Solving PDE’s with FEniCS

Laplace and Poisson

L. Ridgway Scott

The Institute for Biophysical Dynamics,
The Computation Institute, and the
Departments of Computer Science and Mathematics,
The University of Chicago
Define a function $u$ in a domain $\Omega \subset \mathbb{R}^d$ by

$$-\Delta u = f \text{ in } \Omega$$  \hspace{1cm} (1)

together with boundary conditions

$$u = 0 \text{ on } \Gamma \subset \partial \Omega \text{ \hspace{1cm} (Dirichlet)}$$

$$\frac{\partial u}{\partial n} = 0 \text{ on } \partial \Omega \setminus \Gamma \text{ \hspace{1cm} (Neumann)}$$  \hspace{1cm} (2)

where $\frac{\partial u}{\partial n}$ denotes the derivative of $u$ in the direction normal to the boundary, $\partial \Omega$ ($\frac{\partial u}{\partial n} = \mathbf{n} \cdot \nabla u.$)

This is known variously as Poisson’s equation or Laplace’s equation (especially when $f \equiv 0$).
This equation forms the basis of a remarkable number of physical models.

It serves as a basic equation for diffusion, elasticity, electrostatics, gravitation, and many more domains.

In potential flow, the gradient of $u$ is the velocity of incompressible, inviscid, irrotational fluid flow.

- The boundary $\partial \Omega$ is the surface of an obstacle moving in the flow, and one solves the equation on the exterior of $\Omega$. 
We will see that the right place to look for such the solution of such an equation is a Sobolev space denoted $H^1(\Omega)$ defined by

$$H^1(\Omega) = \{ v \in L^2(\Omega) : \nabla v \in L^2(\Omega)^d \}, \quad (3)$$

where by $L^2(\Omega)$ we mean functions which are square integrable on $\Omega$:

$$\|v\|_{L^2(\Omega)} = \left( \int_{\Omega} v(x)^2 \, dx \right)^{1/2} < \infty,$$

and $L^2(\Omega)^d$ means $d$ copies of $L^2(\Omega)$ (Cartesian product).
There is a natural inner-product on $L^2(\Omega)$ defined by

$$\left( v, w \right)_{L^2(\Omega)} = \int_{\Omega} v(x) w(x) \, dx,$$

(4)

For a vector-valued function $w$, e.g., $w = \nabla v$, we define

$$\| w \|_{L^2(\Omega)} = \| \| w \|_{L^2(\Omega)} = \left( \int_{\Omega} |w(x)|^2 \, dx \right)^{1/2},$$

where $|\xi|$ denotes the Euclidean norm of the vector $\xi \in \mathbb{R}^d$, and

$$\left( v, w \right)_{L^2(\Omega)} = \int_{\Omega} v(x) \cdot w(x) \, dx.$$
Figure 1: (a) Domain $\Omega$ with $\Gamma$ indicated in red. (b) Triangulation of $\Omega$.

Now consider equation (1) with boundary conditions (2). Assume $\Gamma$ has nonzero measure (length or area, or even volume, depending on dimension). Later, we will return to the case when $\Gamma$ is empty, the pure Neumann case.

Typical $\Omega$ shown in Figure 1, with $\Gamma$ shown in red.
To formulate the variational equivalent of (1) with boundary conditions (2), we define a variational space that incorporates the essential, i.e., Dirichlet, part of the boundary conditions in (2):

$$ V := \{ v \in H^1(\Omega) : v|_\Gamma = 0 \} . $$

See Table 1 for an explanation of the various names used to describe different boundary conditions.

<table>
<thead>
<tr>
<th>generic name</th>
<th>example</th>
<th>honorific name</th>
</tr>
</thead>
<tbody>
<tr>
<td>essential</td>
<td>$u = 0$</td>
<td>Dirichlet</td>
</tr>
<tr>
<td>natural</td>
<td>$\frac{\partial u}{\partial n} = 0$</td>
<td>Neumann</td>
</tr>
</tbody>
</table>

Table 1: Nomenclature for different types of boundary conditions.
The appropriate bilinear form for the variational problem is determined by multiplying Poisson's equation by a suitably smooth function, integrating over $\Omega$ and then integrating by parts:

$$ (f, v)_{L^2(\Omega)} = \int_\Omega (-\Delta u)v \, dx $$

$$ = \int_\Omega \nabla u \cdot \nabla v \, dx - \oint_{\partial \Omega} v \frac{\partial u}{\partial n} \, ds $$

$$ = \int_\Omega \nabla u \cdot \nabla v \, dx :\!:= a(u, v). \quad (6) $$

The boundary term in (6) vanishes for $v \in V$ because either $v$ or $\frac{\partial u}{\partial n}$ is zero on any part of the boundary.
The integration-by-parts formula derives from the divergence theorem

\[ \int_{\Omega} \nabla \cdot w(x) \, dx = \oint_{\partial \Omega} w(s) \cdot n(s) \, ds \]  

(7)

applied to \( w = v \nabla u \), together with \( \frac{\partial u}{\partial n} = (\nabla u) \cdot n \) and \( \Delta u = \nabla \cdot (\nabla u) \) (\( n \) is the outward-directed normal to \( \partial \Omega \)). More precisely, we observe that

\[ \nabla \cdot (v \nabla u) = \sum_{i=1}^{d} ((v \nabla u)_i)_i = \sum_{i=1}^{d} (v u)_i , i \]

\[ = \sum_{i=1}^{d} v , i u , i + v u , ii = \nabla v \cdot \nabla u + v \Delta u. \]
Thus the divergence theorem applied to $w = v\nabla u$ gives

$$\oint_{\partial \Omega} v \nabla u(s) \cdot n(s) \, ds = \int_{\Omega} (\nabla \cdot (v \nabla u))(x) \, dx$$

$$= \int_{\Omega} (\nabla v \cdot \nabla u + v \Delta u)(x) \, dx,$$

which means that

$$\int_{\Omega} -v(x) \Delta u(x) \, dx = \int_{\Omega} \nabla v(x) \cdot \nabla u(x) \, dx$$

$$- \int_{\partial \Omega} v \frac{\partial u}{\partial n} \, ds$$

$$= a(u, v) - \int_{\partial \Omega} v \frac{\partial u}{\partial n} \, ds.$$

(8)
Thus, \( u \) can be characterized via

\[
\forall v \in V \quad a(u, v) = (f, v)_{L^2(\Omega)}
\tag{9}
\]

The companion result, namely

an a solution to the variational problem in (9) solves Poisson’s equation

can also be proved [2], under suitable regularity conditions on \( u \).

These regularity conditions guarantee the relevant expressions in (1) and (2) are well defined.

We will show how this is done in detail in the one-dimensional case.
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 (9) can be written using

$$F(v) = (f, v)_{L^2(\Omega)} \quad \forall v \in V.$$  \hspace{1cm} (10)

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 $u, v \in V$. 
Critical condition on a linear functional for success in a variational formulation: \textit{bounded} (a.k.a. \textit{continuous}).

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

$$|F(v)| \leq C_F \|v\|_V \quad \forall v \in V. \quad (11)$$

A natural norm $\|\cdot\|_V$ for the space $V$ defined in (5) is

$$\|v\|_a = \sqrt{a(v,v)}. \quad \text{The smallest possible constant } C_F \text{ for which } (11) \text{ holds is called the dual norm of } F.$$
The **dual norm** of $F$ is defined by

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

(12)

The linear form (10)

$$F(v) = (f, v)_{L^2(\Omega)} \quad \forall v \in V$$

is bounded on $H^1(\Omega)$: $L^2$ norm is part of $H^1$ norm.

But other linear forms 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 $v(x) := |x - x_0|^{2/3}$. Note that $v \in H^1([0, 1])$. 
Our definition of the variational space $V$ ensures that $a(v, v) < \infty$ for all $v \in V$.

But our variational formulation involves $a(u, v)$ for $u, v \in V$.

Why is $a(u, v)$ well defined for all $u, v \in V$?

The Cauchy-Schwarz inequality guarantees that

$$|a(u, v)| \leq \|u\|_a \|v\|_a,$$

where $\|v\|_a = \sqrt{a(v, v)}$ for all $v \in V$.

The proof is a simple calculation.
Cauchy-Schwarz inequality proof

\[ a(u - tv, u - tv) = a(u, u) - 2ta(v, u) + t^2a(v, v) \]
\[ = a(u, u) - 2ta(u, v) + t^2a(v, v). \]  
(13)

In particular, since \( a(u - tv, u - tv) \geq 0 \),

\[ 2ta(u, v) \leq a(u, u) + t^2a(v, v). \]  
(14)

For example, suppose that \( a(v, v) = 0 \).

