有限元方法

Finite Element Methods

Chapter 1: Overview of the Finite Element Method

主讲人: 李琦

liqihao@chd.edu.cn

School of Science, Chang'an University

Outline

Variational form of elliptic PDEs

Consider for a given function $f : (0,1) \to \mathbb{R}$ the solution $u : (0,1) \to \mathbb{R}$ of the two-point boundary value problem

$$\text{(BVP)} \qquad \begin{cases} -u''(x) = f(x) \quad \text{for } x \in (0,1), \\ \quad u(0) = 0, \quad u'(1) = 0. \end{cases}$$

The idea is to pass from (BVP) to a system of linear equations — which can be solved on a computer — by projection onto a finite-dimensional subspace.

By multiplying the differential equation with any sufficiently regular test function v with $v(0) = 0$, integrating over $x \in (0,1)$, and integrating by parts. Then any solution u of (BVP) satisfies

$$\begin{aligned} (f, v) := \int_0^1 f(x)v(x)dx &= -\int_0^1 u''(x)v(x)\,dx \\ &= \int_0^1 u'(x)v'(x)\,dx \\ &=: a(u,v), \end{aligned}$$

where we have used that $u'(1) = 0$ and $v(0) = 0$.

Let us (formally for now) define the space

$$V := \{v \in L^2(0,1) : a(v,v) < \infty, v(0) = 0\}.$$

Then we can pose the following problem: Find $u \in V$ such that

$$\text{(W)} \qquad a(u,v) = (f,v) \quad \forall v \in V,$$

holds. This is called the **weak or variational form** of (BVP) (since v varies over all V). If the solution u of (W) is twice continuously differentiable and f is continuous, one can prove (by taking suitable test functions v) that u satisfies (BVP). **On the other hand, there are solutions of (W) even for a discontinuous right-hand side f.** Since then the second derivative of u is discontinuous, u is not necessarily a solution of (BVP). For this reason, $u \in V$ satisfying (W) is called a **weak solution** of (BVP).

Theorem

Suppose $f \in C^0([0,1])$ and $u \in C^2([0,1])$ satisfy (W). Then u solves (BVP).

- Boundary Conditions

Boundary Condition	**Variational Name**	**Proper Name**
$u(x) = 0$	essential	Dirichlet
$u'(x) = 0$	natural	Neumman

Note that the Dirichlet boundary condition $u(0) = 0$ appears **explicitly** in the definition of V, while the Neumann condition $u'(1) = 0$ is **implicitly** incorporated in the variational formulation. In the context of finite element methods, Dirichlet conditions are therefore frequently called **essential conditions**, while Neumann conditions are referred to as **natural conditions**.

Ritz-Galerkin approximation

The fundamental idea is now to approximate u by considering (W) on a finite-dimensional subspace $S \subset V$. We are thus looking for $u_S \in S$ satisfying

$$(\mathrm{W}_S) \qquad a(u_S, v) = (f, v) \quad \forall\, v \in S,$$

Note that this is still the same equation; only the function spaces have changed.

This is a crucial point in (conforming) finite element methods.

(Nonconforming methods, for which $S \nsubseteq V$, will be treated in later content.)

We first have to ask whether (W_S) has a unique solution.

Theorem

Given $f \in L^2(0,1)$, (W_S) has a unique solution.

Proof.

Since S is finite-dimensional, there exists a basis $\phi_1, \phi_2, \cdots, \phi_n$ of S. Due to the bilinearity of $a(\cdot,\cdot)$, it suffices to require that $u_S = \sum_{i=1}^{n} U_i \phi_i \in S$, $U_i \in \mathbb{R}$ for $i = 1, \ldots, n$, satisfies

$$a(u_S, \phi_j) = (f, \phi_j), \quad \forall\ 1 \le j \le n.$$

This is now a system of linear equations for the unknown coefficients U_i. If we define

$$\begin{aligned}
\mathbf{U} &= (U_1, \ldots, U_n)^T \in \mathbb{R}^n, \\
\mathbf{F} &= (F_1, \ldots, F_n)^T \in \mathbb{R}^n, \quad F_i = (f, \phi_i), \\
\mathbf{K} &= (K_{ij}) \in \mathbb{R}^{n\times n}, \qquad K_{ij} = a(\phi_i, \phi_j),
\end{aligned}$$

we have that u_S satisfies (W_S) if and only if ("iff") $\mathbf{KU} = \mathbf{F}$. For a square system we know that uniqueness is equivalent to existence, as this is a finite dimensional system. This linear system has a unique solution iff $\mathbf{KV} = 0$ implies $\mathbf{V} = 0$. □

Continued proof.

