Quantity	$h = 1/40$	$h = 1/80$	$h = 1/160$	$h = 1/320$	Order
$u(L) - u_h(L)$	0.6720	0.1628	0.04038	0.01008	$\mathscr{O}(h^2)$
$u'(L) - u'_h(L)$	-11.469	-7.5888	-4.1838	-2.1783	$\mathscr{O}(h)$
$u'(L) - \mathscr{I}u'(L)$	-13.706	-8.0704	-4.2939	-2.2045	$\mathscr{O}(h)$
$\int(u - u_h)\,dx$	0.2909	0.07064	0.01752	4.373E-3	$\mathscr{O}(h^2)$
$\int(u - \mathscr{I}u)\,dx$	-8.670E-4	-2.135E-4	-5.318E-5	-1.327E-5	$\mathscr{O}(h^2)$

All of these experimental results are consistent with the theory. Notice that $u'_h(L)$, for which we have proved no theoretical estimate, converges with order $\mathscr{O}(h)$. The energy $U(u) = \frac{1}{2}a(u,u)$ in this case equals $\frac{1}{2}(\|u'\|_0^2 + \|u\|_0^2)$ and converges with order $\mathscr{O}(h^2)$.

3.4 Fourth-order problems in one dimension

Finite element methods for fourth-order problems can also be analyzed using Theorem 3.1. This will explain, in particular, the need for continuously differentiable basis functions and thus justify the introduction of the Hermite finite element space.

To make things precise, let $u : [0, L] \to \mathbb{R}$ satisfy $u(0) = g_0$, $u'(0) = d_0$, $q(L)u'''(L) + q'(L)u''(L) = -W$ (W: applied load), $q(L)u''(L) = T$ (T: applied torque), together with

$$\big(q(x)u''(x)\big)'' + c(x)u(x) = f(x), \qquad \forall x \in \Omega\,. \tag{3.37}$$

The coefficient $q(x)$ must be greater than some $q_{\min} > 0$, while $c(x) \geq 0$. For the distributed load $f(x)$ we assume it to be square-integrable again.

We have already seen, in Problem 1.6, that appropriate bilinear and linear forms for this problem are

$$a(u,v) = \int_0^L \big(qu"v" + cuv\big)\,dx, \qquad \ell(v) = \int_0^L f\,v\,dx + W\,v(L) + T\,v'(L)\,.$$

The essential boundary conditions are those imposing $u(0)$ and $u'(0)$.

Let us now consider some finite element space $\mathscr{W}_h$ from which we define the trial and test spaces as

$$\mathscr{S}_h = \{w_h \in \mathscr{W}_h | w_h(0) = g_0,\ w'_h(0) = d_0\}, \qquad \mathscr{V}_h = \{w_h \in \mathscr{W}_h | w_h(0) = 0,\ w'_h(0) = 0\}.$$

The finite element solution $u_h \in \mathscr{S}_h$ is computed from

$$a(u_h, v_h) = \ell(v_h), \qquad \forall v_h \in \mathscr{V}_h\,. \tag{3.38}$$

We now retrace the steps followed in Sections 3.3.1-3.3.4 to analyze the convergence of u_h towards the exact solution u.

Exact consistency. The residual is given by

$$r(u, v_h) = \int_0^L \left(q u" v"_h + c u v_h \right) \, dx - \int_0^L f v_h \, dx - W v_h(L) - T v'_h(L)$$

which upon integrating twice by parts in each element transforms into

$$\begin{aligned} r(u, v_h) \quad = \quad & \int_0^L \left((q u")" + c u - f \right) v_h \, dx \\ & - \left(q'(L) u''(L) + q(L) u'''(L) + W \right) v_h(L) \\ & + \left(q(L) u''(L) - T \right) v'_h(L) \\ & + \left(q'(0) u''(0) + q(0) u'''(0) \right) v_h(0) \\ & - q(0) u''(0) v'_h(0) \\ & - \sum_z \left(q'(z) u''(z) + q(z) u'''(z) \right) \left(v_h(z^-) - v_h(z^+) \right) \\ & + \sum_z q(z) u''(z) \left(v'_h(z^-) - v'_h(z^+) \right) \; , \end{aligned}$$

with z again running over all interelement boundaries. The first, second and third lines of the right-hand side above are zero because u satisfies the differential equation and the (natural) boundary conditions at L. For the fourth and fifth lines to be zero one needs to impose $v_h(0) = v'_h(0) = 0$, which justifies the definition of $\mathcal{V}_h$ above. For the sixth line to be zero, the function v_h must be continuous at z (zero jump in the function). Finally, for the last line to be zero, **the derivative v'_h must be continuous at** z (zero jumps in the derivative across interelement boundaries). This is a generalization of the result we obtained after evaluating consistency in §1.5.3.2.