Choose the sign of \( t \) to be the sign of \( a(u, v) \). From (14) we conclude that

\[ 2|t| |a(u, v)| \leq a(u, u). \]  
(15)

Since \( 2|t| |a(u, v)| \leq a(u, u) \) holds for all \( t \in \mathbb{R} \), we can let \( |t| \to \infty \) to conclude that \( a(u, v) = 0 \).
Cauchy-Schwarz inequality

If \( a(v, v) \neq 0 \), define \( t = \text{sign}(a(u, v)) \| u \|_a / \| v \|_a \).

If by chance \( a(u, u) = 0 \), then we reverse the previous argument to conclude that again \( a(u, v) = 0 \).

If not zero, and thus \( t \neq 0 \), divide by \( |t| \) in (14) to get

\[
2|a(u, v)| \leq \frac{1}{|t|} a(u, u) + |t| a(v, v) = 2\| u \|_a \| v \|_a. \tag{16}
\]

Thus we have proved the Cauchy-Schwarz inequality

\[
|a(u, v)| \leq \| u \|_a \| v \|_a. \tag{17}
\]

The Cauchy-Schwarz inequality is generally true for any non-negative, symmetric bilinear form.
Cauchy-Schwarz often stated for an inner-product. Our bilinear form $a(\cdot, \cdot)$ is almost an inner-product except that it lacks one condition, non-degeneracy. In our case $a(v, v) = 0$ if $v$ is constant, and for an inner-product, this is not allowed.

One example of an inner-product is the bilinear form

$$ (u, v)_{L^2(\Omega)} = \int_{\Omega} u(x) v(x) \, dx. \quad (18) $$

Here we see that $(v, v)_{L^2(\Omega)} = 0$ implies that $v \equiv 0$. But the Cauchy-Schwarz inequality does not require this additional property to be valid.
The case of pure Neumann (or natural) boundary conditions

\[ \frac{\partial u}{\partial n} = 0 \text{ on } \partial \Omega \quad (19) \]

(i.e., when \( \Gamma = \emptyset \)) is a bit different.

In particular, solutions are unique only up to an additive constant, and they can exist only if the right-hand side \( f \) in (1) satisfies

\[ \int_{\Omega} f(x) \, dx = \int_{\Omega} -\Delta u(x) \, dx \]
\[ = \int_{\Omega} \nabla u(x) \cdot \nabla 1 \, dx - \int_{\partial \Omega} \frac{\partial u}{\partial n} \, ds = 0. \quad (20) \]
A variational space appropriate for the present case is

$$V = \left\{ v \in H^1(\Omega) : \int_\Omega v(x) \, dx = 0 \right\}. \quad (21)$$

For any integrable function $g$, we define its mean, $\bar{g}$, as follows:

$$\bar{g} := \frac{1}{\text{meas}(\Omega)} \int_\Omega g(x) \, dx. \quad (22)$$

For any $v \in H^1(\Omega)$, note that $v - \bar{v} \in V$.

Then $u - \bar{u}$ satisfies the variational formulation (9) with $V$ defined as in (21).
Conversely, if \( u \in H^2(\Omega) \) solves the variational equation (9) with \( V \) defined as in (21), then \( u \) solves Poisson’s equation (1) with a right-hand-side given by

\[
\tilde{f}(x) := f(x) - \bar{f} \quad \forall x \in \Omega
\]  

(23)

with boundary conditions (19).

The variational form \( a(\cdot, \cdot) \) coercive on the spaces \( V \) [2]: there is a constant \( C \) depending only on \( \Omega \) and \( \Gamma \) such that

\[
\|v\|^2_{L^2(\Omega)} \leq C a(v, v) \quad \forall v \in V.
\]  

(24)
We show later how to prove such a result in the one-dimensional case.

From (24), it follows that the problem (9) 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:

$$0 = (f, u)_{L^2} = a(u, u) \geq \frac{1}{C} \|u\|_{H^1(\Omega)}^2.$$
In the finite-dimensional case, this uniqueness also implies existence, and a similar result (Lax-Milgram Theorem) holds in the setting of infinite dimensional Hilbert spaces such as $V$.

Moreover, the coercivity condition immediately implies a stability result, namely

$$
\|u\|_{H^1(\Omega)} \leq \frac{Ca(u, u)}{\|u\|_{H^1(\Omega)}} = C\frac{(f, u)_{L^2}}{\|u\|_{H^1(\Omega)}} \leq C\|f\|_{V'},
$$

(25)

where the dual norm $\|f\|_{H^1(\Omega)'}$ is defined in (12).
When coercivity holds, the Lax-Milgram theorem guarantees variational problem (9) has unique solution.

There is an additional continuity condition that usually is straight-forward, namely that the form $a(\cdot,\cdot)$ is bounded on $V$, that is,

$$|a(u,v)| \leq C\|u\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)} \quad \text{for all } u, v \in V.$$ \hspace{1cm} (26)

In most cases, this condition is evident, but not in all as we describe later.
Often easy to see that

\[ a(v, v) \leq C\|v\|_{H^1(\Omega)}^2 \]

for all \( v \in V \).

Connection between this condition and (26) given by the Cauchy-Schwarz inequality (17).

The continuity bound in (25) holds for discrete approximations as well under a very simple condition: \( V_h \subset V \).
Lax-Milgram Theorem Suppose that the variational form $a(\cdot, \cdot)$ is coercive (24) and continuous (26) (bounded) on $H^1(\Omega)$. Then the variational problem (9) has a unique solution $u$ for every continuous (bounded) $F$ defined on $H^1(\Omega)$. Moreover,

$$\|u\|_{H^1(\Omega)} \leq c_1 c_0 \sup_{v \in H^1(\Omega)} \frac{|F(v)|}{\|v\|_{H^1(\Omega)}},$$  \hspace{1cm} (27)

where $c_0$ is the constant in (24) and $c_1$ is the constant in (26).

The combination of continuity and coercivity correspond to the **stability** of the numerical scheme.
Let $T_h$ denote a subdivision of $\Omega$; typically this will be what is called a triangulation, made of triangles in two dimensions or tetrahedra in three dimensions.

Triangulation of the domain in (a) is shown in (b).

The main requirement for a triangulation is that no vertex of a triangle can be in the middle of an edge. However, more general subdivisions can be used that violate this property [1, 8, 9].
Using elements

Figure 2: (a) Domain $\Omega$ with $\Gamma$ indicated in red. (b) Nodal positions for $V_h$ are indicated by the black dots; note that vertices in $\Gamma$ are not included, to respect the essential (Dirichlet) boundary condition.

Main concept of finite element method: use each element of the subdivision as a separate domain in which to reason about the balance of forces.

Mathematically, corresponds to choosing functions on each element to represent variables used in the model.
Often, same functions used on each element, but not necessary \([8, 9, 5]\).

In this way, one constructs a finite dimensional space \(V_h\) which can be used in what is known as the **Galerkin method** to approximate the variational formulation (9), as follows:

\[
\text{find } u_h \in V_h \text{ satisfying } a(u_h, v) = (f, v) \quad \forall v \in V_h. \quad (28)
\]

Here we can think of \(h\) as designating the subdivision, or perhaps as a parameter that denotes the size of the elements of the subdivision.
The coercivity condition implies stability for the discrete approximation, namely

\[
\|u_h\|_{H^1(\Omega)} \leq \frac{Ca(u_h, u_h)}{\|u_h\|_{H^1(\Omega)}} = C \frac{(f, u_h)}{\|u_h\|_{H^1(\Omega)}} \leq C \|f\|_{V'}, \tag{29}
\]

where we will explain the meaning of \(\|f\|_{V'}\) later.

In particular, if \(f \equiv 0\), then \(u_h \equiv 0\).

Provided \(V_h\) is finite dimensional, this implies that (28) always has a unique solution.