To show this, we set $v_S = \sum_{i=1}^{n} V_i \phi_i \in S$, Then,

$$0 = \mathbf{KV} = (a(v_S, \phi_1), \cdots, a(v_S, \phi_n))^T$$

implies that

$$0 = \sum_{i=1}^{n} V_i a(v_S, \phi_i) = a(v_S, v_S) = \int_0^1 v_S'(x)^2 \, dx.$$

This means that v_S' must vanish almost everywhere and thus that v_S is constant. Since $v \in S \subset V$ implies $v_S(0) = 0$, we deduce that $v_S \equiv 0$, and hence it follows for the linear independence of the ϕ_i that $V_i = 0$ for all $1 \leq i \leq n$, this means that $\mathbf{V} = 0$. Therefore, the solution u_S to (W_S) must also exist and be unique. □

Error estimates

Now that we have an approximate solution $u_S \in S$, we are interested in estimating the **discretization error** $\|u_S - u\|$, which of course depends on the choice of S. The fundamental observation is that by subtracting (W) and (W_S) for the same test function $v_S \in S$, we obtain

$$a(u - u_S, v_S) = 0, \quad \forall v \in S. \tag{1.1}$$

This key property is called **Galerkin orthogonality**, and expresses that the discretization error is (in some sense) orthogonal to S. This can be exploited to derive error estimates in the **energy norm**

$$\|v\|_E^2 = a(v, v) \quad \forall v \in V.$$

It is straightforward to verify that this indeed defines a norm, which satisfies the **Cauchy-Schwarz** inequality

$$|a(v, w)| \leq \|v\|_E \|w\|_E \quad \forall v, w \in V.$$

Energy norm

We can thus show that for any $v_S \in S$,

$$\begin{aligned}\|u - u_S\|_E^2 &= a(u - u_S, u - v_S) + a(u - u_S, v_S - u_S) \\ &= a(u - u_S, u - v_S) \leq \|u - u_S\|_E \|u - v_S\|_E,\end{aligned}$$

due to the Galerkin orthogonality for $v_S - u_S \in S$. Taking the infimum over all $v_S \in S$, we obtain

$$\|u - u_S\|_E \leq \inf_{v_S \in S} \|u - v_S\|_E.$$

Since $u_S \in S$, we have

$$\inf_{v_S \in S} \|u - v_S\|_E \leq \|u - u_S\|_E,$$

Therefore,

$$\|u - u_S\|_E = \inf_{v_S \in S} \|u - v_S\|_E.$$

hence this **infimum is attained**. The discretization error is thus completely determined by the approximation error of the solution u of (W) by functions in S :

$$\|u - u_S\|_E = \min_{v_S \in S} \|u - v_S\|_E. \tag{1.2}$$

$L^2(0,1)$-norm

To derive error estimates in the $L^2(0,1)$-norm

$$\|v\|^2 = (v,v) = \int_0^1 v(x)^2\,dx,$$

we apply a **duality argument** (also called **Aubin-Nitsche trick**). Let w be the solution of the **dual** (or **adjoint**) problem

$$\begin{cases} -w''(x) = u(x) - u_S(x) \quad \text{for } x \in (0,1), \\ \quad w(0) = 0, \quad w'(1) = 0. \end{cases} \tag{1.3}$$

Inserting this into the error and integrating by parts (using $(u-u_S)(0) = 0 = w'(1)$ and adding the productive zero), we obtain for all $v_S \in S$ the estimate

$$\begin{aligned} \|u-u_S\|^2 &= (u-u_S, u-u_S) = (u-u_S, -w'') \\ &= ((u-u_S)', w') = a(u-u_S, w) - a(u-u_S, v_S) \\ &= a(u-u_S, w-v_S) \\ &\le \|u-u_S\|_E \|w-v_S\|_E. \end{aligned}$$

Dividing by $\|u - u_S\| = \|w''\|$ from above inequality and taking the infimum over all $v_S \in S$ yields

$$\|u - u_S\| \leq \|u - u_S\|_E \inf_{v_S \in S} \|w - v_S\|_E \|w''\|^{-1}.$$

To continue, we require an approximation property for S : There exists a constant $\epsilon > 0$ such that

$$\inf_{v_S \in S} \|g - v_S\|_E \leq \epsilon \|g''\|, \tag{1.4}$$

holds for sufficiently smooth $g \in V$. If we can apply this estimate to w and u, we obtain

$$\|u - u_S\| \leq \epsilon \|u - u_S\|_E = \epsilon \min_{v_S \in S} \|u - v_S\|_E \leq \epsilon^2 \|u''\| = \epsilon^2 \|f\|.$$