As a consequence, for consistency **the finite element space $\mathcal{W}_h$ must consist of functions that are continuous and have continuous derivative**. This is why the simplest space for this problem is the Hermite H_3 piecewise cubic finite element space.

Coercivity. The bilinear form $a(\cdot,\cdot)$ turns out to be coercive in the so-called H^2-norm, which is given by

$$\| v_h \|_2 = \left[\int_0^L \left(v''_h(x)^2 + v'_h(x)^2 + v(x)^2 \right) dx \right]^{\frac{1}{2}} = \left(\| v_h \|_0^2 + \| v'_h \|_0^2 + \| v''_h \|_0^2 \right)^{\frac{1}{2}} \; . \quad (3.39)$$

This requires that $q_{\min} > 0$, in agreement with physical constraints.

Continuity. The continuity of $a(\cdot,\cdot)$ and $\ell(\cdot)$ in the H^2-norm also holds, and can be proved with the same arguments used for the second order case. Naturally, some requirements appear in the coefficients: q and c must be bounded and the distributed load f must have a finite integral. These requirements are physically sound and hold in most cases.

All hypothesis of Thm. 3.1 have been checked, from which we conclude:

Theorem 3.4. *If the finite element space $\mathscr{W}_h$ consists of* **continuously differentiable functions** *that in addition are C^2 in each element, then the finite element solution u_h defined by* (3.38) *exists, is unique, and satisfies*

$$\|u - u_h\|_2 \leq C \min_{w_h \in \mathscr{S}_h} \|u - w_h\|_2 \tag{3.40}$$

with the constant C independent of the adopted mesh.

We now take as $\mathscr{W}_h$ the **Hermite H_3 finite element space** introduced in section 1.5.4, which satisfies the hypotheses of the theorem since it is a piecewise cubic polynomial space contained in $C^1([0,L])$.

There is a corresponding **Hermite interpolant** of the exact solution u (assumed smooth), given by

$$\mathscr{I} u(x) = u(x_1) H_1(x) + u'(x_1) H_2(x) + u(x_2) H_3(x) + u'(x_2) H_4(x) + \ldots \tag{3.41}$$

Using this interpolant and following analogous steps to those in the proof of Theorem 3.3 one arrives at

Theorem 3.5. *Let u be a C^4 function in $\Omega = [0,L]$ such that $u(0) = g_0$ and $u'(0) = d_0$. Let $\mathscr{W}_h$ be the cubic Hermite finite element space built on a mesh $\mathscr{T}_h$ of Ω. Let h denote the length of the largest element in $\mathscr{T}_h$. Then, there exists a constant C_I such that*

$$E_2(\mathscr{S}_h, u) = \min_{w_h \in \mathscr{S}_h} \|u - w_h\|_2 \leq C_I \, h^2 \, \|u^{(4)}\|_0 \,, \tag{3.42}$$

$$E_1(\mathscr{S}_h, u) = \min_{w_h \in \mathscr{S}_h} \|u - w_h\|_1 \leq C_I \, h^3 \, \|u^{(4)}\|_0 \,, \tag{3.43}$$

$$E_0(\mathscr{S}_h, u) = \min_{w_h \in \mathscr{S}_h} \|u - w_h\|_0 \leq C_I \, h^4 \, \|u^{(4)}\|_0 \,, \tag{3.44}$$

where $u^{(4)} = d^4 u / dx^4$.

Combining Theorems 3.4 and 3.5 we conclude that the finite element solution u_h satisfies, for some constant $C > 0$,

$$\|u - u_h\|_2 \leq C \, h^2 \, \|u^{(4)}\|_0 \,, \tag{3.45}$$

and thus converges toward u as the mesh is refined. The order of convergence is $\mathscr{O}(h^2)$ in the H^2-norm. It is possible to prove that the convergence is also optimal in the H^1- and L^2-norms, with orders $\mathscr{O}(h^3)$ and $\mathscr{O}(h^4)$, respectively.

