Let \((H, \langle \cdot, \cdot \rangle_H)\) be a real Hilbert space and \(B : H \times H \to \mathbb{R}\) is a bilinear functional with the properties:

1) \(|B(u, v)| \leq c_1 \|u\|_H \|v\|_H \quad \forall u, v \in H, \quad [\text{Bounded}]\)

2) \(B(u, u) \geq c_2 \|u\|_H^2 \quad \forall u \in H. \quad [\text{Coercive}]\)

So far, all the weak formulations of our elliptic equations have been of the form:

Find \(u \in H\) such that

\[ B(u, \varphi) = \varepsilon(\varphi) \quad \forall \varphi \in H, \]

where \(\varepsilon \in H^*\) (and \(H = H^1(\Omega)\) or \(H_0^1(\Omega)\) depending on the boundary conditions at hand).

The Galerkin method

In order to obtain a computational approximation \(u_h \approx u\), we replace the infinite dimensional space \(H\) by a sequence of finite dimensional subspaces

\[ \{H_h\}_{0<h \leq h_{\text{max}}} \]

where \(\dim H_h \to \infty\) as \(h \to 0\) and

\[ \lim_{h \to 0} H_h^\# = H. \]
Next, we replace the weak formulation of our elliptic problem by its galerkin discretization:

\[ \text{Find } u_h \in H_h \text{ such that } \quad B(u_h, \psi) = l(\psi) \quad \forall \psi \in H_h. \]

Lemma: For every \( l \in H^* \), there exists a unique galerkin approximation \( u_h \in H_h \).

Proof: \( (H_h, \langle \cdot , \cdot \rangle_H) \) is a Hilbert space (\( H_h \) is finite dimensional \( \Rightarrow \) \( H_h \) is closed \( \Rightarrow \) \( H_h \) is complete). and \( l \big|_{H_h} \in H^*_h \). The proof then follows by Lax–Milgram's Lemma \( [F3, \text{p. 9}] \).

How do we compute \( u_h \)?

Let \( \{ \phi_i \}_{i = 0}^{M_h} \) be a basis in \( H_h \), then

\[ u_h = \sum_{i = 0}^{M_h} c_i \phi_i \quad \text{and} \quad \psi = \sum_{j = 0}^{M_h} a_j \phi_j \]

which implies that

\[ \rightarrow \]
\[ B(u_h, \varphi) = \ell(\varphi) \quad \forall \varphi \in H_h \implies \]
\[
\sum_{j=0}^{M_h} a_j B(u_h, \phi_j) = \sum_{j=0}^{M_h} a_j \ell(\phi_j) \quad \forall a_j \in \mathbb{R} \implies \]
\[ B(u_h, \phi_j) = \ell(\phi_j) \quad \forall j \in [0, \ldots, M_h] \implies \]
\[
\sum_{i=0}^{M_h} c_i B(\phi_i, \phi_j) = \ell(\phi_j) \quad \text{for} \quad \implies \]

\[ K c = b, \quad \text{where} \quad \begin{cases} 
k = B(\phi_i, \phi_j) & i, j = 0, \\
b = [\ell(\phi_0), \ldots, \ell(\phi_{M_h})]^T, \\
c = [c_0, \ldots, c_{M_h}]^T. \end{cases} \]

Remark. \( K \) is referred to as the **stiffness matrix**.

When can we expect convergence?

Theorem (Céa's Lemma)

\[ \|u - u_h\|_H \leq \frac{C_1}{C_2} \inf_{\varphi \in H_h} \|u - \varphi\|_H. \]

Remark. If the subspaces \( \{H_h\} \) are chosen in such a way that
\[
\lim_{h \to 0} \inf_{\varphi \in H_h} \|u - \varphi\|_H = 0,
\]
then \( u_h \) converges to \( u \) in \( H \) as \( h \) tends to zero.
Proof (of Céa's Lemma)

\[ B(u, \varphi) = \ell(\varphi) \quad \forall \varphi \in H \quad \text{and} \quad B(u_h, \varphi) = \ell(\varphi) \quad \forall \varphi \in H_h \]

\[ \Rightarrow \quad B(u - u_h, \varphi) = 0 \quad \forall \varphi \in H_h \quad (\text{but not for all } \varphi \in H). \]