Theorem

Assumption (1.4) *implies that*

$$\|u - u_S\| \leq \epsilon \|u - u_S\|_E \leq \epsilon^2 \|u''\| = \epsilon^2 \|f\|.$$

Remark

The point of course is that $\|u-u_S\|_E$ is of order ϵ whereas $\|u-u_S\|$ is of order ϵ^2.

Remark

This is another key observation: The error estimate depends on the regularity of the weak solution u, and hence on the data f. The smoother u, the better the approximation. Of course, we wish that ϵ can be made arbitrarily small by choosing S sufficiently large. The finite element method is characterized by a special class of subspaces – of **piecewise polynomials** – which have these approximation properties.

Approximation by piecewise polynomials

Let $0 = x_0 < x_1 < \cdots < x_n = 1$ be a partition of $[0, 1]$, and let S be the linear space

$$S := \{v \in C^0(0,1) : v|_{[x_{i-1},x_i]} \in \mathcal{P}_1 \quad \text{and} \quad v(0) = 0\},$$

where $\mathcal{P}_1$ is the space of all linear polynomials. (The fact that $S \subset V$ is not obvious, and will be proved later.) A basis of S, which is especially convenient for the implementation, is formed by the linear hat functions

$$\phi_i = \begin{cases} \dfrac{x - x_{i-1}}{x_i - x_{i-1}} & \text{if } x \in [x_{i-1}, x_i], \\ \dfrac{x_{i+1} - x}{x_{i+1} - x_i} & \text{if } x \in [x_i\ x_{i+1}] \text{ and } i < n, \\ 0 & \text{else} \end{cases}$$

for $1 \le i \le n$, which satisfy $\phi_i(0) = 0$ and hence $\phi_i \in S$. Furthermore,

$$\phi_i(x_j) = \delta_{ij} := \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Remark

$\{\phi_i\}$ is called a **nodal basis** for S, and $\{v(x_i)\}$ are the **nodal values** of a function v. (The points $\{x_i\}$ are called the **nodes**.)

This nodal basis property immediately yields linear independence of the ϕ_i.

- **Linear independence:** The set $\{\phi_i\}$ is **linearly independent** since $\sum_{i=1}^{n} c_i \phi_i(x_j) = 0$ implies $c_j = 0$.
- **Span:** To show that the $\{\phi_i\}$ span S, we consider the **interpolant** $v_I \in S$ of a given $v \in V$, defined via

$$v_I := \sum_{i=1}^{n} v(x_i)\phi_i(x).$$

For $v_S \in S$, the interpolation error $v_S - (v_S)_I$ is piecewise linearas well, and since $(v_S)_I(x_i) = v_S(x_i)$ for all $1 \leq i \leq n$, this implies that $v_S - (v_S)_I \equiv 0$. Any $v_S \in S$ can thus be written as a unique linear combination of ϕ_i (given by its interpolant), and hence the $\{\phi_i\}$ form a basis of S. We also note that this implies that the **interpolation operator** $\mathcal{I} : V \to S$, $v \mapsto v_I$ is a projection (i.e., $\mathcal{I} \circ \mathcal{I} = \mathcal{I}$).

We will now prove the following approximation theorem for the interpolant.

Theorem

Let $h = \max_{1\le n\le n}(x_i - x_{i-1})$. *Then*

$$\|u - u_I\|_E \le Ch\|u''\|$$

for all $u \in V$, *where* C *is independent of* h *and* u.

Proof.

