operator from Theorem 4.2(ii). We estimate

$$
\begin{aligned}
\|\partial_n v\|_{H^{-1/2}(\Gamma)} &= \sup_{w\in H^{1/2}(\Gamma)\setminus\{0\}} \frac{\langle \partial_n v, w\rangle_\Gamma}{\|w\|_{H^{1/2}(\Gamma)}} \le C \sup_{w\in H^{1/2}(\Gamma)\setminus\{0\}} \frac{\langle \partial_n v, w\rangle_\Gamma}{\|\mathcal{E}w\|_{H^1(\Omega)}} \\
&= C \sup_{w\in H^{1/2}(\Gamma)\setminus\{0\}} \frac{\int_\Omega \nabla v\cdot\nabla\mathcal{E}w + \int_\Omega \Delta v\,\mathcal{E}w}{\|\mathcal{E}w\|_{H^1(\Omega)}} \le C \sup_{W\in H^1(\Omega)\setminus\{0\}} \frac{\int_\Omega \nabla v\cdot\nabla W + \int_\Omega \Delta v\, W}{\|W\|_{H^1(\Omega)}} \\
&\le C \sup_{W\in H^1(\Omega)\setminus\{0\}} \frac{\|v\|_{H^1(\Omega)}\,\|W\|_{H^1(\Omega)} + \|\Delta v\|_{\tilde H^{-1}(\Omega)}\,\|W\|_{H^1(\Omega)}}{\|W\|_{H^1(\Omega)}} \\
&= C\left(\|v\|_{H^1(\Omega)} + \|\Delta v\|_{\tilde H^{-1}(\Omega)}\right).
\end{aligned}
$$

□

Remark 4.3 $\Delta v \in L_2(\Omega)$ *implies* $\Delta v \in \tilde H^{-1}(\Omega)$.

5 Finite element error analysis for elliptic problems

In this section we deal with the error analysis of the finite element method. Key steps in the error analysis are the Lax-Milgram lemma (Theorem 2.1), which proves the unique existence of u_h and its stability, and Céa's lemma (Theorem 3.2) proving

$$
\|u - u_h\| \le \frac{C_a}{\alpha}\|u - v\| \quad \forall v \in V_h.
$$

Here, several assumptions are needed, in particular the boundedness of a (with bound C_a) and its V-ellipticity (with ellipticity constant α). Therefore, to bound the error in the energy norm (or the norm of V) we only need to select an appropriate function $v \in V_h$ for which we are able to further estimate $\|u - v\|$. If V_h consists of continuous, piecewise linear functions then a standard candidate is the piecewise linear interpolant $I_h u \in V_h$ (defined below). First, in §5.1, we deal with approximation theory in a more general and abstract form. Then, in §5.2, we apply the approximation results to the finite element method.

5.1 Approximation theory

Definition 5.1 *Let* $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ *be normed linear spaces, and* $A \in \mathcal{L}(X, Y)$, *where* $\mathcal{L}(X, Y)$ *denotes the space of bounded linear operators* $X \to Y$. *Then,* A *is* compact *if and only if* $(Ax_n)_{n\in\mathbb{N}} \subset Y$ *has a convergent subsequence for any bounded sequence* $(x_n)_{n\in\mathbb{N}} \subset X$.

This can be equivalently formulated as: A is compact if and only if every bounded subset of X is mapped to a relatively compact subset of Y.

Proposition 5.1 (Rellich's embedding theorem) *Let* Ω *be a Lipschitz domain. Then for any* $t > s$, *the injection* $i\colon H^t(\Omega) \to H^s(\Omega)$ *is compact.*

Proposition 5.2 (Sobolev's embedding theorem) *Let Ω be a Lipschitz domain in $\mathbf{R}^n$. Then, the injection $i\colon H^{n/2+\varepsilon}(\Omega) \to C^0(\bar{\Omega})$ is continuous for all $\varepsilon > 0$, that is,*

$$\sup_{x\in\Omega} |u(x)| \le C_\varepsilon \|u\|_{H^{n/2+\varepsilon}(\Omega)} \quad \textit{for all } u \in H^{n/2+\varepsilon}(\Omega).$$