\[ c_2 \| u - u_h \|^2_H \leq B(u - u_h, u - u_h + \varphi - \varphi) \]

\[ = B(u - u_h, u - \varphi) + B(u - u_h, \varphi - u_h) \]

\[ \leq c_1 \| u - u_h \|_H \| u - \varphi \|_H \quad \forall \varphi \in H_h. \]

\[ \Rightarrow \]

\[ \| u - u_h \|_H \leq \frac{c_1}{c_2} \| u - \varphi \|_H \quad \forall \varphi \in H_h \]

\[ \Rightarrow \]

\[ \| u - u_h \|_H \leq \frac{c_1}{c_2} \inf_{\varphi \in H_h} \| u - \varphi \|_H. \quad \Box \]

Exercise

Let \( B(u, \varphi) := (u, \varphi)_H \). Prove that the resulting Galerkin approximation \( u_h \) is the best possible approximation in \( H_h \), i.e.,

\[ \| u - u_h \|_H = \inf_{\varphi \in H_h} \| u - \varphi \|_H. \]
Error estimates: \( \| u - u_h \|_H \leq \text{const.} \, h^r \) \hfill (5)

As a direct consequence of Céa's Lemma we have that:

**Corollary** - If there exists a \( X_h \in H_h \) such that

\[ \| u - X_h \|_H \leq C h^r, \]

then \( \| u - u_h \|_H \leq C h^r \).

**Proof** -\[ \| u - u_h \|_H \leq \frac{c_1}{c_2} \inf_{\varphi \in H_h} \| u - \varphi \|_H \leq \frac{c_1}{c_2} \| u - X_h \|_H \leq C h^r. \]

**Comments**

* The task of deriving error estimates for the Galerkin discretization is reduced to finding any approximation \( X_h \in H_h \) of the exact solution \( u \) such that

\[ \| u - X_h \|_H \leq C h^r. \]

* From Céa's Lemma the "best" error estimate would be obtained if

\[ \| u - X_h \|_H = \inf_{\varphi \in H_h} \| u - \varphi \|_H. \]

i.e., \( X_h := P_h u \) where \( P_h \) is the orthogonal projection of \( H \) onto \( H_h \).

**Problem** \( P_h \) is often difficult to work with!

**Alternative choice** of \( X_h \) is needed!
How do we choose \( H_n, \{ \phi_i \} \text{ and } X_h \)?

Let's start with the 1D-case, i.e., \( \Omega = (a,b) \subset \mathbb{R} \).

We would like to have the following:

1) \( H_n \) is a subspace of \( H^1(\Omega) \) or \( H_0^1(\Omega) \).

2) \( \{ \phi_i \} \) gives rise to a linear system \( Kc = b \)
   which is cheap to solve, e.g., if \( K \) becomes a sparse matrix.

3) \( u_h(x_j) = c_j \)

4) An approximation \( X_h \) which is "easy to analyze"
   and which satisfies the inequality \( \| u - X_h \|_{H^1(\Omega)} \leq C h \).

The finite element spaces

\[
\begin{align*}
\Omega &= [x_0, x_1, \ldots, x_{i-1}, x_i, \ldots, x_{M_h} = b] \\
\phi_i &= \max h_i. \\
S_h &= \left\{ u \in C(\Omega) : u(x) = s_i x + t_i \text{ for } x \in [x_{i-1}, x_i] \right\} \subset H^1(\Omega) \\
S_{ho} &= \left\{ u \in S_h : u = 0 \text{ on } \partial \Omega \right\} \subset H_0^1(\Omega).
\end{align*}
\]

Remark
\* \( \dim S_h = M_h + 1 \).
\* \( \dim S_{ho} = M_h - 1 \).
The hat functions \( \{ \phi_i \}_{i=0}^{M_h} \) is a basis for \( S_h \), where

![Diagram showing hat functions \( \phi_i \) for \( i = 0, 1, ..., M_h \).]