We now consider this error separately on each element $[x_{i-1}, x_i]$, i.e., we show that

$$\int_{x_{i-1}}^{x_i} (u - u_I)'(x)^2\,dx \le C^2 h_i^2 \int_{x_{i-1}}^{x_i} u''(x)^2\,dx.$$

First, since u_I is piecewise linear, the error $e := u - u_I$ satisfies $(e|_{[x_{i-1},x_i]})'' = (u|_{[x_{i-1},x_i]})''$. Using the affine transformation $\tilde{e}(t) := e(x(t))$ with $x(t) = x_{i-1} + t(x_i - x_{i-1})$ (a scaling argument), the previous estimate is equivalent to

$$\int_0^1 \tilde{e}'(t)^2\,dt \le C^2 \int_0^1 \tilde{e}''(t)^2\,dt. \tag{1.5}$$

□

Continued proof:

(This is an elementary version of *Poincaré's inequality*). Since u_I is the nodal interpolant of u, the error satisfies $e(x_{i-1}) = e(x_i) = 0$. In addition, u_I is linear and u continuously differentiable on $[x_{i-1}, x_i]$. Hence, e is continuously differentiable on $[0,1]$ with $\tilde{e}(0) = \tilde{e}(1)$, and Rolle' s theorem yields a $\xi \in (0,1)$ with $\tilde{e}'(\xi) = 0$. Thus, for all $y \in [0,1]$ we have (with $\int_a^b f(t)\,dt = -\int_b^a f(t)\,dt$ for $a > b$)

$$\tilde{e}'(y) = \tilde{e}'(y) - \tilde{e}'(\xi) = \int_\xi^y \tilde{e}''(t)\,dt.$$

We can now use the Cauchy-Schwarz inequality to estimate

$$|\tilde{e}'(y)|^2 = \left|\int_\xi^y \tilde{e}''(t)\,dt\right|^2 \le \left|\int_\xi^y 1^2\,dt\right| \cdot \left|\int_\xi^y \tilde{e}''(t)^2\,dt\right| \le |y-\xi| \int_0^1 \tilde{e}''(t)^2\,dt.$$

Integrating both sides with respect to y and taking the supremum over all $\xi \in (0,1)$ yields (1.5) with

$$C^2 = \sup_{\xi\in(0,1)} \int_0^1 |y-\xi|\,dy = \frac{1}{2}.$$

□

Continued proof:

Summing over all elements and estimating h_i by h shows the approximation property (1.4) for S with $\epsilon := Ch$. For this choice of S, the solution u_S of (W_S) satisfies

$$\|u - u_S\|_E \leq \min_{v_S \in S} \|u - v_S\|_E \leq \|u - u_I\|_E \leq Ch\|u''\|,$$

as well as

$$\|u - u_S\| \leq C^2 h^2 \|u''\|. \tag{1.6}$$

□

These are called a **priori estimates**, since they only require knowledge of the given data $f = -u''$ but not of the solution u_S. They tell us that if we can make the mesh size h arbitrarily small, we can approximate the solution u of (W) arbitrarily well. Note that the power of h is one order higher for the $L^2(0,1)$ norm compared to the energy norm, **which represents the fact that it is more difficult to control errors in the derivative than errors in the function value**.

Implementation

As seen in Frame 9, the numerical computation of $u_S \in S$ boils down to solving the linear system $\mathbf{KU} = \mathbf{F}$ for the vector of coefficients $\mathbf{U}$. The missing step is the computation of the elements $\mathbf{K}_{ij} = a(\phi_i, \phi_j)$ of $\mathbf{K}$ and the entries $\mathbf{F}_j = (f, \phi_j)$ of $\mathbf{F}$. (This procedure is called **assembly**.) In principle, this can be performed by computing the integrals for each pair (i, j) in a nested loop (node-based assembly).

A more efficient approach (especially in higher dimensions) is **element-based assembly**: The integrals are split into sums of contributions from each element, e.g.,

$$a(\phi_i, \phi_j) = \int_0^1 \phi_i'(x)\phi_j'(x)\,dx = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} \phi_i'(x)\phi_j'(x)\,dx =: \sum_{k=1}^{n} a_k\left(\phi_i, \phi_j\right),$$

and the contributions from a single element for all (i, j) are computed simultaneously.

Here we can exploit that by its definition, ϕ_i is non-zero only on the two elements $[x_{i-1}\ x_i]$ and $[x_i\ x_{i+1}]$. Hence, for each element $[x_{k-1}\ x_k]$, the integrals are non-zero only for pairs (i, j) with $k-1 \leq i, j \leq k$. Note that this implies that $\mathbf{K}$ is **tridiagonal** and therefore **sparse** (meaning that the number of non-zero elements grows as n, not n^2), which allows efficient solution of the linear system even for large n, e.g., by the method of conjugate gradients (since $\mathbf{K}$ is also symmetric and positive definite).

Another useful observation is that except for an affine transformation, the basis functions are the same on each element. We can thus use the substitution rule to transform the integrals over $[x_{i-1}\ x_i]$ to the **reference element** $[0, 1]$.