We can see this more clearly by choosing a basis \(\{\phi_i \in V_h : i = 1, \ldots, N_h\}\). Write \(u_h = \sum_i U_i \phi_i\).
Using the linearity of the form $a(\cdot, \cdot)$ in each of its variables, we obtain the linear system $AU = F$ where

$$A_{ij} = a(\phi_i, \phi_j), \quad F_i = \int_{\Omega} f(x) \phi_i(x) \, dx \quad \forall i, j = 1, \ldots, N_h.$$
Since $A$ is symmetric ($A_{ji} = a(\phi_j, \phi_i) = a(\phi_i, \phi_j) = A_{ij}$), we have, for all $j = 1, \ldots, N_h$,

$$F_j = \int_{\Omega} f(x) \phi_j(x) \, dx = a(u_h, \phi_j) = a\left(\sum_i U_i \phi_i, \phi_j\right)$$

$$= \sum_i U_i a(\phi_i, \phi_j) = \sum_i U_i A_{ij} = \sum_i A_{ji} U_i = (AU)_j$$

From linear algebra, solution to linear system $AU = F$ exists uniquely $\iff$ only solution for $F = 0$ is $U = 0$.

The latter is guaranteed by the coercivity condition (25):

$$\|u_h\|_{H^1(\Omega)}^2 \leq Ca(u_h, u_h) = C(f, u_h) = 0.$$
Given a triangulation, the simplest space $V_h$ that we can construct is the set of continuous piecewise linear functions.

On each triangle (or tetrahedron), such functions are linear, and moreover we contrive to make them continuous.

A linear function is determined by its values at the vertices of a simplex.

Easy to see in one or two dimensions; graph of function is line or plane through specified values at vertices.

If values of $v \in V_h$ at vertices agree in all of the triangles meeting there, then resulting function is continuous.
In two dimensions, the values along edges are specified completely by the values at the vertices.

A basis for $V_h$ satisfies $\phi_i(x_j) = \delta_{ij}$ (Kronecker $\delta$) as depicted in Figure 3.

The vertices of a triangulation provide the nodes of the space $V_h$; these are shown as black dots in Figure 2. Only vertices where nodal values are non-zero have black dots, where the boundary condition $v = 0$ holds on $\Gamma$ for $v \in V_h \subset V$. 
What determines the accuracy of the approximation?

Céa’s Theorem [2, 2.8.1] says the following.

**Theorem 0.1** Suppose that $V_h \subset V$, that the variational form $a(\cdot, \cdot)$ is bounded and coercive on $V$, and that $F$ is bounded on $V$. Then

$$
\|u - u_h\|_a \leq C \inf_{v \in V_h} \|u - v\|_a.
$$

(30)

Thus the accuracy of the finite element method is determined by the accuracy of approximation.

Often called quasi-optimal.
Nodal values for quadratics

Figure 4: Nodes for quadratics: vertices and edge midpoints.

For more accurate, cost effective approximation, often useful to use higher-order polynomials in each element.

In Figure 4 we see the nodes for piecewise quadratic functions (compare Figure 1).
Again, we can define a basis for the space $V_h$ of continuous piecewise quadratics in terms of functions that satisfy

$$\phi_i(x_j) = \delta_{ij} \text{ (Kronecker } \delta),$$

where the $x_j$’s are the nodes in Figure 4.

But now it is not so clear how we can be sure that this is a valid representation.

What we need to know is that this nodal representation is **unisolvent** on each triangle, meaning that

on each triangle you can solve uniquely for a quadratic given the values at the specified nodes, the vertices and edge midpoints.
Proof of unisolvence: degree reduction

On each edge, we have three distinct points that determine uniquely a quadratic, simply by invoking the fundamental theorem of algebra.

If all nodal values on one edge vanish, then the quadratic \( q(x, y) \) must vanish on that edge.

WLOG suppose that edge lies on the \( x \)-axis. Then \( q(x, y) = y \ell(x, y) \) where \( \ell \) is a linear polynomial in \( x, y \).

Can be verified by expanding \( q \) in powers of \( x \) and \( y \) (there are 6 terms) and invoking \( q(x, y) \equiv 0 \) on the edge lying on the \( x \)-axis.
$q$ also vanishes on the other two edges of the triangle, neither of which can lie on the $x$-axis, so that means that $\ell$ must also vanish on these edges.

But this clearly implies that $\ell \equiv 0$, and thus $q \equiv 0$.

By linear algebra, uniqueness of the representation implies existence of a representation:

We have exactly 6 nodal variables matching the dimension of space of quadratic polynomials in two dimensions.

Complete details are found in [2, Chapter 3].
No limit on the degree of polynomials that can be used.

General family of elements called **Lagrange elements**.

Regular pattern of nodes shown in Figure 5.

Figure 5: Varying mesh number $M$ and polynomial degree $k$ with the same number of nodes: (a) $M = 4, k = 1$ (linears), (b) $M = 2, k = 2$ (quadratics), (c) $M = 1, k = 4$ (quartics).
Using the method of **manufactured solutions**, consider

\[-\Delta u = 2\pi^2 \sin(\pi x) \sin(\pi y) \text{ in } \Omega = [0, 1] \times [0, 1] \]

\[u = 0 \text{ on } \partial \Omega,\]

whose solution is \(u(x, y) = \sin(\pi x) \sin(\pi y)\). Of course, we started with the solution \(u\) and then computed its Laplacian to get \(f = 2\pi^2 \sin(\pi x) \sin(\pi y)\).

The errors

\[\|u_h - u\|_{L^2(\Omega)} = \|u_h - (2\pi^2)^{-1} f\|_{L^2(\Omega)}\]

for different meshes (of the type shown in Figure 5) and polynomial degrees, together with execution times, are given in Table 2.
Computational experiments with (31).

<table>
<thead>
<tr>
<th>degree</th>
<th>mesh number</th>
<th>$L^2$ error</th>
<th>time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>32</td>
<td>2.11e-03</td>
<td>0.09</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>5.65e-04</td>
<td>0.08</td>
</tr>
<tr>
<td>1</td>
<td>128</td>
<td>1.32e-04</td>
<td>0.31</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>6.93e-05</td>
<td>0.08</td>
</tr>
<tr>
<td>1</td>
<td>256</td>
<td>3.31e-05</td>
<td>1.07</td>
</tr>
<tr>
<td>2</td>
<td>64</td>
<td>1.08e-06</td>
<td>0.23</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>7.78e-07</td>
<td>0.08</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>7.29e-08</td>
<td>0.08</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>2.44e-08</td>
<td>0.11</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>1.61e-09</td>
<td>0.09</td>
</tr>
<tr>
<td>4</td>
<td>32</td>
<td>7.64e-10</td>
<td>0.23</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>1.42e-10</td>
<td>0.09</td>
</tr>
<tr>
<td>4</td>
<td>64</td>
<td>2.39e-11</td>
<td>0.74</td>
</tr>
<tr>
<td>4</td>
<td>128</td>
<td>4.95e-12</td>
<td>3.0</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>3.98e-12</td>
<td>0.13</td>
</tr>
<tr>
<td>8</td>
<td>16</td>
<td>1.67e-11</td>
<td>0.33</td>
</tr>
</tbody>
</table>

Table 2: Computational experiments with solving the problem (31).
Arbitrary degree polynomials

What we see in Table 2 is that the error can be reduced substantially by using higher-order polynomials.

Increasing the mesh number for linear Lagrange elements does reduce the error, but the execution time grows commensurately with the error reduction.

Using linears on a mesh of size 256 gives half the error of quadratics on a mesh of size 16, but the latter computation requires one-tenth of the time.

For the same amount of time as this computation with quadratics, using quartics on a mesh of size 8 gives an error almost two orders of magnitude smaller.
Arbitrary degree polynomials

Each mesh with double the number of mesh points was derived from the one with the smaller number of points by subdividing each triangle into four similar triangles.

To get the highest accuracy, the best strategy is to use higher polynomial order, up to a point.

The most accurate computation occurs with polynomial degree 8 with a mesh number of 8.

But the error quits decreasing at a certain point due to round-off error.

Will discuss the effects of finite precision arithmetic in more detail later.
Arbitrary degree polynomials

The times presented here should be viewed as approximate.

There is significant variation due to system load from run to run.

These computations were done on a MacBook Pro with 2.3 GHz Intel Core i7 and 16 GB 1600 MHz DDR3 memory.

However, we do see order of magnitude variation depending on the mesh size and polynomial degree.
It is easy to understand the basic error behavior for the finite element method.

In the case of both piecewise linear and piecewise quadratics, we described the nodal basis functions $\phi_i$ which satisfy $\phi_i(x_j) = \delta_{ij}$ (Kronecker $\delta$), where $x_j$ denotes a typical node.

For linears, the nodes are the vertices, and for quadratics the edge midpoints are added.

For higher degree Lagrange elements, more edge nodes are involved, as well as interior nodes.
For example, with cubics, the centroid of each triangle is a node.

Nodal representation is basis for defining an interpolant.