Remark 5.1 *A simple argument for the influence of the dimension n is the following: Using the Fourier transform, we see for $u \in C_0^\infty(\mathbf{R}^n)$ that*

$$\begin{aligned} |u(x)| &\le \int_{\mathrm{R}^n} |\hat{u}(\xi)|\, d\xi = \int_{\mathrm{R}^n} (1+|\xi|^2)^{-s/2}(1+|\xi|)^{s/2}|\hat{u}(\xi)|\, d\xi \\ &\le \left(\int_{\mathrm{R}^n} (1+|\xi|^2)^{-s}\, d\xi\right)^{1/2} \|u\|_{H^s(\mathrm{R}^n)}, \end{aligned}$$

where the last inequality follows from the Cauchy-Schwarz inequality. Thus, $\int_{\mathrm{R}^n}(1+|\xi|^2)^{-s}\, d\xi < \infty$ will be sufficient. As the integrand is bounded on every bounded set, we only need to study the behaviour as $|\xi| \to \infty$. Transforming to polar coordinates and choosing some $r^* > 0$,

$$\int_{\mathrm{R}^n} (1+|\xi|^2)^{-s}\, d\xi \sim \int_{r^*}^{\infty} r^{-2s} r^{n-1}\, dr = \int_{r^*}^{\infty} r^{n-2s-1}\, dr.$$

The last integral is finite if and only if $n-2s-1 < -1$. This corresponds exactly to the condition $s > \frac{n}{2}$.

Lemma 5.1 *Let $\Omega \subset \mathbf{R}^2$ be a Lipschitz domain, $t \ge 2$ integer, $s = \frac{t(t+1)}{2}$, and $\{z_1, z_2, \ldots, z_s\} \subset \Omega$ be given points such that the interpolation operator $I\colon H^t(\Omega) \to P_{t-1}$ is well-defined. Here, P_{t-1} are the polynomials of degree up to $t-1$. Then, there exists $C \ge 0$ such that*

$$\|u - Iu\|_{H^t(\Omega)} \le C|u|_{H^t(\Omega)} \quad \textit{for all } u \in H^t(\Omega).$$

Proof. We first prove that $\|v\|_{H^t(\Omega)}$ and $|||v||| := |v|_{H^t(\Omega)} + \sum_{i=1}^s |v(z_i)|$ are equivalent norms. Then it follows that

$$\begin{aligned} \|u - Iu\|_{H^t(\Omega)} &\le C\, |||u - Iu||| = C\left(|u - Iu|_{H^t(\Omega)} + \sum_{i=1}^s |u(z_i) - Iu(z_i)|\right) \\ &= C|u - Iu|_{H^t(\Omega)} = C|u|_{H^t(\Omega)}, \end{aligned}$$

since the tth derivatives of $Iu \in P_{t-1}$ vanish.

1. As $t \ge 2$ we see that the embedding $H^t(\Omega) \to H^2(\Omega)$ is continuous and by Proposition 5.2 we have that $H^2(\Omega) \to C^0(\bar{\Omega})$ is continuous. Therefore, the injection $H^t(\Omega) \to C^0(\bar{\Omega})$ is continuous. Thus, $|v(z_i)| \le C\|v\|_{H^t(\Omega)}$, $i = 1, \ldots, s$, and

$$|||v||| = |v|_{H^t(\Omega)} + \sum_{i=1}^s |v(z_i)| \le (1 + sC)\|v\|_{H^t(\Omega)} \quad \text{for all } v \in H^t(\Omega).$$