Setting $\xi(x) = \frac{x - x_{k-1}}{x_k - x_{k-1}}$ and

$$\hat{\phi}_1(\xi) = 1 - \xi, \quad \hat{\phi}_2(\xi) = \xi,$$

we have that $\phi_{k-1}(x) = \hat{\phi}_1(\xi(x))$ and $\phi_k(x) = \hat{\phi}_2(\xi(x))$. Using $\xi'(x) = (x_k - x_{k-1})^{-1} = h_k^{-1}$, the integrals for $i, j \in \{k-1, k\}$ can therefore be computed via

$$\int_{x_{k-1}}^{x_k} \phi_i'(x)\phi_j'(x)\,dx = h_k^{-1} \int_0^1 \hat{\phi}'_{\tau(i)}(\xi)\hat{\phi}'_{\tau(j)}(\xi)d\xi,$$

where

$$\tau(i) = \begin{cases} 1 & \text{if } i = k-1, \\ 2 & \text{if } i = k, \end{cases}$$

is the so-called global-to-local index. (Correspondingly, the inverse mapping τ^{-1} is called the local-to-global index.) Since the derivatives of $\hat{\phi}_1$, $\hat{\phi}_2$ are constant, the contribution from the element $[x_{i-1}\ x_i]$ to $\mathbf{K}_{ij} = a(\phi_i, \phi_j)$ for $i, j \in \{k-1, k\}$ (the contribution for all other pairs (i, j) being zero) is thus

$$a_k(\phi_i, \phi_j) = \begin{cases} h_k^{-1} & \text{if } i = j, \\ -h_k^{-1} & \text{if } i \neq j. \end{cases}$$

The right-hand side (f, ϕ_j) can be computed in a similar way, using numerical quadrature if necessary. Alternatively, one can replace f by its nodal interpolant $f_I = \sum_{i=0}^{n} f(x_i)\phi_i$ and use

$$(f, \phi_j) \approx (f_I, \phi_j) = \sum_{i=0}^{n} f(x_i)(\phi_i, \phi_j).$$

The elements $M_{ij} := (\phi_i, \phi_j)$ of the mass matrix $\mathbf{M}$ are again computed elementwise using transformation to the reference element:

$$\int_{x_{k-1}}^{x_k} \phi_i(x)\phi_j(x)dx = h_k \int_0^1 \hat{\phi}_{\tau(i)}(\xi)\hat{\phi}_{\tau(j)}(\xi)d\xi = \begin{cases} \frac{h_k}{3} & \text{if } i = j, \\ \frac{h_k}{6} & \text{if } i \neq j. \end{cases}$$

This can be done at the same time as assembling $\mathbf{K}$. Setting $\mathbf{f} := (f(x_1), \ldots, f(x_n))^T$, the right-hand side of the linear system is then given by $\mathbf{F} = \mathbf{Mf}$.

Finally, the Dirichlet condition $u(0) = 0$ can be enforced by replacing the first equation in the linear system by $U_0 = 0$, i.e., replacing the first row of $\mathbf{K}$ by $(1, 0, \dots)$ and the first element of $\mathbf{F}$ by 0. The main advantage of this approach is that it can easily be extended to non-homogeneous Dirichlet conditions $u(0) = g$ (by replacing the first element with g).

The full algorithm (in MATLAB-like notation) is given in Algorithm 1.

Algorithm Finite element method in 1D

Require: $0 = x_0 < \cdots < x_n = 1, F := (f(x_0), \ldots, f(x_n))^T$

1: Set $K_{ij} = M_{ij} = 0$

2: **for** $k = 1, 2, \cdots, n$ **do**

3: Set $h_k = x_k - x_{k-1}$

4: Set $K_{k-1:k,k-1:k} \leftarrow K_{k-1:k,k-1:k} + \frac{1}{h_k} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$

5: Set $M_{k-1:k,k-1:k} \leftarrow M_{k-1:k,k-1:k} + \frac{h_k}{6} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

6: **end for**

7: $K_{0,1:n} = 0, \ K_{0,0} = 1, \ M_{0,0:n} = 0$

8: Solve $\mathbf{KU} = \mathbf{MF}$

Ensure: $\mathbf{U}$

Posteriori estimates

The a priori estimate (1.6) is important for proving convergence as the mesh size $h \to 0$, but often pessimistic in practice since it depends on the global regularity of u''. If $u(x)''$ is large only in some parts of the domain, it would be preferable to reduce the mesh size locally.