All the consequences of convergence in the H^1-norm thus hold, in particular that u_h converges **uniformly** to u in Ω, with order at least $\mathscr{O}(h^3)$ but in fact higher.

Additionally, we have that u'_h converges **uniformly** to u' in Ω, with order at least $\mathscr{O}(h^2)$. So, in fourth-order problems, $u'_h(x)$ approximates $u'(x)$ at all points and uniformly.

To confirm these statements numerically, let us again apply the method of manufactured solutions. We select the constants $q = 1$, $c = c_0$ and specify the solution

$$u(x) = \sin(\alpha x^2)$$

in the domain $[0, L]$ (with $L = 1$) so that $g_0 = u(0) = 0$ and $d_0 = u'(0) = 0$. By differentiating u we compute the appropriate distributed load f, end load W and end torque T.

$$f = u^{(4)} + c_0 u = 4\alpha^2 \left((4\alpha^2 x^4 - 3)\sin(\alpha x^2) - 12\alpha x^2 \cos(\alpha x^2) \right) + c_0 \sin(\alpha x^2) \ ,$$

$$W = -u^{(3)}(1) = 4\alpha^2 \left(3 \sin\alpha + 2\alpha\cos\alpha\right) \ ,$$

$$T = u''(1) = 2\alpha \left(\cos\alpha - 2\alpha\sin\alpha\right) \ .$$

We consider $c_0 = 10^6$ and $\alpha = 20$. We then run the code developed in Section 1.5, slightly modified so as to consider f not constant within each element. We begin with a uniform mesh of 20 Hermite cubic elements, so that $h = 1/20$. In Figure 3.3 we show the exact solution u, the finite element solution u_h and the H_3 interpolant $\mathscr{I}u$, together with their derivatives. The corresponding errors are shown in Fig. 3.4.

The convergence of u_h and $\mathscr{I}u$ towards u as h tends to zero, in L^2, H^1 and H^2 norms, can be seen in the following table:

Quantity	$h = 1/10$	$h = 1/20$	$h = 1/40$	$h = 1/80$	$h = 1/160$	Order
$\Vert u - u_h \Vert_0$	0.0553	4.029E-3	2.750E-4	1.734E-5	1.086E-6	$\mathscr{O}(h^4)$
$\Vert u - \mathscr{I}u \Vert_0$	0.0983	5.302E-3	3.688E-4	2.350E-5	1.476E-6	$\mathscr{O}(h^4)$
$\Vert u' - u_h' \Vert_0$	2.7243	0.3536	0.0506	6.499E-3	8.175E-4	$\mathscr{O}(h^3)$
$\Vert u' - \mathscr{I}u' \Vert_0$	3.4610	0.3707	0.0512	6.517E-3	8.180E-4	$\mathscr{O}(h^3)$
$\Vert u'' - u_h'' \Vert_0$	225.92	48.424	13.291	3.3800	0.8484	$\mathscr{O}(h^2)$
$\Vert u'' - \mathscr{I}u'' \Vert_0$	222.44	48.385	13.290	3.3800	0.8484	$\mathscr{O}(h^2)$

Numerical convergence in other quantities is as follows:

Quantity	$h = 1/10$	$h = 1/20$	$h = 1/40$	$h = 1/80$	$h = 1/160$	Order
$u(L) - u_h(L)$	0.1726	-2.198E-4	-2.661E-4	-2.007E-5	-1.305E-6	$\mathscr{O}(h^4)$
$u'(L) - u_h'(L)$	-1.0843	-0.1365	-0.0158	-1.098E-3	-7.026E-5	$\mathscr{O}(h^4)$?
$u''(L) - u_h''(L)$	-676.63	103.00	72.175	23.577	6.5204	$\mathscr{O}(h^2)$
$u''(L) - \mathscr{I}u''(L)$	-631.51	102.25	72.137	23.576	6.5204	$\mathscr{O}(h^2)$
$\int (u - u_h)\, dx$	-3.037E-6	-3.461E-8	-1.061E-9	4.528E-11	-8.328E-12	$\mathscr{O}(h^5)$?
$\int (u - \mathscr{I}u)\, dx$	-8.554E-3	-3.034E-4	-1.709E-5	-1.042E-6	-6.476E-8	$\mathscr{O}(h^4)$