2. Assume that $\|v\|_{H^t(\Omega)} \leq C\,|||v|||$ for every $v \in H^t(\Omega)$ is false for any constant $C > 0$. Then, there exists a sequence $(v_n)_{n\in\mathbb{N}} \subset H^t(\Omega)$ such that $\|v_n\|_{H^t(\Omega)} = 1$ and $|||v_n||| \leq \frac{1}{n}$ for all $n \in \mathbb{N}$. We see that (v_n) is bounded in $H^t(\Omega)$, so by Proposition 5.1 there is a subsequence of (v_n) which converges in $H^{t-1}(\Omega)$. We assume without loss of generality that this subsequence is (v_n). In particular, it follows that (v_n) is a Cauchy sequence in $H^{t-1}(\Omega)$ and thus, since $|v_n|_{H^t(\Omega)} \leq |||v_n||| \to 0$ for $n \to \infty$,

$$\|v_k - v_l\|^2_{H^t(\Omega)} \leq \|v_k - v_l\|^2_{H^{t-1}(\Omega)} + (|v_k|_{H^t(\Omega)} + |v_l|_{H^t(\Omega)})^2 \to 0 \quad \text{for } k, l \to \infty.$$

Therefore, (v_n) is a Cauchy sequence in $H^t(\Omega)$ and by completeness there exists $v^* \in H^t(\Omega)$ such that $v_n \to v^*$ in $H^t(\Omega)$ for $n \to \infty$. By the continuity of the norms it follows from $\|v_n\|_{H^t(\Omega)} = 1$ that $\|v^*\|_{H^t(\Omega)} = 1$, and from $|||v_n||| \leq \frac{1}{n}$ that $|||v^*||| = 0$ since, by the first part,

$$|||v^*||| \leq |||v^* - v_n||| + |||v_n||| \leq C\|v^* - v_n\|_{H^t(\Omega)} + |||v_n||| \to 0 \quad \text{as } n \to \infty.$$

By definition of $|||\cdot|||$ it follows that $|v^*|_{H^t(\Omega)} = 0$, that is, $v^* \in P_{t-1}$, and $|v^*(z_i)| = 0$, $i = 1, \dots, s$. Therefore, v^* vanishes at $\frac{t(t+1)}{2}$ distinct points. It follows that $v^* = 0$ and this is a contradiction to $\|v^*\|_{H^t(\Omega)} = 1$.

Therefore, there exists a constant C such that $\|v\|_{H^t(\Omega)} \leq C\,|||v|||$.

□

Theorem 5.1 (Bramble-Hilbert Lemma) *Let $\Omega \subset \mathbb{R}^2$ be a Lipschitz domain, and $t \geq 2$ integer. For a normed linear space Y let $L \in \mathcal{L}(H^t(\Omega), Y)$.*

If $P_{t-1} \subset \ker L$ then there exists a constant $C \geq 0$ such that

$$\|Lv\|_Y \leq C|v|_{H^t(\Omega)} \quad \textit{for all } v \in H^t(\Omega).$$

Proof. As L is bounded and linear there exists $D \geq 0$ such that $\|Lv\|_Y \leq D\|v\|_{H^t(\Omega)}$. Let $I : H^t(\Omega) \to P_{t-1}$ be an interpolation operator as in Lemma 5.1. Then, $Iv \in P_{t-1} \subset \ker L$ for all $v \in H^t(\Omega)$ and

$$\|Lv\|_Y = \|L(v - Iv)\|_Y \leq D\|v - Iv\|_{H^t(\Omega)} \leq CD|v|_{H^t(\Omega)}$$

by Lemma 5.1. □

5.2 Finite element error estimate for elliptic problems

We deal with the case $V = H^1(\Omega)$ and $V_h = \{v \in V : \ v|_K \in P_{t-1}(K) \ \forall K \in \mathcal{T}_h\}$ where $\mathcal{T}_h = \{K\}$ is a triangulation of Ω, which is assumed to be polygonal (so that it can be discretised by triangular meshes). Here, $P_{t-1}(K)$ denotes the space of polynomials of degree $t-1$ on K. The mesh needs to satisfy certain conditions. We define