Using such a nodal representation, we can define what is known as a **global interpolant** $I_h$ defined on continuous functions, by

$$I_h u = \sum_i u(x_i) \phi_i.$$  \hspace{1cm} (33)

Thus $I_h$ maps continuous functions into the space $V_h$ used in finite element computations.
Interpolant approximation

Let $I_h$ denote a global interpolant for a family of finite elements based on the components of $\mathcal{T}^h$.

Suppose that $I_h u$ is continuous, as is true for the Lagrange family of elements.

Further, suppose that the corresponding shape functions have an approximation order, $m$, that is

$$
\| u - I_h u \|_{H^1(\Omega)} \leq C h^{m-1} |u|_{H^m(\Omega)}.
$$

(34)

In order to have good approximation, we need to have

$$
I_h \left( V \cap C^k(\Omega) \right) \subset V_h,
$$

(35)

where $k = 0$ for Lagrange elements.
However, we allow for the possibility that \( k > 0 \) since this holds for other element families.

Condition \( I_h \left( V \cap C^k(\Omega) \right) \subset V_h \), is a mesh constraint.

If \( V_h \subset V \) and the form \( a(\cdot, \cdot) \) is bounded and coercive, then the unique solution, \( u_h \in V_h \), to the variational problem

\[
a(u_h, v) = (f, v) \quad \forall v \in V_h
\]

satisfies

\[
\|u - u_h\|_{H^1(\Omega)} \leq C \inf_{v \in V_h} \|u - v\|_{H^1(\Omega)}.
\]

(36)

by Céa’s theorem.
If conditions (34) and (35) hold, then

\[ \| u - u_h \|_{H^1(\Omega)} \leq Ch^{m-1} |u|_{H^m(\Omega)}. \]

The requirements \( V_h \subset V \) and (35) place a constraint on the subdivision in the case that \( \Gamma \) is neither empty nor all of the boundary.

These requirements provide the **consistency** of the numerical approximation.

Necessary to choose the mesh so that it aligns properly with the points where the boundary conditions change from Dirichlet to Neumann.
If the points where the boundary conditions change are vertices in the triangulation, then

\[ V_h := I_h \left( V \cap C^0(\Omega) \right) \]

is same as space of piecewise polynomials that vanish on edges contained in \( \Gamma \).

Since we have chosen the mesh so that the edges contained in \( \Gamma \) form a subdivision of the latter, it follows that \( V_h \subset V \) holds.

On the other hand, if the set of edges where functions in \( V_h \) vanish is too small, then \( V_h \not\subset V \).
Matching Boundary Conditions

If the set of edges where functions in $V_h$ vanish is too big, (35) fails to hold.

In the case of pure Dirichlet data, i.e., $\Gamma = \partial \Omega$, then $V_h$ is just the set of piecewise polynomials that vanish on the entire boundary.

In the case of pure Neumann data, i.e., $\Gamma = \emptyset$, $V_h$ is the entire set of piecewise polynomials with no constraints at the boundary.

Even if finite element space matches $\Gamma$ correctly, there is a singularity associated with changing boundary condition type along a straight boundary, as we detail subsequently.
When boundary conditions are equal to zero, we often call them homogeneous, whereas we refer to nonzero boundary conditions as inhomogeneous.

Inhomogeneous boundary conditions are easily treated.

For example, suppose that we wish to solve (1) with boundary conditions

$$u = g_D \text{ on } \Gamma \subset \partial \Omega \quad \text{and} \quad \frac{\partial u}{\partial n} = g_N \text{ on } \partial \Omega \setminus \Gamma, \quad (37)$$

where $g_D$ and $g_N$ are given.

For simplicity, let us assume that $g_D$ is defined on all of $\Omega$, with $g_D \in H^1(\Omega)$ and that $g_N \in L^2(\partial \Omega \setminus \Gamma)$. 
Recall the space (5): \( V = \{ v \in H^1(\Omega) : v|_\Gamma = 0 \} \).

Then the variational formulation of (1), (37) is as follows: find \( u \) such that \( u - g_D \in V \) and such that

\[
a(u, v) = (f, v)_{L^2(\Omega)} + \oint_{\partial\Omega \setminus \Gamma} g_N v \, ds \quad \forall v \in V.
\]  

This is well-posed since the linear form

\[
F(v) := (f, v)_{L^2(\Omega)} + \oint_{\partial\Omega \setminus \Gamma} g_N v \, ds
\]

is well defined (and continuous) for all \( v \in V \).
Equivalence of formulations

Equivalence of these formulations follows from (8): for any $v \in V$,

$$
\int_{\Omega} (-\Delta u)v \, dx = \int_{\Omega} \nabla u \cdot \nabla v \, dx - \oint_{\partial\Omega} v \frac{\partial u}{\partial n} \, ds
$$

(39)

$$
= a(u, v) - \oint_{\partial\Omega \setminus \Gamma} v \frac{\partial u}{\partial n} \, ds.
$$

Thus, if $u$ solves (1) with boundary conditions (37), then (38) follows as a consequence.

Conversely, if $u$ solves (38) then choosing $v$ to vanish near $\partial\Omega$ shows that (1) holds, and thus

$$
\oint_{\partial\Omega \setminus \Gamma} g_N v \, ds - \oint_{\partial\Omega \setminus \Gamma} v \frac{\partial u}{\partial n} \, ds = 0 \quad \forall v \in V.
$$

Choosing $v$ to be arbitrary proves (37) follows.
Finite element approximation of (38) involves, typically, use of an interpolant, $I_h g_D$, of the Dirichlet data.

We pick a subspace $V_h$ of $V$ just as before, and we seek $u_h$ such that $u_h - I_h g_D \in V_h$ and such that

$$a(u_h, v) = (f, v)_{L^2(\Omega)} + \int_{\partial\Omega \setminus \Gamma} g_N v \, ds \quad \forall v \in V_h.$$  \hfill (40)

We can cast this in a more standard form as: find \( \hat{u}_h = u_h - I_h g_D \in V_h \) such that

$$a(\hat{u}_h, v) = (f, v)_{L^2(\Omega)} + \int_{\partial\Omega \setminus \Gamma} g_N v \, ds + a(I_h g_D, v) \quad \forall v \in V_h.$$  \hfill (41)

Then we can set $u_h = \hat{u}_h + I_h g_D$.

Fortunately, the dolfin built-in function solve automates all of this, so that the data $g_D$ just needs to be specified.
Robin boundary conditions

It is frequently the case that more complex boundary conditions arise in physical models.

The so-called Robin boundary conditions take the form

\[ \alpha u + \frac{\partial u}{\partial n} = 0 \text{ on } \partial \Omega \setminus \Gamma, \]

where \( \alpha \) is a positive measurable function.

This will be coupled as before with a Dirichlet condition on \( \Gamma \).

A variational formulation for this problem can be derived as follows.
Let $V$ be the space defined in (5) with the added proviso that $V = H^1(\Omega)$ in the case that $\Gamma = \emptyset$.

From (8), we get

$$
(f, v)_{L^2(\Omega)} = \int_{\Omega} (-\Delta u(x))v(x) \, dx
$$

$$
= \int_{\Omega} \nabla u(x) \cdot \nabla v(x) \, dx - \int_{\partial\Omega} v(s) \frac{\partial u}{\partial n}(s) \, ds
$$

$$
= \int_{\Omega} \nabla u(x) \cdot \nabla v(x) \, dx + \int_{\partial\Omega} \alpha(s) v(s) u(s) \, ds,
$$

after substituting the boundary condition $\frac{\partial u}{\partial n} = -\alpha u$ on $\partial\Omega \setminus \Gamma$ and using the condition (5) that $v = 0$ on $\Gamma$. 

Brown Feb 2017

58/110
Thus we define a new variational form

\[ a_{\text{Robin}}(u, v) : = \int_{\Omega} \nabla u(x) \cdot \nabla v(x) \, dx \]

\[ + \int_{\partial \Omega} \alpha(s) v(s) u(s) \, ds. \]  \hspace{1cm} (43)

Variational formulation for the equation (1) together with the Robin boundary condition (42) takes the usual form

\[ u \in V \text{ satisfies } a_{\text{Robin}}(u, v) = (f, v)_{L^2(\Omega)} \quad \forall v \in V. \]  \hspace{1cm} (44)

A solution to the variational problem (44) solves both (1) and (42) under suitable smoothness conditions.
Note that $a_{\text{Robin}}(\cdot, \cdot)$ is coercive on $H^1(\Omega)$, that is there is a constant $C < \infty$ such that

$$\|v\|_{H^1(\Omega)}^2 \leq C a_{\text{Robin}}(v, v) \quad \forall v \in H^1(\Omega).$$

(45)

This holds even if $\Gamma = \emptyset$.

