Solving PDE's with FEniCS

One-dimensional and nonlinear problems

Chapters 8–9

Introduction to

Automated Modeling

with FEniCS

by L. Ridgway Scott

Exact solution

Consider the two-point boundary value problem

$$\begin{aligned} -\frac{d^2u}{dx^2} &= f \text{ in } (0,1) \\ u(0) = g_0, \quad u'(1) &= g_1. \end{aligned} \tag{1}$$

The solution can be determined from f via two integrations.

First of all, we can write

$$\frac{du}{dx}(t) = \int_t^1 f(s)\,ds + g_1$$

using the boundary condition at $x = 1$.

Exact solution

Integrating again shows that

$$u(x) = \int_0^x \int_t^1 f(s)\, ds\, dt + g_1 x + g_0$$

using the boundary condition at $x = 0$.

In particular, this shows that (1) is well posed.

Not so easy to check well-posedness in general, but

every investigation should (in principle)

begin with this step.

Integral representations often used to do this.

Suppose that u is the solution of (1) with $g_0 = 0$. Let v be any (sufficiently regular) function such that $v(0) = 0$. Then integration by parts yields

$$(f, v)_{L^2([0,1])} = \int_0^1 f(x)v(x)dx = \int_0^1 -u''(x)v(x)dx$$
$$= \int_0^1 u'(x)v'(x)dx - g_1 v(1).$$

Define

$$a(u, v) := \int_0^1 u'(x)v'(x)dx \tag{2}$$

and

$$V = \left\{ v \in L^2([0,1]) \ : \ a(v,v) < \infty \text{ and } v(0) = 0 \right\}. \tag{3}$$

Then we can say that the solution u to (1) is characterized by

$$\begin{gathered} u \in V \quad \text{such that} \\ a(u, v) = (f, v)_{L^2([0,1])} + g_1 v(1) \quad \forall v \in V, \end{gathered} \tag{4}$$

which is called the **variational formulation** or **weak formulation** of (1).

The relationship (4) is called “variational” because the function v is allowed to vary arbitrarily.

It has a natural interpretation in the setting of Hilbert spaces [1].

Variational Formulation of Boundary Conditions

The Dirichlet boundary condition $u(0) = 0$ is called an **essential boundary condition** because it appears in the variational space.

The Neumann boundary condition $u'(1) = 0$ is called an **natural boundary condition** because it does not appear in the variational space but rather is implied in the formulation.

Inhomogeneous Dirichlet boundary conditions are handled as follows in the variational formulation.

Let u_0 be some function satisfying the inhomogeneous Dirichlet boundary conditions (but not necessarily the Neumann boundary conditions).

Then

$$u - u_0 \in V \quad \text{such that}$$
$$a(u, v) = (f, v)_{L^2([0,1])} + g_1 v(1) \quad \forall v \in V. \tag{5}$$

Equivalently, this can be written as $u = w + u_0$ where $w \in V$ such that

$$a(w, v) = (f, v)_{L^2([0,1])} + g_1 v(1) - a(u_0, v) \qquad \forall v \in V. \tag{6}$$

Note that the general problem (5) can be written

$$w \in V \quad \text{such that} \quad a(w, v) = F(v) \qquad \forall v \in V \tag{7}$$

where F denotes a **linear functional** on the space V.

Linear functionals

A function F defined on V is called a linear functional if

it is a linear function defined for any $v \in V$ having

a real number as its value.

The right-hand side of (6) can be written succintly as

$$F(v) = (f, v)_{L^2([0,1])} + g_1 v(1) - a(u_0, v) \quad \forall v \in V.$$

The expression F is called a linear functional because (a) it is linear and (b) it has scalar values.

By linear, we mean that

$$F(u + av) = F(u) + aF(v)$$

for any scalar a and any $u, v \in V$.

Continuous linear functionals

Critical condition on a linear functional for success in a variational formulation: *bounded* or *continuous*.

A linear functional F is bounded (equivalently continuous) on a normed space V if

$$|F(v)| \leq C_F \|v\|_V \quad \forall v \in V. \tag{8}$$

A natural norm $\|\cdot\|_V$ for the space V defined in (3) is

$$\|v\|_a = \sqrt{a(v,v)}.$$

The smallest possible constant C_F for which

(8) holds is called the dual norm of F.

Dual norm

The **dual norm** of F is defined by

$$\|F\|_{V'} := \sup_{0 \neq v \in V} \frac{|F(v)|}{\|v\|_V}.$$

The main point is that all the linear forms considered so far *are* bounded

In particular the Dirac δ-function, defined by $\delta(v) = v(1)$.

But also easy to think of others which are not, such as

$$F(v) := v'(x_0)$$

for some $x_0 \in [0, 1]$. This form is linear, but consider what it should do for the function $v(x) := |x - x_0|^{2/3}$.

Equivalence of formulations

The general variational formulation (7) can be shown to be completely equivalent to the orignal differential equation (see Theorem 0.1.4 of [1]).

A key tool is the following: suppose that f is continuous and

$$\int_0^1 f(x)v(x)\,dx = 0 \quad \text{for all } v \in V.$$

Then $f \equiv 0$.

Moreover, the variational formulation provides a framework that allows less regular data (arbitrary continuous linear functionals for F) as required by important physical applications.

One detail

The expression $a(\cdot, \cdot)$ is called a **bilinear functional** on the space V, since it is a bilinear function defined on the Cartesian product $V \times V$ having a single real number as its value.

If we fix one of the variables of a bilinear form, it yields a linear form in the remaining variable.

Recall that we have proved the Cauchy-Schwarz inequality

$$|a(u, v)| \leq \|u\|_a \|v\|_a. \tag{9}$$

where $\|v\|_a = \sqrt{a(v, v)}$.

Another detail

Consider the linear form $F(v) = v(x_0)$ for some $x_0 \in [0, 1]$.

We want to prove that this is bounded on V.