$$\begin{aligned} h_k &= \text{diameter of } K = \text{length of longest side of } K, \\ \rho_k &= \text{diameter of the largest circle in } K, \\ h &= \max_{K \in \mathcal{T}_h} h_k \end{aligned}$$

and require that there exists $\beta > 0$ which is independent of h such that

$$\frac{\rho_k}{h_k} \geq \beta \qquad \forall K \in \mathcal{T}_h. \tag{5.1}$$

This means that the elements $K \in \mathcal{T}_h$ are not too thin, i.e. the interior angles of K are not too small (they are bounded from below by a positive constant). One also says that the elements of $\mathcal{T}_h$, or $\mathcal{T}_h$, are *shape regular.* Since we are interested in a sequence of meshes $\{\mathcal{T}_h\}$ such that we can study the behaviour of the finite element error $\|u - u_h\|$ for a sequence of mesh sizes $\{h\}$, the constant β in (5.1) must be independent of h.

We now apply the Bramble-Hilbert Lemma to prove a piecewise polynomial approximation result.

Theorem 5.2 *For a Lipschitz domain $\Omega \subset \mathbf{R}^2$ with polygonal boundary and a given integer $t \geq 2$ let $\{T : \ T \in \mathcal{T}_h\}$ be a shape regular triangulation of Ω.*

Then, for a piecewise polynomial interpolation operator I_h of degree $t-1$ (piecewise with respect to $\mathcal{T}_h$) there holds

$$\left(\sum_{T \in \mathcal{T}_h} \|u - I_h u\|^2_{H^m(T)} \right)^{1/2} \leq C h^{t-m} |u|_{H^t(\Omega)} \quad \textit{for all } u \in H^t(\Omega) \textit{ and all } 0 \leq m \leq t.$$

Here, the constant C is independent of h and u.

Proof. The idea of the proof is to transform to the reference element $\hat{T}$, make a transition from H^m to H^t, and transform back. The transformations give the required powers of h since the Bramble-Hilbert Lemma gives the transition to a semi-norm on $\hat{T}$.

By the assumption of shape regularity it is enough to consider the case that T_h is congruent to $\hat{T}$. Then we can assume without loss of generality that $T_h = h\hat{T} := \{(x_1, x_2) : \ 0 \leq x_1, x_2 \leq h, x_1 + x_2 \leq h\}$. For $v \in H^t(T_h)$ we define $\hat{v} \in H^t(T)$ by $\hat{v}(x_1, x_2) := v(hx_1, hx_2)$. For a multi-index $\alpha = (\alpha_1, \alpha_2)$ of order $|\alpha| = \alpha_1 + \alpha_2$ with non-negative integers α_1, α_2 let D^α denote the partial derivative operator defined by

$$D^\alpha v(x_1, x_2) := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2}} v(x_1, x_2).$$

We see that $D^\alpha v = h^{-\alpha} D^\alpha \hat{v}$ for all multi-indices α with $|\alpha| \le t$. Thus,

$$|v|^2_{H^l(T_h)} = \sum_{|\alpha|=l} \int_{T_h} (D^\alpha v)^2\, dx = \sum_{|\alpha|=l} h^{-2l} h^2 \int_{\hat{T}} (D^\alpha \hat{v})^2\, d\xi = h^{2-2l} |\hat{v}|^2_{H^l(\hat{T})}.$$

Since the transform $\hat{I}$ of the interpolation operator I_h is again an interpolation operator, we can transform to the reference element, apply the Bramble-Hilbert Lemma and transform back to obtain

$$\begin{aligned} \|v - I_h v\|^2_{H^m(T_h)} &\le Ch^{2-2m} \|\hat{v} - \hat{I}\hat{v}\|^2_{H^m(\hat{T})} \le Ch^{2-2m} \|\hat{v} - \hat{I}\hat{v}\|^2_{H^t(\hat{T})} \\ &\le Ch^{2-2m} |\hat{v}|^2_{H^t(\hat{T})} = Ch^{2-2m} h^{2t-2} |v|^2_{H^t(T_h)} = Ch^{2(t-m)} |v|^2_{H^t(T_h)}. \end{aligned}$$