Coercivity follows because $\alpha > 0$:

$$a_{\text{Robin}}(v, v) = a(v, v) + \int_{\partial\Omega} \alpha(s) v(s)^2 \, ds.$$

(Proof is a Sobolev space exercise.)

If $\alpha$ were negative, it might not be coercive.
The geometry of the domain boundary has a significant impact on the regularity of the solution.

We begin by considering the problem

\[-\Delta u = 0 \text{ in } \Omega\]
\[u = g \text{ on } \partial \Omega,\]

where \( \Omega \) is a polygonal domain in \( \mathbb{R}^2 \).

We will see that the principal singularity of the solution can be identified, associated with what are often called re-entrant vertices.
The **L-shaped** domain $\Omega$ is depicted in Figure 6(a):

$$\Omega = \left\{ (x, y) \in [-1, 1]^2 : (x, y) = (r \cos \theta, r \sin \theta), \ 0 \leq r \leq 1, \ 0 < \theta < \frac{3}{2} \pi \right\}$$

defined using polar coordinates $(x, y) = r(\cos \theta, \sin \theta)$. 

Figure 6: (a) L-shaped domain, (b) re-entrant corner of angle $\kappa$. 

L-shaped domain
Re-entrant corners

Again using polar coordinates, define

\[ g(r(\cos \theta, \sin \theta)) = r^{2/3} \sin \left( \frac{2}{3} \theta \right). \]  \hspace{1cm} (47)

\( \partial \Omega \) has two parts: the convex part \( \Gamma_c = A \cup B \cup C \cup D \) where

\[ A = \{(1, y) : 0 \leq y \leq 1\}, \quad B = \{(x, 1) : -1 \leq x \leq 1\}, \]
\[ C = \{(-1, y) : -1 \leq y \leq 1\}, \quad D = \{(x, -1) : 0 \leq x \leq 1\}, \]  \hspace{1cm} (48)

(see Figure 6) and the re-entrant part

\[ \Gamma_r = \{(0, y) : -1 \leq y \leq 0\} \cup \{(x, 0) : 0 \leq x \leq 1\}. \]  \hspace{1cm} (49)

Then our data \( g = 0 \) on \( \Gamma_r \). Moreover, \( g \) is \textbf{harmonic}, meaning \( \Delta g = 0 \).

This follows immediately from complex analysis, since \( g \) is the imaginary part of the complex analytic function \( e^{(2/3)z} \).
Partial derivation

Deriving such a result is not easy using calculus. First of all, using polar coordinates \((x, y) = r(\cos \theta, \sin \theta)\), we find

\[
\nabla r = \frac{(x, y)}{r} \quad \text{and} \quad \nabla \theta = \frac{(-y, x)}{r^2}.
\]

This means that

\[
\nabla g(x, y) = \frac{2}{3} \left( (\nabla r)r^{-1/3} \sin(\frac{2}{3}\theta) + (\nabla \theta)r^{2/3} \cos(\frac{2}{3}\theta) \right)
\]

\[
= \frac{2}{3}r^{-4/3} \left( (x, y) \sin(\frac{2}{3}\theta) + (-y, x) \cos(\frac{2}{3}\theta) \right)
\]

\[
= \frac{2}{3}r^{-4/3} \left( x \sin(\frac{2}{3}\theta) - y \cos(\frac{2}{3}\theta), y \sin(\frac{2}{3}\theta) + x \cos(\frac{2}{3}\theta) \right)
\]

\[
= \frac{2}{3}r^{-1/3} \left( - \sin(\frac{1}{3}\theta), \cos(\frac{1}{3}\theta) \right),
\]

where trigonometric identities flow from the expressions \((\iota = \sqrt{-1})\)

\[
\cos(\frac{1}{3}\theta) - \iota \sin(\frac{1}{3}\theta) = \cos(-\frac{1}{3}\theta) + \iota \sin(-\frac{1}{3}\theta) = e^{-\iota(1/3)\theta} = e^{-\iota\theta} e^{\iota(2/3)\theta}
\]

\[
= (\cos \theta - \iota \sin \theta) (\cos(\frac{2}{3}\theta) + \iota \sin(\frac{2}{3}\theta))
\]

\[
= (\cos \theta \cos(\frac{2}{3}\theta) + \sin \theta \sin(\frac{2}{3}\theta)) + \iota ( - \sin \theta \cos(\frac{2}{3}\theta) + \cos \theta \sin(\frac{2}{3}\theta))
\]

(50)
The immediate result of the calculation (50) is that, for $0 < \theta < \frac{3}{2}\pi$,

$$|\nabla g(x, y)| \text{ blows up like } |(x, y)|^{-1/3}, \text{ since}$$

$$|\nabla g(x, y)| = |\nabla g(r \cos \theta, r \sin \theta)| = \frac{2}{3} r^{-1/3} = \frac{2}{3} |(x, y)|^{-1/3}.$$

Therefore $|\nabla g(x, y)|$ is square integrable, but it is obviously not bounded.

Benefit of working with Sobolev spaces:

allows $g$ to be considered a reasonable function even though it has an infinite gradient.
We can in principle use the vector calculus identity

\[ \nabla \cdot (\phi \psi) = \nabla \phi \cdot \psi + \phi \nabla \cdot \psi \]

to compute

\[ \Delta g = \nabla \cdot (\nabla g) \]

to verify that \( \Delta g = 0 \), but the algebra is daunting.

Exercise: compute solution via variational problem (40), see if \( u = g \) throughout \( \Omega \).

Another exercise: verify that \( \Delta g = 0 \) analytically using polar coordinates.
Singularity like the L-shaped domain occurs for any domain with **non-convex vertex** depicted in Figure 6(b), where the angle of the **re-entrant vertex** is $\kappa$.

The L-shaped domain corresponds to $\kappa = \frac{3}{2}\pi$.

The principle singularity for such a domain is of the form

$$g_\kappa(r(\cos \theta, \sin \theta)) = r^{\pi/\kappa} \sin((\pi/\kappa)\theta).$$

(51)

Note that when $\kappa < \pi$ (a convex vertex), the gradient of $g_\kappa$ is bounded.

**Exercise:** explore this general case for various values of $\kappa$. 
The largest that $\kappa$ can be is $2\pi$ which corresponds to a slit domain.

We have $g_{2\pi} = \sqrt{r} \sin(\frac{1}{2} \theta)$, which is still in $H^1(\Omega)$.

The slit domain is often a model for crack propagation.

An illustration of a problem on a slit domain is given by

$$-\Delta u = 1 \text{ in } [0, 1] \times [-1, 1]$$
$$u = 0 \text{ on } \Gamma, \quad \frac{\partial u}{\partial n} = 0 \text{ on } \partial \Omega \setminus \Gamma,$$

where $\Gamma = \{(x, 0) : x \in [\frac{1}{2}, 1]\}$.  

(52)
The solution of (52) is depicted in Figure 7, where only the top half of the domain (that is, \([0, 1] \times [0, 1]\)) is shown.

The solution in the bottom half of the domain can be obtained by symmetric reflection across the \(x\)-axis.

The square-root singularity is clearly visible.
Figure 7: Illustration of the singularity that can occur when boundary condition types are changed, cf. (53), as well as a cut-away of the solution to slit problem (52). Computed with piecewise linears on the indicated mesh.
General non-convex domains

The range of $\kappa$ values for a realistic polygonal domain excludes a region around $\kappa = 0$ and $\kappa = \pi$.

In particular, we see that $\kappa = \pi$ does not yield a singularity; the boundary is a straight line in this case, and $g_\pi(x, y) = r \sin \theta = y$, which is not singular.

When $\kappa = 0$, there is no interior in the domain near this point.

Thus for any polygonal domain with a finite number of vertices with angles $\kappa_j$, there is some $\epsilon > 0$ such that

$$\kappa_j \in [\epsilon, \pi - \epsilon] \cup [\pi + \epsilon, 2\pi] \text{ for all } j.$$
In three dimensions, the set of possible singularities is much greater [4].

Edge singularities correspond to the vertex singularities in two dimensions, but in addition, vertex singularities appear [10].

The effect of smoothing singular boundaries is considered in [6].
Changing boundary condition type

When boundary conditions change type along a straight line, singularity same as slit domain.