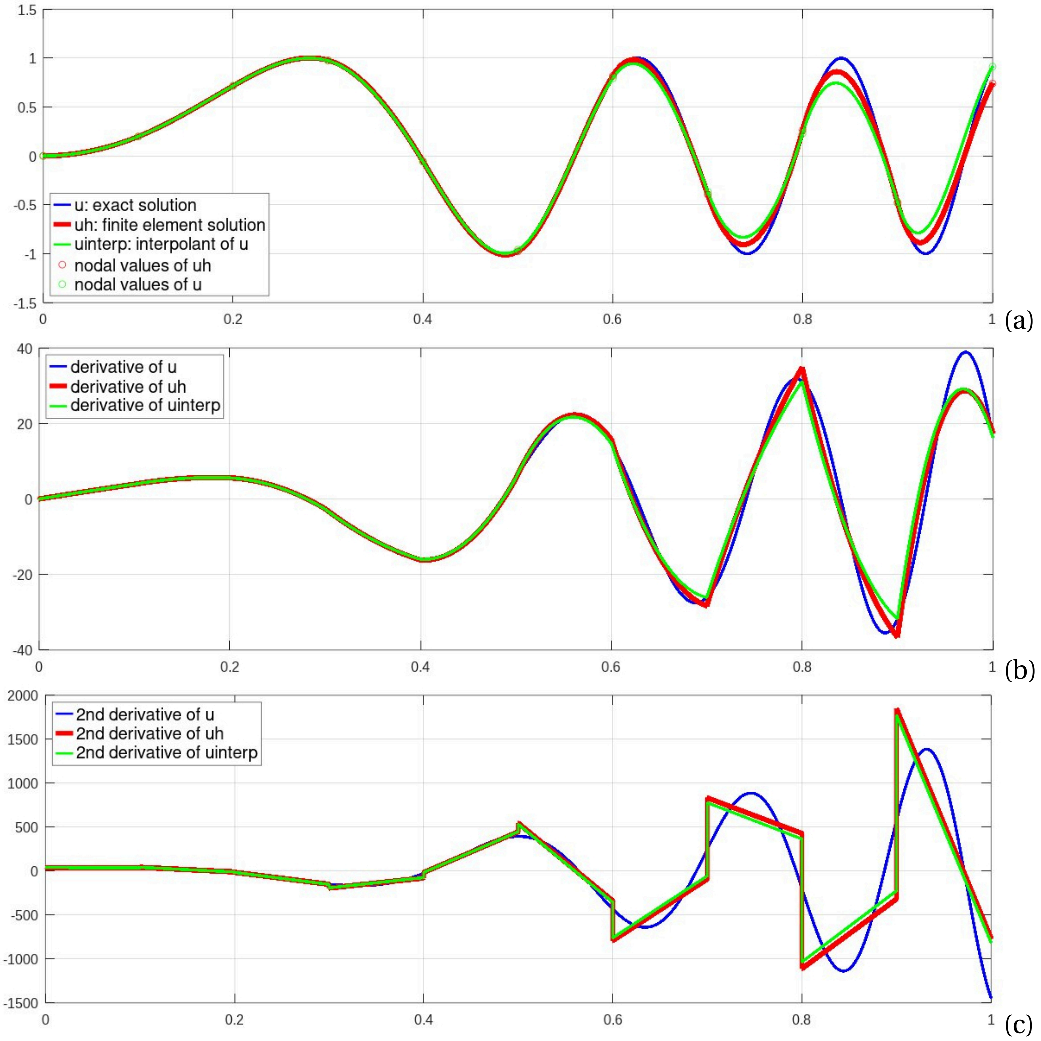


Figure 3.3 Example of numerical convergence. Results on a mesh of 10 Hermite cubic elements, all of size $h = 1/10$. (a) Exact solution u, finite element solution u_h and interpolant of exact solution $\mathscr{I}u$. (b) First derivative of u, u_h and $\mathscr{I}u$. (c) Second derivative of u, u_h and $\mathscr{I}u$.