For this, a **posteriori estimates** are useful, which are localized error estimates for each element but involve the computed solution u_S. This gives information on which elements should be refined (i.e., replaced by a larger number of smaller elements).

We consider again the space S of piecewise linear finite elements on the nodes $x_0, \cdots, x_n$ with mesh size h. We once more apply a duality trick: Let w be the solution of

$$\begin{cases} -w''(x) = u(x) - u_S(x) & \text{for } x \in (0,1), \\ w(0) = 0, \quad w'(1) = 0, \end{cases} \tag{1.7}$$

and proceed as before, yielding

$$\|u - u_S\|^2 = a(u - u_S, w - v_S)$$

for all $v_S \in S$. We now choose $v_S = w_I \in S$, the interpolant of w. Then we have

$$\begin{aligned} \|u - u_S\|^2 =& a(u - u_S, w - w_I) = a(u, w - w_I) - a(u_S, w - w_I) \\ =& (f, w - w_I) - a(u_S, w - w_I). \end{aligned}$$

Note that the unknown solution u of (W) no longer appears on the right-hand side. We now use the specific choice of v_S to localize the error inside each element $[x_{i-1}\ x_i]$. Writing the integrals over $[0,1]$ as sums of integrals over the elements, we can integrate by parts on each element and use the fact that $(w - w_I)(x_i) = 0$ to obtain

$$\begin{aligned}
\|u - u_S\|^2 &= \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} f(x)\,(w - w_I)\,(x)\,dx - \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} u_S'(x)\,(w - w_I)'\,(x)\,dx \\
&= \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} (f + u_S'')\,(x)\,(w - w_I)\,(x)\,dx \\
&\leq \sum_{i=1}^{n} \left(\int_{x_{i-1}}^{x_i} (f + u_S'')\,(x)^2\,dx \right)^{\frac{1}{2}} \left(\int_{x_{i-1}}^{x_i} (w - w_I)\,(x)^2\,dx \right)^{\frac{1}{2}}
\end{aligned}$$

by the Cauchy-Schwarz inequality.

The first term contains the **finite element residual**

$$R_h := f + u_S'',$$

which we can evaluate after computing u_S. For the second term, one can show (similarly as in the proof of the a priori error estimate (1.6)) that

$$\left(\int_{x_{i-1}}^{x_i} (w - w_I)(x)^2 \, dx \right)^{\frac{1}{2}} \leq \frac{h_i^2}{2} \|w''\|$$

holds, from which we obtain

$$\begin{aligned} \|u - u_S\|^2 &\leq \frac{1}{2} \|w''\| \sum_{i=1}^{n} h_i^2 \|R_h\|_{L^2(x_{i-1}, x_i)} \\ &= \frac{1}{2} \|u - u_S\| \sum_{i=1}^{n} h_i^2 \|R_h\|_{L^2(x_{i-1}, x_i)} \end{aligned}$$

by the definition of w. This yields the a **posteriori estimate**

$$\|u - u_S\| \leq \frac{1}{2} \sum_{i=1}^{n} h_i^2 \|R_h\|_{L^2(x_{i-1}, x_i)}.$$

This estimate can be used for an adaptive procedure: Given a tolerance $\tau > 0$,

1: choose initial mesh $0 = x_0^{(0)} < \cdots < x_{n^{(0)}}^{(0)} = 1$, compute corresponding solution $u_{S^{(0)}}$, evaluate $R_{h^{(0)}}$, set $m = 0$

2: **while** $\sum_{i=1}^{n^{m+1}} (h_i^{(m)})^2 \|R_{h^{(m)}}\|_{L^2(x_{i-1}^{(m)}, x_i^{(m)})} > \tau$ **do**

3: choose new mesh $0 = x_0^{(m+1)} < \cdots < x_{n^{(m+1)}}^{(m+1)} = 1$,

4: compute corresponding solution $u_{S^{(m+1)}}$

5: evaluate $R_{h^{(m+1)}}$

6: set $m \leftarrow m + 1$

7: **end while**

There are different strategies to choose the new mesh. A common requirement is that the strategy should be

- **reliable**, meaning that the error on the new mesh in a certain norm can be guaranteed to be less than a given tolerance $\tau > 0$,
- as well as **efficient**, meaning that the number of new nodes should not be larger than necessary.
- One (**simple**) possibility is to refine those elements where $\|R_h\|$ is largest (or larger than a given threshold) by replacing them with two elements of half size.