Suppose that we have a domain

$$\Omega = \{(x, y) \in \mathbb{R}^2 : x \in [-1, 1], \ y \in [0, 1]\}$$

and we impose homogeneous Dirichlet conditions on

$$\Gamma = \{(x, 0) \in \mathbb{R}^2 : x \in [0, 1]\}$$

and Neumann conditions on $\partial \Omega \setminus \Gamma$.

We can reflect the domain $\Omega$ around the line $y = 0$, and we get the domain $[-1, 1]^2$ with a slit given by $\Gamma$. 
Changing boundary condition type

\[ g_{2\pi} = \sqrt{r} \sin\left(\frac{1}{2} \theta\right) \] satisfies Dirichlet conditions on \( \Gamma \) and Neumann conditions on \( \Gamma^* \).

Such singularities occur any time we switch from Dirichlet to Neumann boundary conditions along a straight boundary segment.

We illustrate this with the following problem:

\[-\Delta u = 0 \text{ in } [0, 1]^2 \]
\[ u = 0 \text{ on } \Gamma, \quad \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega \setminus \Gamma, \]

where \( \Gamma = \{(x, 0) : x \in \left[\frac{1}{2}, 1\right]\} \), cf. Figure 7.

Exercise: explore this problem in more detail.
Suppose that, in general,

$$|\nabla^k u(r)| \approx C|r - r_0|^{-k+\gamma} \text{ for } r \in \Omega,$$  \hspace{1cm} (54)

where $\nabla^k u$ denotes tensor of partial derivatives of order $k$, and $|T|$ denotes the Frobenius norm of a tensor $T$.

For the L-shaped problem, we saw this holds for $k = 1$ and $\gamma = 2/3$.

(54) holds for all $k \geq 1$ and $\gamma = \pi/\kappa$ for boundary vertices with angle $\kappa$.

For simplicity, we assume that $r_0 = 0$ from now on.
Local error estimates

From (36) we have

$$\| u - u_h \|_{H^1(\Omega)} \leq C \inf_{v \in V_h} \| u - v \|_{H^1(\Omega)}.$$  

For a non-uniform mesh,

$$\| u - I_h u \|_{H^1(\Omega)}^2 \leq C \sum_e \left( h_e^{m-1} \| u \|_{H^m(e)} \right)^2,$$  \hspace{1cm} (55)

where summation is over all of elements $e$ of mesh and $h_e =$ size of $e$.

Since we are assuming that the derivatives of the solution degrade in a radial fashion, let us also assume that the mesh is refined in a radial fashion.
For each element $e$, let $r_e$ denote its centroid $r_e$.

Assume there is a monotonic mesh function $\mu$ such that $h_e \approx (1/n)\mu(|r_e|)$, where $n$ is a mesh size parameter.

For example, we will consider $\mu(r) = r^\beta$ for $\beta > 0$.

Let $|e|$ denote the volume of an element $e$.

With such a mesh and under the assumption (54), the error expression (55) takes the form

\[
\begin{align*}
n^{2-2m} \sum_\Omega (\mu(|r_e|)^{m-1} |r_e|^{-m+\gamma} \sqrt{|e|})^2 \\
\approx n^{2-2m} \int_\Omega (\mu(|r|)^{m-1} |r|^{-m+\gamma})^2 \, dr.
\end{align*}
\]
Taking $\mu(r) = r^\beta$, the integrand in (56) simplifies to $|r|^p$ where
\[ p = 2(\beta(m-1) - m + \gamma). \]

Such an expression is integrable in $d$ dimensions if and only if $p > -d$, that is, if
\[ \beta > \frac{m - \gamma - d/2}{m - 1}. \]

If $d = 2$ and $m = 2$ (piecewise linears in 2 D), requirement is $\beta > 1 - \gamma$.

For the L-shaped domain, this means $\beta > \frac{1}{3}$. 
For higher-order approximations, mesh conditions are different.

Other corners of the L-shaped domain can also require mesh refinement.

For these, $\gamma = 2$, and cubics ($m = 4$) requires $\beta > \frac{1}{3}$.

In this case, $\beta > \frac{7}{9}$ is required at the re-entrant corner (for $m = 4$).

When $\gamma = \frac{\pi}{\kappa}$ is sufficiently large (comparable with $m - d/2$), we can take $\beta \approx 0$, meaning a mesh of uniform size.
localized behavior

How to model localized behavior in a physical system:

what is shape of drum head if you push down on it with a sharp pin?

Tempting to model as effect occurring at a single point.

The Laplace equation models to a reasonable extent the deformation of the drum head (for small deformations), so one might consider

\[-\Delta u = 0 \text{ in } \Omega\]

\[u(x_0) = u_0\]

(57)

where \(u_0\) denotes the prescribed position of the pencil.
An ill-posed problem?

However, this problem is not well posed. Difficulty: one cannot constrain a function in $H^1(\Omega)$ at a single point.

This is illustrated by the function [2, Example 1.4.3]

$$v(x) = \log|\log|x||$$  \hspace{1cm} (58)

which satisfies $v \in H^1(B)$ where

$$B = \{x \in \mathbb{R}^2 : |x| < \frac{1}{2}\}.$$  

This function does not have a well-defined point value at the origin.
Thus setting a point value for a function in $H^1$ does not make sense.

By shifting this function around, we realize that functions in $H^1$ may not have point values on a dense set of points $\{x_i\}$:

$$u(x) = \sum_{i=1}^{\infty} 2^{-i} \log |\log |x - x_i||$$  \hspace{1cm} (59)

For example, take the set of points $\{x_i\}$ to be dense in

$$D = \{x \in \mathbb{R}^2 : |x| < \frac{1}{4}\}.$$
It is possible to change to a Dirichlet problem

$$\begin{align*}
-\Delta u &= 0 \text{ in } \Omega \\
u &= u_0 \text{ on } \Gamma
\end{align*}$$

where $\Gamma$ is a small curve representing the point of contact of a knife edge with the drum head, and $u_0$ is some function defined on $\Gamma$.

As long as $\Gamma$ has positive length, this problem is well posed.

However, its behavior will degenerate as the length of $\Gamma$ is decreased.
Another approach to modeling such phenomena is using the Dirac $\delta$-function [2]:

$$-\Delta u = \delta_{x_0} \text{ in } \Omega$$

$$u = 0 \text{ on } \partial\Omega,$$

(61)

where $\delta_{x_0}$ is the linear functional $\delta_{x_0}(v) = v(x_0)$.

Again, there is an issue since this linear functional is not bounded on $V$, as the function $v$ defined in (58) illustrates.

On the other hand, the solution to (61) is known as the Green’s function for the Laplacian on $\Omega$ (with Dirichlet conditions).
Possible to make sense of (61) using sophisticated Sobolev spaces [2].

However, rather than taking that approach, we take one that effectively resolves the issue in conventional spaces.

What we do is replace $\delta_{x_0}$ by a smooth function $\delta^A_{x_0}$ with the property that

$$\int_{\Omega} \delta^A_{x_0}(x) v(x) \, dx \rightarrow v(x_0) \text{ as } A \rightarrow \infty$$

for sufficiently smooth $v$.  

(62)
We then consider the problem

$$\Delta u^A = \delta_{x_0}^A \text{ in } \Omega$$
$$u^A = g \text{ on } \partial \Omega.$$  (63)

Note that we can pick $g$ to be the fundamental solution, and thus we have $u^A \to g$ as $A \to \infty$.

For example, we can choose $\delta_{x_0}^A$ to be Gaussian function of amplitude $A$ and integral 1. In particular, in two dimensions,

$$\delta_{x_0}^A = A e^{-\pi A|x-x_0|^2}.$$
We check our requirement that the integral is 1 via the change of variables $y = \sqrt{\pi A} \, x$:

$$
\int_{\mathbb{R}^2} \pi A \, e^{-\pi A |x-x_0|^2} \, dx = \int_{\mathbb{R}^2} e^{-|y-y_0|^2} \, dy
$$

$$
= 2\pi \int_0^\infty e^{-r^2} \, r \, dr
$$

$$
= \pi \int_0^\infty e^s \, ds = \pi.
$$

In our experiments, $x_0$ was chosen to be near the middle of the square $\Omega = [0, 1]^2$, that is, $x_0 = (0.50001, 0.50002)$ to avoid having the singularity at a grid point.
The fundamental solution for the Laplace equation in two dimensions is

\[ g(x) = -\frac{1}{2\pi} \log |x - x_0|, \]

and so we took as boundary conditions \( g(x) \) for \( x \in \partial\Omega \).

Approximation of singular Green’s function only first order accurate.