Summing up this yields

$$\sum_{T \in \mathcal{T}_h} \|v - I_h v\|^2_{H^m(T_h)} \le Ch^{2(t-m)} \sum_{T \in \mathcal{T}_h} |v|^2_{H^t(T_h)} = Ch^{2(t-m)} |v|^2_{H^t(T_h)}.$$

□

Remark 5.2 *Note that in Theorem* 5.2 *we cannot write* $\|u - I_h u\|_{H^m(\Omega)}$ *in general since* $u - I_h u$ *might not be in* $H^m(\Omega)$. *The operator* I_h *represents only a piecewise interpolation from which, in general, no global regularity properties follow.*

Let N_i, $i = 1, \ldots, M$, be the nodes of $\mathcal{T}_h$. For a continuous function $u \in C^0(\bar{\Omega})$ we now consider the piecewise linear interpolant (again using the same operator symbol I_h) $I_h u \in V_h$ by

$$I_h u(N_i) = u(N_i), \qquad i = 1, \ldots, M. \tag{5.2}$$

Note that on $K \in \mathcal{T}_h$, $I_h u$ is the linear interpolant of u.

Corollary 5.1 *Let* $\Omega \subset \mathbf{R}^2$ *be a polygon with a quasi-uniform, regular and shape-regular mesh, and* I_h *be the piecewise linear interpolation operator (piecewise with respect to the triangulation) with respect to the vertices of the mesh.*

Then,

$$\|u - I_h u\|_{H^1(\Omega)} \le Ch|u|_{H^2(\Omega)} \quad \text{for all } u \in H^2(\Omega).$$

Proof. As I_h interpolates at the vertices of the mesh, $I_h u$ is continuous and piecewise linear, i.e. $I_h u \in H^1(\Omega)$. An application of Theorem 5.2 proves

$$\|u - I_h u\|^2_{H^1(\Omega)} = \sum_{T \in \mathcal{T}_h} \|u - I_h u\|^2_{H^1(T)} \le Ch^2 |u|^2_{H^2(\Omega)}.$$

□

Now we are in a position to present an a priori error estimate for the finite element method dealing with elliptic problems of second order. Assume that we are solving a variational problem in $V = H_0^1(\Omega)$ ($\Omega \subset \mathbf{R}^2$ is a Lipschitz continuous polygonal domain),

$$u \in V : \quad a(u, v) = L(v) \quad \forall v \in V, \tag{5.3}$$

where a is a continuous, V-elliptic bilinear form, and L is a continuous linear form on V. We then consider the finite element approximation u_h to u defined by

$$u_h \in V_h : \quad a(u_h, v) = L(v) \quad \forall v \in V_h. \tag{5.4}$$

Selecting any finite-dimensional subspace $V_h \subset V$ there holds Céa's lemma. In particular, selecting V_h to be the space of continuous, piecewise linear functions defined on a mesh $\mathcal{T}_h$ satisfying the shape-regularity condition (5.1) there holds (applying Céa's lemma)

$$\|u - u_h\|_{H^1(\Omega)} \leq \frac{C_a}{\alpha} \|u - I_h u\|_{H^1(\Omega)}. \tag{5.5}$$

Here, I_h is the interpolation operator defined in (5.2).

Therefore, applying the results from §5.1, in particular Corollary 5.1, we conclude that there holds the following *a priori error estimate.*

Theorem 5.3 (a priori error estimate)
Assume that the solution u of (5.3) *satisfies $u \in H^2(\Omega)$ and that $u_h \in V_h$ is the finite element approximation defined by* (5.4) *(using piecewise linear functions on a shape regular mesh). Then there exists a constant $C > 0$ which is independent of h such that*

$$\|u - u_h\|_{H^1(\Omega)} \leq C\, h\, |u|_{H^2(\Omega)}. \tag{5.6}$$

This means that u_h converges linearly in h to u in the $H^1(\Omega)$-norm.