**Comments**

* \( S_h \subset H^1(\mathbb{R}) \) as \( \phi_i' = \frac{1}{h} \text{rect}(x - x_i) \in L^2(\mathbb{R}) \)

* \( K_{ij} = B(\phi_i, \phi_j) = 0 \) for \( i < j - 1 \) and \( i > j + 1 \)

\( \Rightarrow \) \( K \) is tridiagonal, i.e., sparse, in the 1D case.

* \( u_h(x_j) = \sum_{i=0}^{M_h} c_i \phi_i(x_j) = c_j \).

**Exercise**

Consider the Dirichlet problem

\[
\begin{cases}
-\Delta u + u &= f \text{ in } (0,1) \\
u(0) &= u(1) = 0
\end{cases}
\]

and its finite element discretization on an equidistant grid, i.e., \( h_i = h \) for all \( i \).

1) Compute the related stiffness matrix \( K \).

2) Let \( f(x) = 2x \). Compute the vector \( b \) and solve the resulting linear system \( Kc = b \).

3) Have you seen this numerical scheme before?
**Def.** \( J_h : C(\Omega) \to S_h \) is the piecewise linear interpolation operator, i.e.,
\[
J_h u := \sum_{i=0}^{M_h} u(x_i) \phi_i.
\]

**Comments**
* We will choose \( \chi_h := J_h u \) for our analysis.

* \( J_h u \) is well defined for \( u \in H^s(\Omega) \) when \( \Omega \subset \mathbb{R}^d \) and

<table>
<thead>
<tr>
<th>( d )</th>
<th>( s )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

This means that we need more regularity than \( u \in H^2(\Omega) \) for the 2D and 3D-cases. Remedy: Consider \( f \in L^2(\Omega) \) instead of \( u \in H^2(\Omega) \).

**Lemma**
For \( \Omega \subset \mathbb{R}^d \) and \( u \in H^2(\Omega) \) we have that

1. \( \| (u - J_h u)' \|_{L^2(\Omega)} \leq \text{Const.} \cdot h \| u'' \|_{L^2(\Omega)} \),
2. \( \| u - J_h u \|_{L^2(\Omega)} \leq \text{Const.} \cdot h^2 \| u'' \|_{L^2(\Omega)} \).

**Comments:**
* The lemma implies that
\[
\| u - J_h u \|_{H^1(\Omega)} \leq \text{Const.} \cdot (1 + h_{\text{max}}^2)^{1/2} \cdot h \| u'' \|_{L^2(\Omega)} \\
\leq \text{Const.} \cdot h \| u \|_{H^2(\Omega)}.
\]
* As a direct consequence of the lemma and our corollary \([F4 \, p5]\) we obtain that

\[
\| u - u_h \|_{H^2(\Omega)} \leq \text{const.} \| u \|_{H^2(\Omega)},
\]

i.e., the Galerkin discretization is first order convergent for the 1D-case of all the elliptic problems considered so far!

* The extra regularity assumption \( u \in H^3(\Omega) \) is not always fulfilled, e.g., \[
\begin{aligned}
-\Delta u &= \delta \text{ on } (-1,1) \\
\xi(0) &= u(1) = 0
\end{aligned}
\]

has the solution

\[
\begin{aligned}
\xi &= 0 \\
\end{aligned}
\in H^1_0(\Omega) \setminus H^2(\Omega)
\]
Proof (Lemma [F4, p.8])

Let \( u \in \mathcal{C}^2(\overline{a}) \).

1) By the mean value theorem there exists a \( \xi \in (x_{i-1}, x_i) \) such that

\[
    u'(\xi) = \frac{u(x_i) - u(x_{i-1})}{x_i - x_{i-1}} = (J_u)'(x) \quad \forall x \in (x_{i-1}, x_i)
\]

\[
\Rightarrow \quad u'(x) - (J_u)'(x) = u'(\xi) = \int_{x_{i-1}}^{x} u''(s) \, ds \quad \forall -11
\]

\[
\Rightarrow \quad \int_{x_{i-1}}^{x_i} \left( u'(x) - (J_u)'(x) \right)^2 \, dx = \int_{x_{i-1}}^{x_i} \int_{x_{i-1}}^{x} \left( \int_{x_{i-1}}^{x} u''(s) \, ds \right)^2 \, dx
\]