We write a function as the integral of its derivative and begin to estimate:

$$v(t) = \int_0^t v'(x)\,dx = \int_0^1 v'(x)w'(x)\,dx = a(v, w), \quad (10)$$

where the function $w \in V$ is defined by

$$w(x) = \begin{cases} x & 0 \le x \le t \\ t & x \ge t \end{cases}$$

Side benefits

One benefit of our loosened notion of derivative is that such functions are indeed in V, even though the derivative of w is discontinuous.

By the Cauchy-Schwarz inequality (9), for all $t \in [0, 1]$,

$$|v(t)| = |a(v, w)| \leq \|v\|_a \|w\|_a = \sqrt{t}\|v\|_a \leq \|v\|_a. \quad (11)$$

Inequality (11) is called Sobolev's inequality, and V is an example of a Sobolev space.

Note that the first step in (10) uses the fact that for $v \in V$, $v(0) = 0$.

This subtle point is nevertheless essential, since (11) is clearly false if this boundary condition is not available.

Sobolev's inequality

In particular, if v is a constant function, then the right-hand-side of (11) is zero for this v whereas the left-hand-side is not (unless $v \equiv 0$).

Sobolev's inequality holds in a more general setting, not requiring boundary conditions, but only when the bilinear form is an inner-product.

What Sobolev's inequality inequality tells us is that,

- even though the functions in V are not smooth in the classical sense
- (derivatives can even be infinite at isolated points),
- they nevertheless have some type of classical regularity, namely continuity in this case.

Natural boundary conditions

We saw that the 'natural' boundary condition, e.g., $u'(1) = 0$ in (1) when $g_1 = 0$, disappears in the variational formulation (4).

But if these are in some sense equivalent formulations (they are), then the natural boundary condition must be encoded in the variational formulation is some way.

We can see this by reversing the process used to go from (1) to (4).

So suppose that u satisfies (4), and also assume that it is smooth enough for us to integrate by parts:

Equivalence proof

$$
\begin{aligned}
(f, v)_{L^2([0,1])} &= \int_0^1 u'(x)v'(x)\,dx \\
&= \int_0^1 -u''(x)v(x)\,dx + \big(u'v\big)\big|_0^1 \qquad (12) \\
&= \int_0^1 -u''(x)v(x)\,dx + u'(1)v(1).
\end{aligned}
$$

Choosing first $v \in V$ that vanishes at $x = 1$, we conclude that

$$\int_0^1 \big(f + u''(x)\big)v(x)\,dx = 0$$

for all such v.

Equivalence proof

From this, one can show that we necessarily have $-u'' = f$.

E.g., suppose $\big(f + u''\big)(x) \neq 0$ at some x. Then choose v supported near x to be of one sign, yielding a contradition.

Inserting this fact in (12), we conclude that $u'(1) = 0$ simply by taking a single v such that $v(1) \neq 0$, e.g., $v(x) = x$.

Thus the natural boundary condition emerges from the variational formulation "naturally."

And as an intermediate step, we see that u satisfies the first equation in (1), proving equivalence of (1) and (4).

Galerkin Approximation

Let $V_h \subset V$ be any (finite dimensional) subspace. Let us consider (4) with V replaced by V_h, namely

$$u_h \in V_h \text{ such that } a(u_h, v) = (f, v)_{L^2([0,1])} \quad \forall v \in V_h. \quad (13)$$

Then (13) represents a square, finite system of equations for u_h which can easily be seen to be invertible [1].

Note how easily a discrete scheme for approximating (1) can be defined.

A matrix equation is derived by writing (13) in terms of a basis $\{\phi_i \ : \ 1 \leq i \leq n\}$ of V_h.

Write u_h in terms of this basis, i.e.,

$$u_h = \sum_{j=1}^{n} U_j \phi_j$$

where the coefficients U_j are to be determined. Define

$$A_{ij} = a(\phi_j, \phi_i), \quad F_i = (f, \phi_i) \text{ for } i, j = 1, ..., n.$$

Set $\mathbf{U} = (U_j)$, $\mathbf{A} = (A_{ij})$ and $\mathbf{F} = (F_i)$.

Then (13) is equivalent to solving the (square) matrix equation

$$\mathbf{AU} = \mathbf{F}. \tag{14}$$

Galerkin matrices

The matrix $\mathbf{A}$ is often referred to as the **stiffness matrix**, a name coming from corresponding matrices in the context of structural problems.

Another important matrix is the **mass matrix**, namely

$$M_{ij} = (\phi_j, \phi_i)_{L^2([0,1])} \quad \text{for} \quad i, j = 1, ..., n.$$

If $f \in V$ with $f = \sum \tilde{F}_j \phi_j$ then (13) is equivalent to solving the matrix equation

$$\mathbf{AU} = \mathbf{M}\tilde{\mathbf{F}}.$$

We now construct the finite element space V_h, including the mesh.

Let $0 = x_0 < x_1 < ... < x_n = 1$ be a partition of $[0, 1]$, and

let V_h be the linear space of functions v such that

- v is continuous everywhere
- $v|_{[x_{i-1},x_i]}$ is a linear polynomial, $i = 1, ..., n$, and
- $v(0) = 0$.

The function space just defined can be described as the set of **continuous piecewise linear** functions with respect to the mesh (x_i).

Piecewise linear basis

For each $i = 1, .., n$ define ϕ_i by the requirement that

$$\phi_i(x_j) = \delta_{ij}$$

the Kronecker delta. Then

- $\{\phi_i \ : \ 1 \leq i \leq n\}$ is called a **nodal basis** for V_h, and
- $\{v(x_i)\}$ are the **nodal values** of a function v.

The points $\{x_i\}$ are called the **nodes**.

A function space consisting of **continuous piecewise quadratic** functions, with respect to the mesh (x_i), can be defined similarly.

Let V_h be the linear space of functions v such that

- v is continuous everywhere
- $v|_{[x_{i-1},x_i]}$ is a quadratic polynomial, $i = 1, ..., n$, and
- $v(0) = 0$.

However, now there are additional nodes in the middle of each *element* $[x_{i-1}, x_i]$, i.e., at $(x_i + x_{i-1})/2$.

Now the nodal numbering gets a bit complicated.