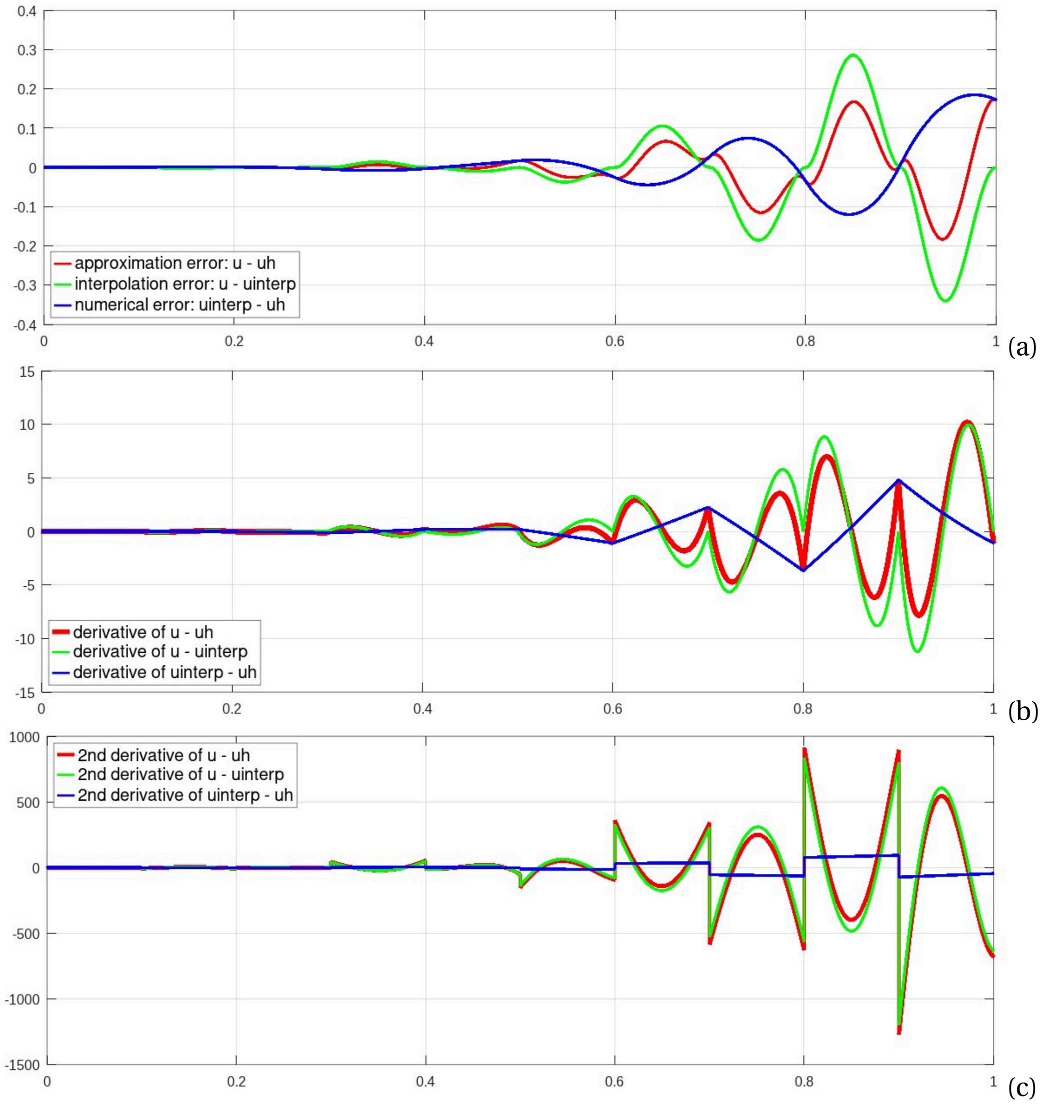


Figure 3.4 Example of numerical convergence. Results on a mesh of 10 Hermite cubic elements, all of size $h = 1/10$. (a) Approximation error $u - u_h$, interpolation error $u - \mathscr{I}u$ and numberical error $\mathscr{I}u - u_h$. (b) First derivative of the errors in (a). (c) Second derivative of the errors in (a).

These experimental results are consistent with the theory. Though for the second derivative $u_h''(L)$ we have no theoretical estimate, it converges with order $\mathcal{O}(h^2)$. The energy $U(u) = \frac{1}{2}a(u,u)$ equals $\frac{1}{2}(\|u''\|_0^2 + c_0\,\|u\|_0^2)$ and converges with order $\mathcal{O}(h^4)$. It is interesting to notice the error of $\int u_h\,dx$, which is much smaller than that of $\int \mathscr{I}\,u\,dx$. The integral of the numerical solution is superconvergent (it converges faster than the interpolation estimate).

