{R}}\, w_x\, v_x \, dx \ , \\ \ell(w) &= \int_0^{1/2} f^{\mathrm{L}}\, v \, dx + \int_{1/2}^1 f^{\mathrm{R}}\, v \, dx \ . \end{aligned}$$

(b) In this problem, which boundary/interface conditions are essential, and which are natural?

(c) Is the solution to this problem in $H^2(\Omega)$? in $H^1(\Omega)$?

■

▷ ***Exercise 10*** Consider the Robin problem (the third standard boundary condition for the Poisson problem, and more generally second order elliptic PDEs)

$$\begin{aligned} -\nabla^2 u &= f \qquad \text{in } \Omega \\ u &= 0 \qquad \text{on } \Gamma^D \\ -\frac{\partial u}{\partial n} &= h_c u \qquad \text{on } \Gamma^R \quad (\overline{\Gamma} = \overline{\Gamma^D} \cup \overline{\Gamma^R}) \end{aligned}$$

where $h_c > 0$ (recall that $\frac{\partial u}{\partial n}$ refers to the outward normal on Γ).

(a) Find the functional J (and hence a and ℓ) such that

$$u = \arg\min_{w \in X} \; J(w) = \tfrac{1}{2}\, a(w,w) - \ell(w) \;,$$

and

$$a(u,v) = \ell(v), \qquad \forall\, v \in X \;,$$

where $X = \{v \in H^1(\Omega) \mid v|_{\Gamma^D} = 0\}$. *Hint*: multiply the equation by v, integrate by parts, and substitute $-h_c u$ for $\frac{\partial u}{\partial n}$ on the boundary; identify a and ℓ; verify your results *a posteriori* — prove a is SPD, u is a minimizer, ...

(b) In this problem, which boundary conditions are essential, and which are natural?

■

▷ ***Exercise 11*** Consider the fourth-order problem

$$u_{xxxx} = f \qquad \text{in } \Omega = (0,1) \;,$$
$$u(0) = u_{xx}(0) = u(1) = u_{xx}(1) = 0 \;.$$

(a) Show that the minimization statement of Exercise 7 still applies, but that now members v of X need only satisfy $v(0) = v(1) = 0$ (not $v_x(0) = v_x(1) = 0$ as before, or $v_{xx}(0) = v_{xx}(1) = 0$).

(b) Which boundary conditions are essential, and which are natural?

■

4 Inhomogeneous Dirichlet Conditions

4.1 Strong Formulation

SLIDE 31

Find u such that

$$\boxed{\begin{array}{rcll} -\nabla^2 u & = & f & \text{in } \Omega \\ u & = & u^D & \text{on } \Gamma^D = \Gamma \;; \end{array}}$$

simple extension to mixed Neumann or Robin.

The boundary data u^D must satisfy certain regularity conditions on Γ^D. In fact u^D must be a bit more than $L^2(\Gamma^D)$, but need not be quite as much as $H^1(\Gamma^D)$; discontinuites should be avoided.

4.2 Minimization Statement

SLIDE 32

Find $$\boxed{u = \arg \min_{w \in X^D} J(w)}$$

where
$$X^D = \{v \in H^1(\Omega) \mid v|_{\Gamma^D} = u^D\} ,$$
$$X = \{v \in H^1(\Omega) \mid v|_{\Gamma^D} = 0\} ,$$

Note that X^D is not a linear space. But the difference of any two members in X^D is a member of X, which is of course a space.

$$J(w) = \frac{1}{2} \underbrace{\int_\Omega \nabla w \cdot \nabla w \, dA}_{a(w,w)} - \underbrace{\int_\Omega f \, w \, dA}_{\ell(w)} .$$

4.3 Weak Formulation

SLIDE 33

Find $u \in X^D$ such that E12

$$\boxed{\delta J_v(u) = 0 , \qquad \forall \, v \in X \equiv H_0^1(\Omega)}$$

$$\Updownarrow$$

$$\boxed{\underbrace{\int_\Omega \nabla u \cdot \nabla v \, dA}_{a(u,v)} = \underbrace{\int_\Omega f \, v \, dA}_{\ell(v)} , \qquad \forall \, v \in X .}$$

▷ ***Exercise 12*** Prove the minimization statement and weak statement for the inhomogeneous Dirichlet case. *Hint*: proceed as for the homogeneous case, but note that any $w \in X^D$ can be expressed as $w = u \, (\in X^D) + v \, (\in X)$; v is *still* in X, that is, vanishes on Γ^D. ■

5 Summary

SLIDE 34

- The Poisson problem has a strong formulation;
  a minimization formulation; and a weak formulation.

- The minimization/weak formulations are more general than the strong formulation in terms of regularity and admissible data.

SLIDE 35

- The minimization/weak formulations are defined by: a space X; a bilinear form a; a linear form ℓ.

- The minimization and weak formulations identify

ESSENTIAL boundary conditions,
Dirichlet — reflected in X;

NATURAL boundary conditions,
Neumann — reflected in a, ℓ.

SLIDE 36

- The *points of departure* for the *finite element method* are:

the weak formulation (more generally);
or
the minimization statement (if a is SPD).

References: In addition to the references given in the course syllabus, in particular Strang & Fix and Quarteroni & Valli, the book Linear Operator Theory in Engineering and Science*, by A.W. Naylor and G.R. Sell, Springer-Verlag, 1982, is a very useful introduction to a number of basic concepts (e.g., linear spaces) covered in this lecture.*