Let $y_{2i} = x_i$ and let $y_{2i-1} = (x_i - x_{i-1})/2$ for $i = 1, \ldots, n$.

Then the nodal basis is defined by $\phi_i(y_j) = \delta_{ij}$ for $i, j = 1, \ldots, 2n$

Galerkin matrix entries

The Galerkin method using piecewise polynomials spaces described in terms of nodal values is called the finite-element method.
The stiffness matrix $\mathbf{A}$ as defined in (14), using the Kronecker basis $\{\phi_i\}$, can be interpreted as a difference operator.

Let $h_i = x_i - x_{i-1}$.

Then the matrix entries $A_{ij} = a(\phi_i, \phi_j)$ can be easily calculated to be $A_{nn} = h_n^{-1}$, for $i = 1, ..., n-1$

$$A_{ii} = h_i^{-1} + h_{i+1}^{-1}, \quad A_{i,i+1} = A_{i+1,i} = -h_{i+1}^{-1}, \tag{15}$$

with the rest of the entries of $\mathbf{A}$ being zero.

Similarly, the entries of $\mathbf{F}$ can be approximated if f is sufficiently smooth:

$$(f, \phi_i) = \frac{1}{2}(h_i + h_{i+1})(f(x_i) + \mathcal{O}(h))$$

where $h = \max h_i$.

Thus, the $i - th$ equation of $\mathbf{AU} = \mathbf{F}$ (for $1 \leq i \leq n-1$) can be written as

$$\frac{-2}{h_i + h_{i+1}} \left[\frac{U_{i+1} - U_i}{h_{i+1}} - \frac{U_i - U_{i-1}}{h_i} \right] = \frac{2(f, \phi_i)}{h_i + h_{i+1}} \qquad (16)$$

$$= f(x_i) + \mathcal{O}(h).$$

Relationship to Difference Methods

The difference operator on the left side of this equation can be seen to be an $\mathcal{O}(h)$ accurate approximation to the differential operator $-d^2/dx^2$.

But showing that it is $\mathcal{O}(h^2)$ accurate is not easy [3].

For a uniform mesh, the equations reduce to the familiar difference equations

$$-\frac{U_{i+1} - 2U_i + U_{i-1}}{h^2} = f(x_i) + \mathcal{O}(h^2).$$

Thus the finite difference and finite element discretization techniques can be seen to produce essentially the same set of equations in many cases.

Relationship to Difference Methods

Even though the difference method (16) is formally only first order accurate, the variational framework [1] shows that the resulting error is second order accurate:

$$e_h := \max_{1 \le i \le 2n} |u(y_n) - u_n| \le C_f h^2$$

Useful to view a difference method as a variational method (if possible) for the purposes of analysis.

Because the variational form $a(\cdot, \cdot)$ is symmetric, the Galerkin method will always yield a symmetric matrix.

In applying boundary conditions with finite difference methods, care must be exercised to retain the symmetry of the original differential equation.

Relationship to Difference Methods

The system of equations obtained for the nodal variables (u_n) in the case of the Galerkin method using continuous piecewise quadratics does not look like a conventional finite difference method.

The equations associated with the internal nodes are different from the ones associated with the subdivision points.

On the other hand, they yield a more accurate method, satisfying

$$e_h := \max_{1 \le i \le 2n} |u(y_n) - u_n| \le C_f h^3.$$

Coercivity of the Variational Problem

The variational form $a(\cdot, \cdot)$ introduced in (2) is *coercive* on the corresponding spaces V (see [1]):

there is a constant γ depending only on Ω and Γ such that

$$\|v\|^2_{H^1(\Omega)} \leq \gamma a(v, v) \quad \forall v \in V. \tag{17}$$

The proof of this is elementary.

All we need to show is that

$$\|v\|^2_{L^2(\Omega)} \leq C a(v, v) \quad \forall v \in V, \tag{18}$$

from which (17) follows with constant $\gamma = C + 1$.

Proof of coercivity

To prove (18), we apply Sobolev's inequality (11), which says

$$|v(t)| \leq \sqrt{t}\|v\|_a.$$

Thus

$$\int_0^1 v(t)^2\,dt \leq a(v,v)\int_0^1 t\,dt \leq \tfrac{1}{2}a(v,v)$$

which completes the proof of (18), with $C = 1/2$.

Note that our proof of Sobolev's inequality (9) uses the fact that for $v \in V$, $v(0) = 0$.

If v is a constant function, then the right-hand-side of (17) is zero for this v whereas the left-hand-side is not (unless $v \equiv 0$).

Application of coercivity

From (17), it follows that the problem (7) is well posed.

In particular, we easily see that the solution to the problem must be unique, for if F is identically zero then so is the solution.

In the finite-dimensional case, this uniqueness also implies existence, and a similar result holds in the setting of infinite dimensional Hilbert spaces such as V. Moreover, the coercivity condition immediately implies a stability result, namely

$$\|u\|_{H^1(\Omega)} \leq \frac{\gamma a(u,u)}{\|u\|_{H^1(\Omega)}} = \gamma \frac{F(u)}{\|u\|_{H^1(\Omega)}} \leq \gamma \|F\|_{H^{-1}(\Omega)}.$$

Lipschitz continuity of solution operator

Here we are using the notation $\|F\|_{H^{-1}(\Omega)}$ for the dual norm of F in the dual space of $H^1(\Omega)$, i.e., $H^{-1}(\Omega) := \left(H^1(\Omega)\right)'$ [1].

The same result holds for a discrete approximation as well.

As a byproduct, coercivity proves continuity of the solution as a function of the data since the problem is linear.

In particular, if F_i, $i = 1, 2$, are two bounded linear forms, and u_i denotes the corresponding solutions to (7), then

$$\|u_1 - u_2\|_{H^1(\Omega)} \leq \gamma \|F_1 - F_2\|_{H^{-1}(\Omega)}.$$

Consider the two-point boundary value problem

$$-\frac{d^2u}{dx^2} + \alpha(x)\frac{du}{dx} + \beta(x)u = f \text{ in } (0,1)$$
$$u(0) = g_0, \quad u'(1) = g_1.$$

Then integration by parts can again be used to derive the variational formulation

$$a(u,v) = (f,v)_{L^2([0,1])} \quad \forall v \in V$$

where

$$a(u,v) := \int_0^1 u'(x)v'(x) + \alpha(x)u'(x)v(x) + \beta(x)u(x)v(x)\,dx. \tag{19}$$

Variational problem requires use of quadrature to integrate the expressions involving α and β.