Increasing the order of polynomials used is only of modest benefit.

Increasing the amplitude of the approximate \( \delta \)-function useful up to a point, but making it larger only of value if the mesh is refined.
### Table 3: Data for the solution of (63).

Key: degree is the degree of polynomials used, mesh is the number of triangle pairs in each direction. The amplitude is $A$, error is $\|u_h^A - g\|_{L^2(\Omega)}$, check-sum is the value $1 - \int_\Omega (\delta_{x_0}^A)_h \, dx$ where $(\delta_{x_0}^A)_h$ denotes the interpolant of $\delta_{x_0}^A$ in $V_h$. 

<table>
<thead>
<tr>
<th>degree</th>
<th>mesh</th>
<th>amplitude</th>
<th>error</th>
<th>check-sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>128</td>
<td>10,000</td>
<td>5.27e-03</td>
<td>-2.34e-02</td>
</tr>
<tr>
<td>1</td>
<td>256</td>
<td>10,000</td>
<td>2.50e-03</td>
<td>-4.57e-09</td>
</tr>
<tr>
<td>1</td>
<td>512</td>
<td>10,000</td>
<td>1.47e-03</td>
<td>-2.22e-16</td>
</tr>
<tr>
<td>1</td>
<td>1024</td>
<td>10,000</td>
<td>1.08e-03</td>
<td>5.11e-15</td>
</tr>
<tr>
<td>4</td>
<td>256</td>
<td>10,000</td>
<td>9.73e-04</td>
<td>-1.02e-10</td>
</tr>
<tr>
<td>1</td>
<td>512</td>
<td>100,000</td>
<td>9.67e-04</td>
<td>-1.06e-03</td>
</tr>
<tr>
<td>1</td>
<td>1024</td>
<td>100,000</td>
<td>5.24e-04</td>
<td>-1.98e-14</td>
</tr>
</tbody>
</table>
Adding a potential

We now augment the equation (1) with a potential $Z$, which is simply a function defined on $\Omega$ with real values. The PDE takes the form

$$-\Delta u + Zu = f \text{ in } \Omega$$

(64)

together with the boundary conditions (2).

To formulate the variational equivalent of (1) with boundary conditions (2), we again use the variational space

$$V := \{ v \in H^1(\Omega) : v|_{\Gamma} = 0 \}.$$ 

(65)
The appropriate bilinear form for the variational problem is then

\[ a_Z(u, v) = \int_{\Omega} \nabla u(x) \cdot \nabla v(x) + Z(x)u(x)v(x) \, dx. \quad (66) \]

In the case of homogeneous boundary conditions, we seek a solution \( u \in V \) to

\[ a_Z(u, v) = \int_{\Omega} f(x)v(x) \, dx \quad \forall \, v \in V. \quad (67) \]

The simplest case is when \( Z \) is a constant, in which case (64) is often called the **Helmholtz equation**.
The Helmholtz problem becomes interesting if \( Z \) is large, or equivalently, there is a small coefficient in front of \( \Delta \) in (64).

**Exercise:** explore this problem.

To understand coercivity in such problems, we first consider the eigenvalue problem

\[
-\Delta u = \lambda u \text{ in } \Omega
\]  

(68)

together with the boundary conditions (2).

Denote solution of (68) by \( u_\lambda \).
Potential coercivity

Let $\lambda_0$ be the lowest eigenvalue, and $u_{\lambda_0} \in V$ the corresponding eigenvector, for the eigenproblem problem (68):

$$a_0(u_{\lambda}, v) = \lambda \int_{\Omega} u_{\lambda}(x)v(x) \, dx \quad \forall \, v \in V,$$

(69)

where $a_0(\cdot, \cdot)$ denotes the case $Z \equiv 0$, same as bilinear form $a(\cdot, \cdot)$ in (6).

Coercivity (24) of the bilinear form $a_0(\cdot, \cdot)$ shows that $\lambda_0 > 0$.

Moreover, if $Z(x) > -\lambda_0$ for all $x \in \Omega$, then the problem (67) is well posed since it is still coercive.
For certain unbounded potentials, it is still possible to show that (67) is well posed.

For example, if $Z$ is either the Coulombic or gravitational potential $Z(x) = -|x|^{-1}$, then the eigenvalue problem

$$a_Z(u_\lambda, v) = \lambda \int_{\Omega} u_\lambda(x)v(x) \, dx \quad \forall \, v \in V,$$

is well posed, even in the case $\Omega = \mathbb{R}^3$.

In this case, eigensolutions correspond to the wave functions of the hydrogen atom [7].

Exercise: explore this problem.
Figure 8: Asymptotic wavefunction perturbation computed with quartics on a mesh of size 100, with $L = 7$. 
van der Waals interaction

van der Waals interaction energy between two hydrogen atoms, separated by a distance $R$, is asymptotically of the form $-C_6 R^{-6}$ where the constant $C_6$ can be computed [3] by solving a two-dimensional PDE.

Let $\Omega = [0, \infty] \times [0, \infty]$ and consider the PDE

$$
\begin{align*}
-\frac{1}{2} \Delta u(r_1, r_2) + (\kappa(r_1) + \kappa(r_2)) u(r_1, r_2) &= -\frac{1}{\pi} (r_1 r_2)^2 e^{-r_1-r_2} \text{ in } \Omega, \\
\end{align*}
$$

(71)

where the function $\kappa$ is defined by $\kappa(r) = r^{-2} - r^{-1} + \frac{1}{2}$. The minimum of $\kappa$ occurs at $r = 2$, and we have $\kappa(r) \geq \frac{1}{4}$. 
The problem (71) is well posed in $H^1_0(\Omega)$, i.e., given Dirichlet conditions on the boundary of the quarter-plane $\Omega$.

The variational form for (71) is

$$a_\kappa(u, v) = \int_\Omega \frac{1}{2} \nabla u(r_1, r_2) \cdot \nabla v(r_1, r_2) \, dr_1 dr_2$$

$$+ \int_\Omega (\kappa(r_1) + \kappa(r_2)) u(r_1, r_2) v(r_1, r_2) \, dr_1 dr_2,$$

(72)

defined for all $u, v \in H^1_0(\Omega)$. 
The form (72) is coercive on $H^1_0(\Omega)$, since
\[ \kappa(r_1) + \kappa(r_2) \geq \frac{1}{2}. \]
In particular,
\[ a_\kappa(v, v) \geq \frac{1}{2} \int_{\Omega} |\nabla v(r_1, r_2)|^2 + v(r_1, r_2)^2 \, dr_1 dr_2, \quad (73) \]
for all $v \in H^1_0(\Omega)$.

The form $a_\kappa(\cdot, \cdot)$ is continuous on $H^1_0(\Omega)$ because of the Hardy inequality
\[ \int_0^\infty \left( \frac{u(r)}{r} \right)^2 \, dr \leq 4 \int_0^\infty (u'(r))^2 \, dr \quad (74) \]
for $u \in H^1_0(0, \infty)$. 

van der Waals coercivity
Note that it would not be continuous on all of $H^1(0, \infty)$:

without the Dirichlet boundary condition, the form would be infinite for some functions in $H^1(0, \infty)$.

Here is an example of a variational form where

- coercivity is easy to demonstrate
- but continuity is delicate.
To be able to render this problem computationally feasible, we replace $\Omega$ by a square $\Omega_L$ of side $L$ in length; $\Omega_L = [0, L] \times [0, L]$. Define $U(r_1, r_2) = u(Lr_1, Lr_2)$. Then $\Delta U(r_1, r_2) = L^2 \Delta u(Lr_1, Lr_2)$. Thus

$$-\frac{1}{2} L^{-2} \Delta U(r_1, r_2) = -\frac{1}{2} \Delta u(Lr_1, Lr_2)$$

$$= - (\kappa(Lr_1) + \kappa(Lr_2)) u(Lr_1, Lr_2) - \frac{L^4}{\pi} (r_1 r_2)^2 e^{-Lr_1 - Lr_2}$$

$$= - (\hat{\kappa}_L(r_1) + \hat{\kappa}_L(r_2)) U(r_1, r_2) - \frac{L^4}{\pi} (r_1 r_2)^2 e^{-Lr_1 - Lr_2},$$

(75)

where $\hat{\kappa}_L(r) = L^{-2} r^{-2} - L^{-1} r^{-1} + \frac{1}{2}$. 
van der Waals interaction

Therefore $U$ satisfies

$$-rac{1}{2}L^{-2}\Delta U(r_1, r_2) + (\hat{\kappa}_L(r_1) + \hat{\kappa}_L(r_2)) U(r_1, r_2) = -\frac{L^4}{\pi}(r_1 r_2)^2 e^{-Lr_1-Lr_2},$$

(76)

which we can pose with homogeneous Dirichlet boundary conditions ($u = 0$) on $\Omega_1 = [0, 1] \times [0, 1]$.