Cauchy–Schwarz on \( \mathcal{C}(\mathbb{R}) \): \( (\int_{x_{i-1}}^{x} v \, ds)^2 \leq \int_{x_{i-1}}^{x} 1 \, ds \cdot \int_{x_{i-1}}^{x} v^2 \, ds \).

\[
\Rightarrow \quad \int_{x_{i-1}}^{x_i} \left( \int_{x_{i-1}}^{x} 1 \, ds \cdot \int_{x_{i-1}}^{x} u''(s)^2 \, ds \right) \, dx
\]

\[
\int_{x_{i-1}}^{x} u''(s)^2 \, ds \leq \int_{x_{i-1}}^{x_i} u''(s)^2 \, ds \quad \text{independent of } x!
\]

\[
\leq \int_{x_{i-1}}^{x_i} 1 \, ds \int_{x_{i-1}}^{x} u''(s)^2 \, ds \leq \int_{x_{i-1}}^{x_i} 1 \, x_1 \, dx \cdot \int_{x_{i-1}}^{x_i} u''(s)^2 \, ds
\]

\[
\leq C \, h_i \int_{x_{i-1}}^{x_i} u''(s)^2 \, ds.
\]
\[ \| (u - J_n u)' \|_{L^2(\mathbb{R})}^2 \geq \sum_{i=1}^{M_n} \int_{x_{i-1}}^{x_i} \left( u'(x) - (J_n u)'(x) \right)^2 \, dx \]

\[ \leq \sum_{i=1}^{M_n} Ch_i^2 \int_{x_{i-1}}^{x_i} u''(x)^2 \, dx \leq Ch_i^2 \| u'' \|_{L^2(\mathbb{R})}^2 \]

which proves assertion (2) of the lemma for \( u \in C^2(\mathbb{R}) \).

2) \( u(x) - (J_n u)(x) = \int_{x_{i-1}}^{x} u'(s) - (J_n u)'(s) \, ds \)

\[ \leq \left\| \int_{x_{i-1}}^{x} (x - x_{i-1}) \int_{x_{i-1}}^{x} (u'(s) - (J_n u)'(s))^2 \, ds \, dx \right\| \]

\[ \leq Ch_i \int_{x_{i-1}}^{x_i} u''(s) \, ds \]

\[ \| u - J_n u \|_{L^2(\mathbb{R})}^2 = \sum_{i=1}^{M_n} \int_{x_{i-1}}^{x_i} (u(x) - (J_n u)(x))^2 \, dx \]

\[ \leq C \| u'' \|_{L^2(\mathbb{R})}^2 \]

which proves the assertion (2) for \( u \in C^2(\mathbb{R}) \).
3) What remains is to replace $u \in C^2(\mathbb{R})$ by $u_m \in C^2(\mathbb{R})$ such that $u_m \rightarrow u \in H^2(\mathbb{R})$ and taking the limits in the bounds of assertions 1) and 2). □

**Remark** In order to conduct step 3) of the proof one needs to observe that $J_{h_m} u_m \rightarrow J_h u$ and $(J_{h_m} u_m)' \rightarrow (J_h u)'$ in $L^2(\mathbb{R})$. This holds by the bounds limits:

\[ \| J_{h_m} u_m - J_h u \|^2_{L^2(\mathbb{R})} = \sum_{i=1}^{N_h} \int_{x_{i-1}}^{x_i} \left( \int_{x_{i-1}}^{x_i} \frac{u_{m}(x_{i-1}) - u(x_{i-1}) + u(x_{i+1}) - u(x_{i+1})}{x_i - x_{i-1}} \right)^2 dx \]

\[ \rightarrow 0 \text{ as } m \rightarrow \infty \] (because $\| u_m - u \|_{C(\mathbb{R})} \leq (\| u_m - u \|_{H^2(\mathbb{R})})$).

\[ \| (J_{h_m} u_m)' - (J_h u)' \|^2_{L^2(\mathbb{R})} = \sum_{i=1}^{N_h} \int_{x_{i-1}}^{x_i} \left( \frac{u_{m}(x_i) - u(x_i) + u(x_i) - u(x_{i-1})}{x_i - x_{i-1}} \right)^2 dx \]

\[ \rightarrow 0 \text{ as } m \rightarrow \infty. \]