The question of coercivity of the form (19) can be addressed in at least simple cases.

If $\beta \equiv 0$ and α is constant, then

$$\begin{aligned} a(v,v) &= \int_0^1 v'(x)^2 + \tfrac{1}{2}\alpha(v^2)'(x)\,dx \\ &= \int_0^1 v'(x)^2\,dx + \tfrac{1}{2}\alpha v(1)^2 \quad \forall v \in V. \end{aligned}$$

If $\alpha > 0$, then this is coercive.

Exercise: determine conditions needed on β to retain coercivity.

Spectral Elements – P Method

Use of high-order piecewise polynomials in Galerkin approximations goes by various names.

- Degree P used as approximation parameter
- Convergence achieved by letting P increase

Often called the ‘P’ method.

Also goes by the name “spectral element” method because of similarities with spectral methods

with possibility of subdividing domain using “elements.”

Trigonometric Polynomials – Spectral Methods

Trigonometric polynomials in Galerkin approximation useful for square domains

Discretizations popularly known as Spectral Methods.

Fast Fourier transforms used to solve discrete equations

Lageurre funtions

Useful for infinite domains

and the list goes on

Nonlinear example

Consider the following problem:

$$-\frac{d^2u}{d\xi^2}(\xi) + 4u(\xi) + 6u(\xi)^2 = f(\xi) \text{ for } \xi \in (0,\theta) \qquad (20)$$
$$u(0) = 0, \quad u'(\theta) = 0.$$

With $f = C$ (= constant), this describes the profile of radial component of fluid flow in converging channel (a.k.a. Jeffrey-Hamel flow).

In (20), differentiation is with respect to the polar angle ξ and θ is half of the angle (in radians) of convergence of the channel.

Given a solution u of (20), one can show that

$$\mathbf{u}(x,y) := \nu \, \frac{u(\operatorname{atan}(y/x))}{x^2+y^2} \, \mathbf{x}\,, \quad \mathbf{x} = (x,y) \in \Omega$$

solves the steady Navier-Stokes equations with kinematic viscosity ν over a wedge domain Ω (cf. [2]).

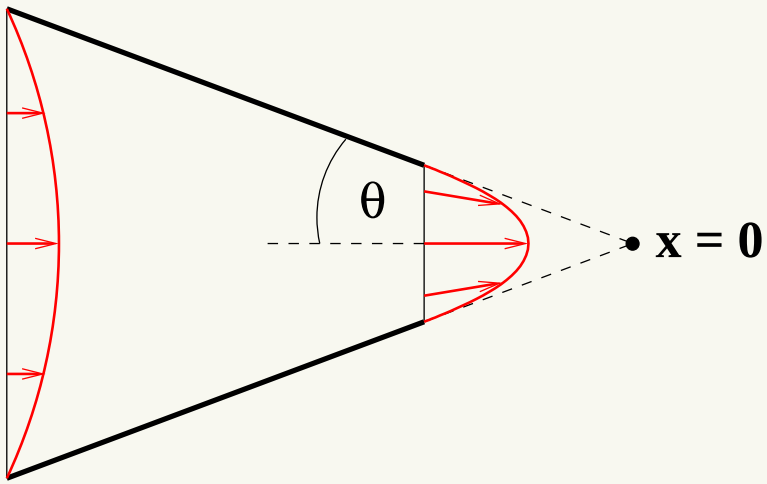


Figure 1: Interpretation of solutions of Jeffrey-Hamel equation.