Multiplying by $2L^2$, we obtain the equation

$$-\Delta U(r_1, r_2) + (\kappa_L(r_1) + \kappa_L(r_2)) U(r_1, r_2) = -\frac{2L^6}{\pi}(r_1 r_2)^2 e^{-Lr_1-Lr_2}$$

$$= f(r_1, r_2),$$

(77)

where $\kappa_L(r) = 2r^{-2} - 2Lr^{-1} + L^2$. 
Thus we introduce the variational form

\[ a_L(u, v) = \int_{[0,1]^2} \nabla u(r_1, r_2) \cdot \nabla v(r_1, r_2) \]

\[ + \left( \kappa_L(r_1) + \kappa_L(r_2) \right) u(r_1, r_2) v(r_1, r_2) \, dr_1 dr_2 \]  

(78)

Variational problem: find \( u_L \in V = H^1_0([0,1]^2) \) such that

\[ a_L(u_L, v) = \int_{[0,1]^2} f(r_1, r_2) v(r_1, r_2) \, dr_1 dr_2 \]  

(79)

for all \( v \in V \).

The solution is shown in Figure 8 with \( L = 7 \) computed on a mesh of size 100 with quartic Lagrange piecewise polynomials.
van der Waals interaction

The main quantity of interest \[3, \text{equation (3.25)}\] is

\[
C_6 = -\frac{32\pi}{3} \int_0^\infty \int_0^\infty r_1^2 r_2^2 e^{-(r_1+r_2)} u(r_1, r_2) \, dr_1 dr_2
\]

\[
\approx -\frac{32\pi}{3} \int_0^L \int_0^L r_1^2 r_2^2 e^{-(r_1+r_2)} u(r_1, r_2) \, dr_1 dr_2
\]

\[
\approx -\frac{32\pi}{3} \int_0^L \int_0^L r_1^2 r_2^2 e^{-(r_1+r_2)} U(r_1/L, r_2/L) \, dr_1 dr_2
\]

\[
= -\frac{32\pi}{3} \int_0^1 \int_0^1 L^4 R_1^2 R_2^2 e^{-(LR_1+LR_2)} U(R_1, R_2) \, L^2 \, dR_1 dR_2
\]

\[
= -\frac{16\pi^2}{3} \frac{2L^6}{\pi} \int_0^1 \int_0^1 R_1^2 R_2^2 e^{-(LR_1+LR_2)} U(R_1, R_2) \, dR_1 dR_2
\]

\[
= \frac{16\pi^2}{3} \int_0^1 \int_0^1 f(R_1, R_2) U(R_1, R_2) \, dR_1 dR_2,
\]

where we made the substitution \(r_i = LR_i, i = 1, 2,\) and \(f\) is defined in \((77)\).
van der Waals interaction

<table>
<thead>
<tr>
<th>degree</th>
<th>quadrature</th>
<th>mesh no.</th>
<th>$C_6$ error</th>
<th>$\epsilon$</th>
<th>$L$</th>
<th>time</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>6</td>
<td>100</td>
<td>4.57e-07</td>
<td>1.00e-09</td>
<td>15.0</td>
<td>1.47</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>100</td>
<td>4.57e-07</td>
<td>1.00e-09</td>
<td>15.0</td>
<td>1.595</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>250</td>
<td>4.56e-07</td>
<td>1.00e-09</td>
<td>15.0</td>
<td>12.5</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>600</td>
<td>4.45e-07</td>
<td>1.00e-09</td>
<td>15.0</td>
<td>4.738</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>250</td>
<td>-1.22e-07</td>
<td>1.00e-09</td>
<td>15.0</td>
<td>0.786</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>255</td>
<td>-2.74e-08</td>
<td>1.00e-09</td>
<td>15.0</td>
<td>0.786</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>265</td>
<td>4.23e-08</td>
<td>1.00e-09</td>
<td>15.0</td>
<td>0.837</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>240</td>
<td>-1.87e-08</td>
<td>1.00e-09</td>
<td>15.0</td>
<td>0.739</td>
</tr>
</tbody>
</table>

Table 4: Using finite element computation of $C_6 = 6.4990267054$ [3]. The potential was modified as in (81). Computations were done with 4 cores via MPI and a PETSc Krylov solver. Error values were the same for $\epsilon = 10^{-9}$ and $\epsilon = 10^{-12}$. 
To avoid singularities in the coefficients, we modified the potential to be

\[
\kappa_\ell^\varepsilon(r) = 2(\varepsilon + r)^{-2} - 2L(\varepsilon + r)^{-1} + L^2.
\]  

(81)

Computational results are shown in Table 4. The results were insensitive to \(\varepsilon\) for \(\varepsilon \leq 10^{-9}\). The singularity \((r_i)^{-2}\) is difficult to deal with. But we can integrate by parts to soften its effect, as follows:

\[
\int_\Omega (r_i)^{-2} u(r_1, r_2) v(r_1, r_2) \, dr_1 dr_2 = -\int_\Omega \left( \frac{\partial}{\partial r_i} (r_i)^{-1} \right) u(r_1, r_2) v(r_1, r_2) \, dr_1 dr_2
\]

\[
= \int_\Omega (r_i)^{-1} \frac{\partial}{\partial r_i} (u(r_1, r_2) v(r_1, r_2)) \, dr_1 dr_2,
\]

(82)

where for simplicity we define \(\Omega = [0, 1]^2\) here and for the remainder of this subsection.
Another formulation

We have assumed that $u, v \in V = H^1_0(\Omega)$ in (82). Thus

$$
\int_\Omega ((r_1)^{-2} + (r_2)^{-2}) u(r_1, r_2) v(r_1, r_2) \, dr_1 dr_2
$$

$$
= \int_\Omega ((r_1)^{-1}, (r_2)^{-1}) \cdot \nabla (u(r_1, r_2) v(r_1, r_2)) \, dr_1 dr_2
$$

$$
= \int_\Omega ((r_1)^{-1}, (r_2)^{-1}) \cdot \left( (\nabla u(r_1, r_2)) v(r_1, r_2) + (\nabla v(r_1, r_2)) u(r_1, r_2) \right) \, dr_1 dr_2
$$

Thus we introduce a new variational form (cf. (78))

$$
\hat{a}_L^\epsilon(u, v) = \int_\Omega \nabla u(r_1, r_2) \cdot \nabla v(r_1, r_2) + (\hat{\kappa}_L^\epsilon(r_1) + \hat{\kappa}_L^\epsilon(r_2)) u(r_1, r_2) v(r_1, r_2) \, dr_1 dr_2
$$

$$
+ 2 \int_\Omega \hat{\beta}(r_1, r_2) \cdot (\nabla u(r_1, r_2)) v(r_1, r_2) + u(r_1, r_2)(\nabla v(r_1, r_2)) \, dr_1 dr_2
$$

where

$$
\hat{\beta}(r_1, r_2) = ((r_1 + \epsilon)^{-1}, (r_2 + \epsilon)^{-1}), \quad \hat{\kappa}_L^\epsilon(r) = -2L(r + \epsilon)^{-1} + 2L^2.
$$
Extended exercise: Let $\epsilon > 0$. Consider the problem

$$-\epsilon \Delta u_\epsilon + u_\epsilon = f \text{ in } \Omega$$

together with the boundary conditions (2), where $f$ is held fixed independent of $\epsilon$.

Give conditions for which you would expect $u_\epsilon \to f$ as $\epsilon \to 0$.

Do a simple example in which $f$ does not satisfy the boundary conditions (2) and see what happens for small $\epsilon$.

Does $u_\epsilon \to f$ except in a small boundary layer?
Boundary layers continued

Compare your results with those in Table 5 which corresponds to the choices $f \equiv 1$ and $\epsilon = 10^{-6}$.

Best results require a large number of nodes to resolve the boundary layer, but among the different choices (linear, quadratics, and so forth), the results are about the same and take about the same time to compute.

High-order polynomial does not provide particular benefit in this case.
Table 5: Boundary layer problem with $\epsilon = 10^{-6}$. Degree refers to the polynomial degree, mesh number indicates the number of edges along each boundary side, $L^2$ difference is $\|u - f\|_{L^2([0,1]^2)}$, and time is in seconds.
References