3.5 Second-Order Problems in Two Dimensions

As you have seen in the previous sections, the strategy for analyzing a finite element method consists of:
(a) Checking the hypotheses of Céa's lemma (Theorem 3.1), followed by
(b) Applying an interpolation result to estimate the convergence of $\min_{w_h\in\mathscr{S}_h}\|u-w_h\|$. The norm to be used in (b) is determined while performing (a).

We will now pursue this strategy for second order problems in 2D, for which a finite element method was defined in Chapter 2. Let Ω be a two-dimensional domain and let u be a smooth function satisfying

$$-\operatorname{div}(K\nabla u) + c\,u = f\,, \qquad \text{in } \Omega, \tag{3.46}$$

together with

$$\begin{aligned} u &= g\,, &&\text{on}\,\partial\Omega_D,\text{ and} && (3.47)\\ (K\nabla u)\cdot\check{n} &= H && \text{on } \partial\Omega_N\,. && (3.48)\end{aligned}$$

We assume, as in Problem 2.1 that $\partial\Omega_D$ and $\partial\Omega_N$ are disjoint and that their union covers $\partial\Omega$. The diffusion matrix $K(x)$ is assumed bounded, symmetric and positive definite, with all eigenvalues greater than some $k_{\min} > 0$. The coefficient $c(x)$ is assumed bounded and non-negative.

The finite element method introduced in 2.4 reads, as usual: *Find $u_h \in \mathscr{S}_h$ such that*

$$a(u_h, v_h) = \ell(v_h) \tag{3.49}$$

for all $v_h \in \mathscr{V}_h$, where

$$\begin{aligned} a(u_h,v_h) &= \int_\Omega (K\nabla u_h\cdot\nabla v_h + c\,u_h\,v_h)\;d\Omega\,, && (3.50)\\ \ell(v_h) &= \int_\Omega f\,v_h\;d\Omega + \int_{\partial\Omega_N} H\,v_h\;d\Gamma\,, && (3.51)\end{aligned}$$

and

$$\begin{aligned} \mathscr{S}_h &= \{w_h \in \mathscr{W}_h \mid w_h = g \text{ on } \partial\Omega_D\}\,, && (3.52)\\ \mathscr{V}_h &= \{w_h \in \mathscr{W}_h \mid w_h = 0 \text{ on } \partial\Omega_D\}\,. && (3.53)\end{aligned}$$

3.5.1 Checking the Hypotheses of Céa's Lemma

Exact consistency. The consistency residual is

$$\begin{aligned} r(u, v_h) &= a(u, v_h) - \ell(v_h) \\ &= \int_\Omega \left(K\nabla u \cdot \nabla v_h + c u v_h - f v_h\right)\, d\Omega - \int_{\partial\Omega_N} H v_h \, d\Gamma\,. \end{aligned}$$

We know that u is smooth by hypotheses and that v_h, belonging to the finite element space $\mathscr{W}_h$, is smooth inside each element. We want to determine what the required continuity at interelement boundaries is, and thus we assume no a priori continuity. We can nonetheless integrate by parts **element-wise**, which results in

$$\begin{aligned} r(u, v_h) = & \sum_e \int_{\Omega_e} \left(-\mathrm{div} K\nabla u + c u - f\right) v_h \, d\Omega \\ & + \int_{\partial\Omega_N} (K\nabla u \cdot \check{n} - H)\, v_h \, d\Gamma \\ & + \int_{\partial\Omega_D} (K\nabla u \cdot \check{n})\, v_h \, d\Gamma \\ & + \sum_a \int_{\gamma_a} [\![v_h]\!] K\nabla u \cdot \check{n} \, d\gamma\,. \end{aligned} \tag{3.54}$$

The first and second terms are automatically zero because u is an exact solution. The third term is zero because v_h is zero on $\partial\Omega_D$.

The fourth term is a sum over all inter-element boundaries (edges), where $[\![v_h]\!]$ is the jump in v_h across the edge γ_a. The properties of u do not guarantee at all that this term vanishes, and thus **for exact consistency to hold, the functions in $\mathscr{V}_h$ must be continuous at inter-element boundaries**.