Jeffrey-Hamel flow

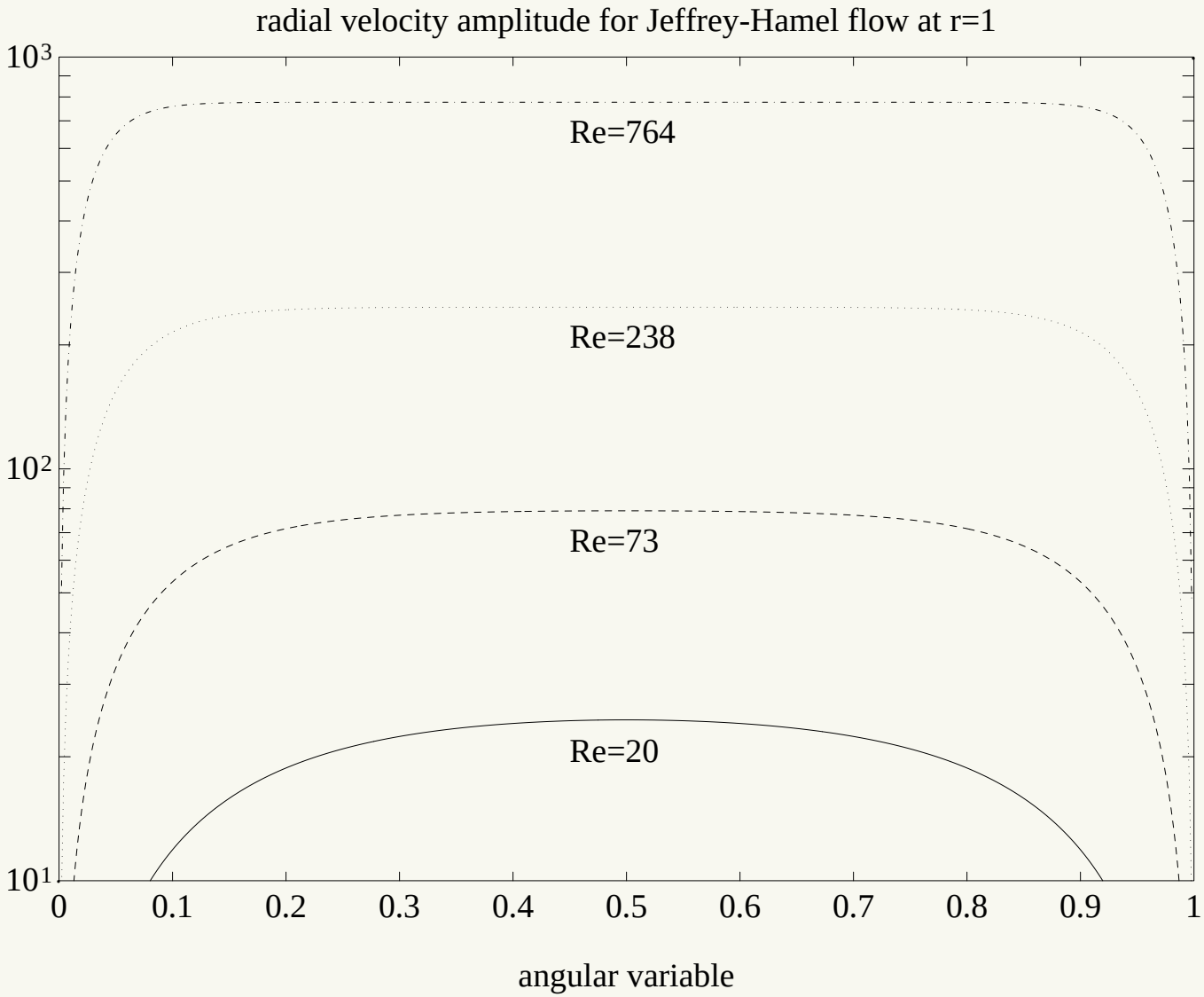


Figure 2: Solutions of Jeffrey-Hamel equation with $C = 10^k$ for $k = 1, 2, 3, 4$.

Reality checks: large C

With nonlinear problems, difficult to derive synthetic solutions using method of manufactured solution.

But nonlinear problems allow useful asymptotic estimates.

In the Jeffrey-Hamel problem (20), if solution u becomes large, then quadratic term u^2 will be even larger.

Dropping smaller terms in $-u'' + 4u + 6u^2 = C$ we get

$$6u^2 \approx C,$$

so that $u \approx \sqrt{C/6}$, with boundary layer near $\xi = 0$.

Reality checks: small C

When u is small, u^2 is smaller than u, so can drop quadratic term u^2 in (20) to get $-u'' + 4u = C$.

Let v be the solution of

$$\begin{aligned} -\frac{d^2v}{d\xi^2} + 4v &= 1 \text{ in } (0,\theta) \\ v(0) = 0, \quad v'(\theta) &= 0. \end{aligned} \tag{21}$$

Then for small C we expect that

$$u \approx Cv.$$

Exact solution

In this case, we can go one step further since it is possible to solve (21) exactly:

$$v(\xi) = \frac{1}{4} - \frac{e^{-2(\theta-\xi)} + e^{2(\theta-\xi)}}{4\left(e^{-2\theta} + e^{2\theta}\right)}$$

Therefore we can compute $\|u - Cv\|_{L^2(\Omega)}$ to see if it goes to zero as $C \to 0$.

Thus there are useful tests that can be done for nonlinear problems that can determine whether or not the code is producing reasonable answers.

Nonlinear problems such as (20) can also be formulated variationally, as

$$a(u,v) + n(u,v) = (f,v)_{L^2([0,1])} \quad \forall v \in V$$

where $a(\cdot,\cdot)$ is as in (19) with $\alpha \equiv 0$ and $\beta(x) \equiv 4$.

The nonlinearity has been separated for convenience in the form

$$n(u,v) = 6\int_0^1 u(x)^2 v(x)\,dx = 6\,(u^2,v)_{L^2([0,1])}.$$

Define a nonlinear form F by

$$F(u,v) = a(u,v) + n(u,v) - (f,v)_{L^2([0,1])} \quad \forall v \in V \quad (22)$$

Nonlinear problems in matrix form

A Galerkin method for a space with a basis $\{\phi_i \ : \ i = 1, \ldots, n\}$ can be written as a system of nonlinear equations

$$F_i(u) := a(u, \phi_i) + n(u, \phi_i) - (f, \phi_i)_{L^2([0,1])} = 0$$

Writing $u = \sum_j U_j \phi_j$, Newton's method for this system of equations for (U_j) can be derived.

However, can also be cast in variational form as follows.

Instead of using a basis function, define a functional $\mathbf{F}$ with coordinates parametrized by an arbitrary $v \in V$:

$$F_v(u) := a(u, v) + n(u, v) - (f, v)_{L^2([0,1])}$$

If $v = \phi_i$ then of course we have the previous function.

Newton's method requires computing derivatives of F with respect to "coordinates" (elements of V).

Derivative of F_v at u in direction of $w \in V$ is, as always, a limit of a difference quotient,

$$\frac{F_v(u + \epsilon w) - F_v(u)}{\epsilon},$$

as $\epsilon \to 0$. Expanding, we find that

$$\begin{aligned} F_v(u + \epsilon w) - F_v(u) &= \epsilon a(w, v) + 6\left((u + \epsilon w)^2 - u^2, v\right)_{L^2([0,1])} \\ &= \epsilon a(w, v) + 6\left(2\epsilon uw + \epsilon^2 w^2, v\right)_{L^2([0,1])}. \end{aligned}$$

Newton variational formulation

$$F_v(u+\epsilon w) - F_v(u) = \epsilon\Big(a(w,v) + 6\left(2uw + \epsilon\, w^2, v\right)_{L^2([0,1])}\Big)$$

Therefore

$$\lim_{\epsilon\to 0} \frac{F_v(u+\epsilon w) - F_v(u)}{\epsilon} = a(w,v) + 12\left(uw, v\right)_{L^2([0,1])}$$

for any $w \in V$.

Then Newton's method characterized by

$$u \leftarrow u - w \quad \text{where } w \text{ solves}$$

$$\begin{aligned} a(w,v) + 12\left(uw,v\right)_{L^2([0,1])} &= a(u,v) + n(u,v) - (f,v)_{L^2([0,1])} \\ &= F_v(u) \\ &= F(u,v) \quad \forall v \in V \end{aligned}$$

Fortunately, `dolfin` automates all of this.

With F as defined in (22),

$$F(u,v) = a(u,v) + n(u,v) - (f,v)_{L^2([0,1])} \quad \forall v \in V$$

we just say

```
solve(F==0, u, bc)
```

`dolfin` will normally report (exhaustively) on the convergence of Newton's method.

Hierarchy of nonlinear problems

Different levels of difficulty for nonlinear problems.

The lowest level have no nonlinearity: they are linear.

Next level are **semilinear** equations: nonlinearities do not appear in leading-order parts of the equation.

Example: the Jeffrey-Hamel flow problem

Equation is of the form

$$\sum_{|\boldsymbol{\alpha}|=k} c_{\boldsymbol{\alpha}}(\mathbf{x}) D^{\boldsymbol{\alpha}} u + c_0(\mathbf{x}, \nabla^{k-1} u, \ldots, \nabla u, u) = 0.$$

k is highest order of derivative in PDE.

Nonlinear hierarchy continued

For the Laplacian, $k = 2$.

Will see equations of order 3 and 4.

Next, **quasilinear** equations have nonlinear coefficients in the highest-order terms, but they cannot depend on the highest-order derivatives:

$$\sum_{|\boldsymbol{\alpha}|=k} c_{\boldsymbol{\alpha}}(\mathbf{x}, \nabla^{k-1}u, \ldots, \nabla u, u) D^{\boldsymbol{\alpha}} u + c_0(\mathbf{x}, \nabla^{k-1}u, \ldots, \nabla u, u) = 0.$$

An example of this is p-Lapalcean

Fully nonlinear problems

Finally, **fully nonlinear** equations can be of the form

$$\sum_{|\boldsymbol{\alpha}|=k} c_{\boldsymbol{\alpha}}(\mathbf{x}, \nabla^k u, \ldots, \nabla u, u) D^{\boldsymbol{\alpha}} u + c_0(\mathbf{x}, \nabla^{k-1} u, \ldots, \nabla u, u) = 0.$$

An example is the Monge-Ampère equation.

These are beyond our scope.

A quasi-linear problem

The **p-Laplacian** is a widely studied equation that illustrates many interesting features of nonlinear PDEs.

It is defined for $p > 1$ and takes the form

$$-\nabla \cdot \left(|\nabla u|^{p-2} \nabla u \right) = f \quad \text{in } \Omega, \tag{23}$$

with a mix of Dirichlet and Neumann boundary conditions possible on $\partial\Omega$.

When $p = 2$, we obtain the standard Poisson equation.

The variational form for the p-Laplacian is

$$a_p(u, v) = \int_\Omega |\nabla u(\mathbf{x})|^{p-2} \nabla u(\mathbf{x}) \cdot \nabla v(\mathbf{x}) \, d\mathbf{x}.$$

p-Laplacean problem

$$a_p(u,v) = \int_\Omega |\nabla u(\mathbf{x})|^{p-2} \nabla u(\mathbf{x}) \cdot \nabla v(\mathbf{x})\, d\mathbf{x}$$

For $p < 2$, coefficient $|\nabla u|^{p-2}$ misbehaves:

- $|\nabla u|^{p-2} \to \infty$ if $\nabla u \to \mathbf{0}$, and
- $|\nabla u|^{p-2} \to 0$ if $\nabla u \to \infty$.

Latter can happen for nonconvex polygonal domains.

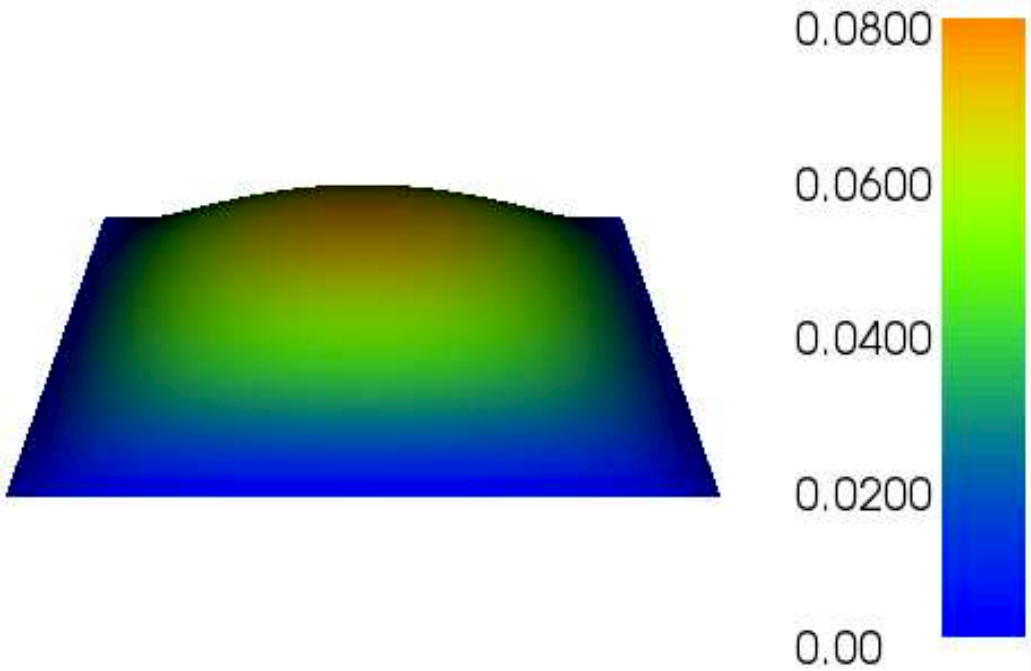
p-Laplacian of interest in image processing for $p < 2$.