Coercivity. Being a second order elliptic problem, the norm in which coercivity can be established is the H^1-norm, which reads

$$\| v_h \|_1 = \left[\int_\Omega \left(\|\nabla v_h(x)\|^2 + v_h(x)^2 \right) \, d\Omega \right]^{\frac{1}{2}}\,. \tag{3.55}$$

The notation $\|\nabla v_h(x)\|$ refers to the euclidean norm of the vector $\nabla v_h(x)$. It is frequent to consider the L^2-norm of a vector field $w : \Omega \to \mathbb{R}^d$, which is defined as

$$\| w \|_0 = \left[\int_\Omega \| w(x) \|^2 \, d\Omega \right]^{\frac{1}{2}}\,. \tag{3.56}$$

With this notation, it holds that $\| v_h \|_1^2 = \|\nabla v_h\|_0^2 + \| v_h \|_0^2$.

If $c(x) \geq c_{\min} > 0$ it is immediate to prove that

$$a(v_h, v_h) = \int_\Omega \left(K\nabla v_h \cdot \nabla v_h + c v_h^2 \right) \geq \min(k_{\min}, c_{\min}) \, \| v_h \|_1^2$$

for all $v_h \in \mathscr{V}_h$. Thus $a(\cdot,\cdot)$ is coercive wich coercivity constant $\alpha = \min(k_{\min}, c_{\min})$.

In pure diffusion problems (i.e., problems with $c(x) = 0$ for all x) the coercivity of $a(\cdot,\cdot)$ is less evident and depends on the boundary conditions. For example, if there is no Dirichlet boundary ($\partial\Omega_D = \emptyset$), then constant functions belong to $\mathcal{V}_h$. Let $v_h(x) = A \in \mathbb{R}$ for all $x \in \Omega$ be a constant function, with $A \neq 0$. Then, when $c(x) = 0$ for $x \in \Omega$,

$$a(v_h, v_h) = \int_\Omega k \|\nabla v_h\|^2 \, d\Omega = 0$$

and

$$\| v_h \|_1^2 = \int_\Omega A^2 \, d\Omega > 0 \, ,$$

contradicting coercivity.

This shows that, if $c(x) \equiv 0$, boundary conditions must be such that **constant functions do not belong to** $\mathcal{V}_h$. Such a condition is generally met because in purely diffusive problems the **problem is ill-posed if the measure (length) of** $\partial\Omega_D$ **is zero**. If $\partial\Omega_D$ has positive length (and the domain is bounded), then the so-called **Poincaré inequality** guarantees that there exists $C_P > 0$ (independent of h) such that

$$\| v_h \|_0^2 \leq C_P \int_\Omega \|\nabla v_h\|^2 \, d\Omega \, . \tag{3.57}$$

Using (3.57) it is easy to prove that $a(\cdot,\cdot)$ is coercive with just the hypothesis $c_{\min} \geq 0$. In fact,

$$\begin{aligned} a(v_h, v_h) &= \int_\Omega \left(K\nabla v_h \cdot \nabla v_h + c v_h^2 \right) \\ &\geq k_{\min} \int_\Omega \|\nabla v_h\|^2 \, d\Omega + c_{\min} \| v_h \|_0^2 \\ &\geq \frac{k_{\min}}{2} \int_\Omega \|\nabla v_h\|^2 \, d\Omega + c_{\min} \| v_h \|_0^2 + \frac{k_{\min}}{2C_P} \| v_h \|_0^2 \\ &\geq \min\left(\frac{k_{\min}}{2}, c_{\min} + \frac{k_{\min}}{2C_P} \right) \| v_h \|_1^2 \, . \end{aligned}$$

We see that there is coercivity, with the coercivity constant independent of the mesh and not depending on $c_{\min} > 0$.

Remark: To preclude non-zero constant functions from belonging to $\mathcal{V}_h$ it would suffice to fix just one nodal value of v_h to zero (that is, to have $\partial\Omega_D$ equal to a single point). In such a case the bilinear form is indeed coercive in $\mathcal{V}_h$, but the coercivity constant tends to zero as $h \to 0$. This reflects a mathematical fact about the exact problem, which is not well-posed if the length of $\partial\Omega_D$ is zero.

Continuity: The continuity of $a(\cdot,\cdot)$ in $H^1(\Omega)$ can be proved in much the same way as done in the one-dimensional case. In fact, one obtains

$$|a(u - w_h, v_h)| \leq (k_{\max} + c_{\max}) \| u - w_h \|_1 \, \| v_h \|_1 \, ,$$

where

$$k_{\max} = \max_{x \in \Omega} \max_{i,j} K_{ij}(x) \, .$$

What about the continuity of the linear functional? Remember that

$$\ell(v_h) = \int_\Omega f\, v_h\, d\Omega + \int_{\partial\Omega_N} H\, v_h\, d\Gamma\ ,$$

The first term is certainly continuous if $f \in L^2(\Omega)$, since by Cauchy-Schwartz inequality in $L^2(\Omega)$ we have that

$$\int_\Omega f\, v_h\, d\Omega \leq \|f\|_0\ \|v_h\|_0 \leq \|f\|_0\ \|v_h\|_1\ .$$

If the source term function $f : \Omega \to \mathbb{R}$ is not square-integrable this term can still be continuous in $H^1(\Omega)$, but the proof is more technical.