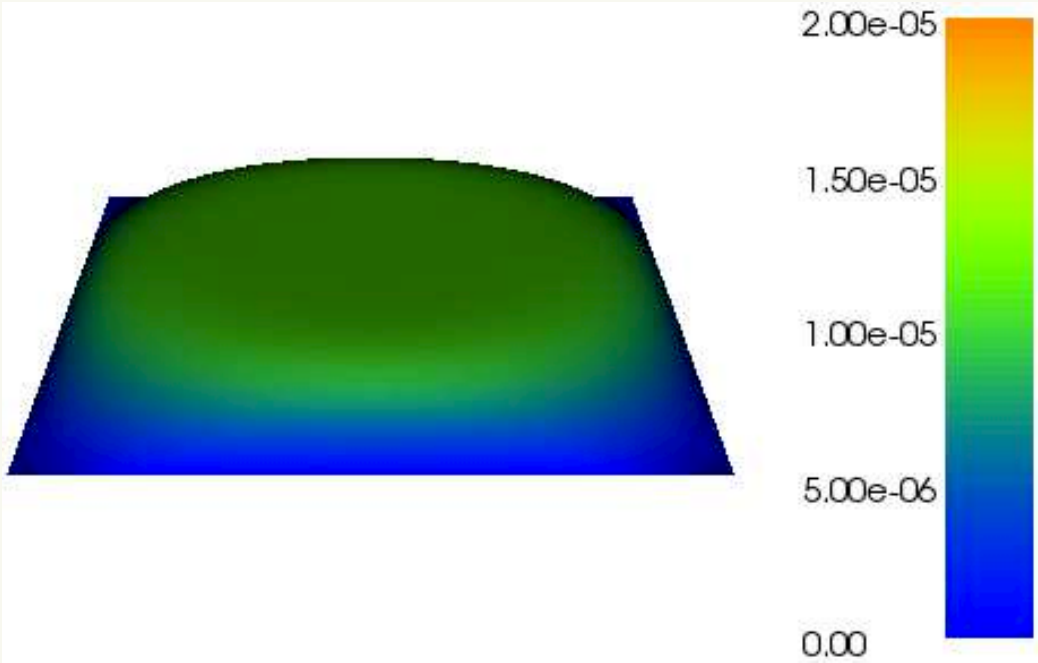
For $p > 2$, coefficient $|\nabla u|^{p-2}$ misbehaves:

- $|\nabla u|^{p-2} \to \infty$ if $\nabla u \to \infty$, and
- $|\nabla u|^{p-2} \to 0$ if $\nabla u \to \mathbf{0}$.

p-Laplacian models granular materials (sand) for $p > 2$.

Computational example of p-Laplacean





Solution of the p-Laplacian on a unit square

with $f \equiv 1$ and $u = 0$ on $\partial\Omega$,

computed using piecewise linears on a 128×128 mesh.

(left) $p = 2$ and (right) $p = 1.15$.

p-Laplacian variational form

The variational form for the p-Laplacian is

$$a_p(u, v) = \int_\Omega |\nabla u(\mathbf{x})|^{p-2} \nabla u(\mathbf{x}) \cdot \nabla v(\mathbf{x}) \, d\mathbf{x}.$$

Thus we seek u such that

$$a_p(u, v) = F(v)$$

for all v in a suitable space V.

Note that continuity and coercivity are problemmatic:

$$a_p(v, v) = \int_\Omega |\nabla v(\mathbf{x})|^p \, d\mathbf{x}.$$

Standard Lax-Milgram theory fails for p-Laplacean since

$$a_p(v, v) = \int_\Omega |\nabla v(\mathbf{x})|^p \, d\mathbf{x}.$$

Need to work with spaces of functions whose gradients are p-th power integrable.

Fortunately, not a restriction for finite element spaces.

Applications of the p-Laplacian are abundant.

There are even applications where p is allowed to be a function of $\mathbf{x}$ and vary over the domain Ω.

p-Laplacian asymptotics

Useful to examine what happens in some limits.

Let us consider a particular problem: $\Omega = [0,1]^2$, $F(v) = \int_\Omega v(\mathbf{x})\, d\mathbf{x}$ ($f \equiv 1$).

In addition, we choose homogeneous Dirichlet boundary conditions ($u = 0$) on $\partial\Omega$.

By symmetry, $\nabla u = 0$ at the middle of the square, so

- if $p < 2$, the expression $|\nabla u|^{p-2}$ is infinite there.
- if $p > 2$, the expression $|\nabla u|^{p-2}$ is zero there.

p-Laplacian computation

To avoid computational difficulties, we redefine

$$a_p(u,v) = \int_\Omega \left(\epsilon^2 + |\nabla u(\mathbf{x})|^2\right)^{(p-2)/2} \nabla u(\mathbf{x}) \cdot \nabla v(\mathbf{x})\, d\mathbf{x}, \quad (24)$$

where ϵ is a small parameter to mollify computations.

Needed for $p < 2$ when $\nabla u = \mathbf{0}$ because

- coefficient of $\nabla u(\mathbf{x}) \cdot \nabla v(\mathbf{x})$ would be infinite if we did not,
- fractional power not differntiable there, so Newton's method could misbehave.

Now coefficient is just very large when ∇u is small.

p-Laplacian asymptotics

When ∇u is small, size of u is depressed, as follows.

When ∇u is small, equation looks like

$$-\epsilon^{p-2}\Delta u = f$$

so $-\Delta u = \epsilon^{2-p} f$.

For $p < 2$, ϵ^{2-p} is small and u will be proportional to ϵ^{2-p}.

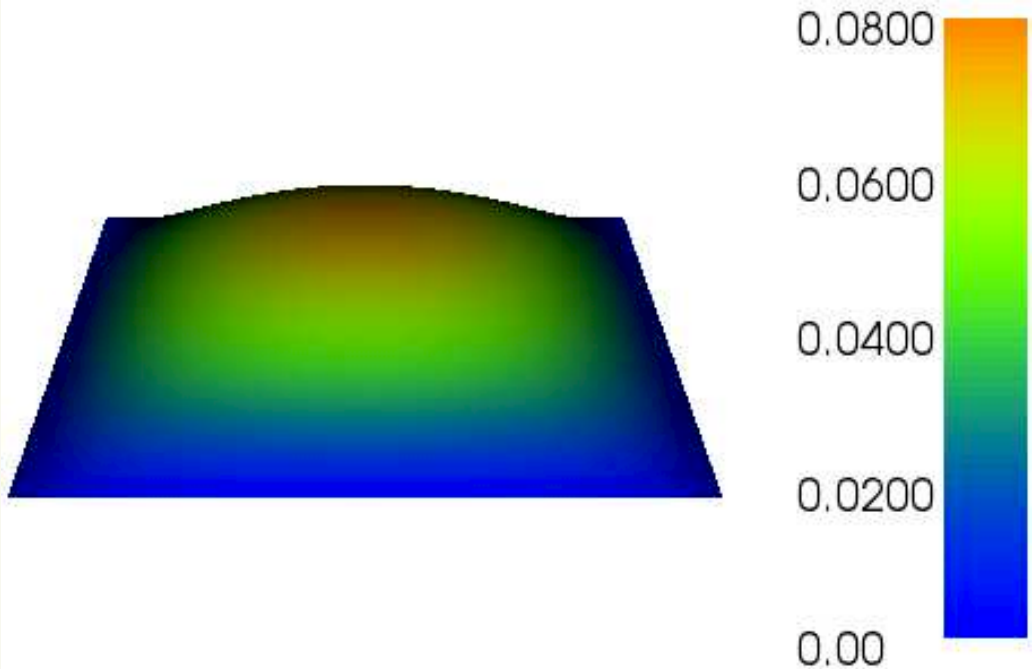
Making u smaller flattens u because $u = 0$ on $\partial\Omega$.

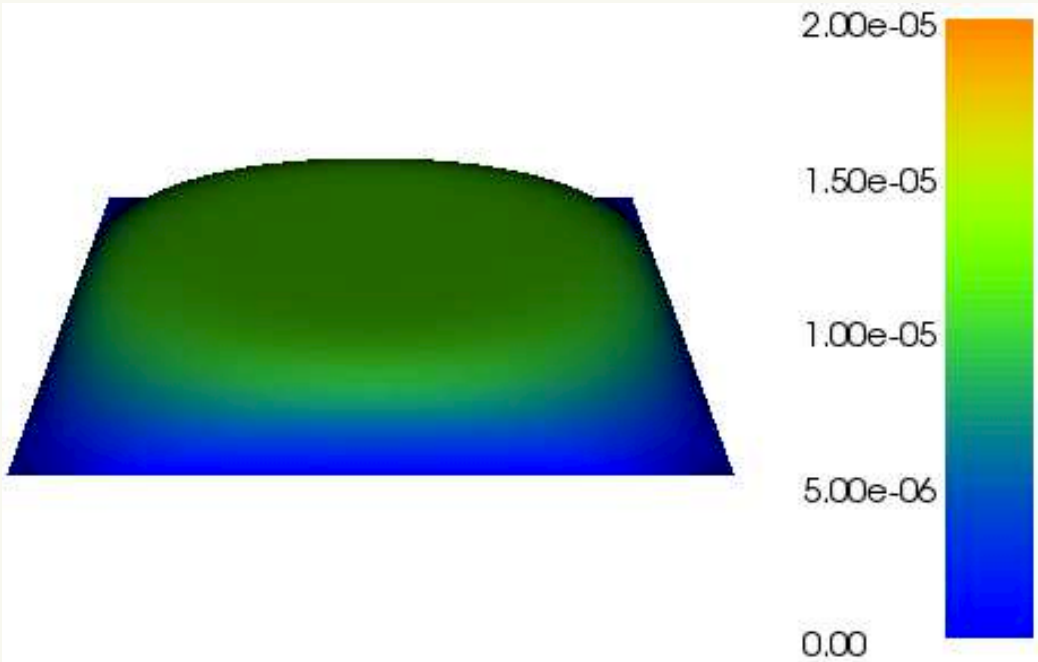
If we flatten u, region where ∇u is small increases.

A vicious cycle.

This analysis is born out in the figure and Table 1.

As p decreases, u becomes both smaller and flatter.





(left) $p = 2$ and (right) $p = 1.15$.

ϵ	p	L2norm
1.00e-05	2.00	4.13e-02
1.00e-06	1.80	2.78e-02
1.00e-06	1.60	1.45e-02
1.00e-06	1.40	4.07e-03
1.00e-06	1.30	1.17e-03
1.00e-07	1.20	1.02e-04
1.00e-07	1.15	9.36e-06
1.00e-07	1.10	1.03e-07

Table 1: $L^2(\Omega)$ norm of solution of the p-Laplacian problem for various exponents p. The fudge factor ϵ appears in the mollified form (24). Computed with piecewise linears on a 128×128 mesh.

Nonlinear solvers

Newton’s method has many benefits.

It converges rapidly, and it has a functorial definition that allows it to be applied automatically and very generally.

However, it is essentially a local method.

It performs well when it is started near a solution.

But it may behave badly otherwise.

Thus other nonlinear solvers are of interest to provide a more global approach to nonlinear PDEs.

Continuation method

One simple approach is called **continuation**.

Method applied to problems with a natural parameter.

For the p-Laplacian, p is the parameter.

To solve for large p, we can start with $p = 2$ (a linear problem) and then increase p, using solution for lower values of p as starting guesses for larger values of p.

Fortunately, in `dolfin`, the solver syntax
`solve(F == 0, u, bc)`
is interpreted to mean that if u is defined before the `solve` call, then it will be used as the initial guess.

Continuation method

Thus we can embed `solve` in a loop to generate a sequence of solutions.

For the p-Laplacian problem these solutions actually converge as p increases, so the Newton method gets faster and faster.

Other problems tend to get harder as the parameter is increased, but continuation is still a very useful algorithm.

References

[1] Susanne C. Brenner and L. Ridgway Scott. *The Mathematical Theory of Finite Element Methods*. Springer-Verlag, third edition, 2008.

[2] L.D. Landau and E.M. Lifshitz. *Fluid Dynamics*. Oxford: Pergammon, 1959.

[3] M. N. Spijker. On the structure of error estimates for finite-difference methods. *Numerische Mathematik*, 18(1):73–100, 1971.