The second term of $\ell(\cdot)$ involves a one-dimensional integral over the boundary segment (or segments) $\partial\Omega_N$. As all one-dimensional integrals, it satisfies Cauchy-Schwartz inequality, so

$$\int_{\partial\Omega_N} H\, v_h\, d\Gamma \leq \|H\|_{L^2(\partial\Omega_N)}\ \|v_h\|_{L^2(\partial\Omega_N)}\ .$$

The last ingredient needed in the proof is a so-called **trace theorem**. The version we will use here, that we accept without proof, reads: "If $v_h \in \mathscr{W}_h$, where $\mathscr{W}_h$ is a continuous finite element space (i.e., $\mathscr{W}_h \subset C^0(\Omega)$ and piecewise polynomial), then for any subset Γ of $\partial\Omega$ there exists $C_T > 0$ **independent of the mesh** such that

$$\|v_h\|_{L^2(\Gamma)} \leq C_T \|v_h\|_1 .\text{"} \tag{3.58}$$

Many trace inequalities exist. They all share the structure of (3.58), in that a norm of the function evaluated at the boundary of Ω is bounded by a norm that considers the function over the interior of Ω.

Collecting the previous results, we have proved that

$$\ell(v_h) \leq m \|v_h\|_1$$

for all $v_h \in \mathscr{W}_h$, with

$$m = \|f\|_0\ +\ C_T \|H\|_{L^2(\partial\Omega_N)}\ . \tag{3.59}$$

3.5.2 Convergence

In the previous section all hypotheses of Céa's lemma have been checked with constants independent of the mesh (with some additional hypotheses that appeared along the way). We thus know that there exists C such that

$$\|u - u_h\|_1\ \leq\ C \min_{w_h \in \mathscr{S}_h} \|u - w_h\|_1\ . \tag{3.60}$$

Let us now consider a specific finite element space. We assume for now that the domain Ω is polygonal. We take as $\mathscr{W}_h$ the space of P_k **Lagrange finite elements** in two dimensions, of which the mesh $\mathscr{T}_h$ consists of triangles (and has no hanging nodes). This family of finite elements is very popular and general (any polygonal domain can be decomposed into triangles), and extends readily to dimensions greater than two.

Under a suitable set of hypotheses, it is possible to prove that, using P_k Lagrange elements,

$$E_1(\mathscr{S}_h, u) = \min_{w_h \in \mathscr{S}_h} \|u - w_h\|_1 \leq C_I\, h^k\, \|D^{k+1} u\|_0\ , \tag{3.61}$$

and

$$E_0(\mathscr{S}_h, u) = \min_{w_h \in \mathscr{S}_h} \|u - w_h\|_0 \leq C_I\, h^{k+1}\, \|D^{k+1} u\|_0\ , \tag{3.62}$$

for some $C_I > 0$ independent of u and of the mesh $\mathscr{T}_h$. In other words, the same estimates of Theorem 3.3 also hold for the two-dimensional case, with the only difference that the $(k+1)$-th derivative is now a tensor.

The necessary hypotheses, however, are much more technical than in 1D. Let us discuss them in some detail:

- **Regularity of the mesh.** The size h_e of each element is defined as the diameter of the subdomain Ω_e. The diameter of its inscribed circle is denoted by ρ_e. The mesh size is defined as $h = \max_e h_e$. A mesh is said to be **regular** if $\max_e h_e/\rho_e$ is bounded by some fixed constant. This essentially means that there are no small angles in the mesh, irrespective of how fine the mesh $\mathscr{T}_h$ is. **We adopt here the hypothesis that $\mathscr{T}_h$ is regular as $h \to 0$**, which is sufficient for E_1 to obey (3.61) if the other hypotheses are met.

  Remark: Regularity of the mesh is not strictly necessary for E_1 to satisfy (3.61) with P_k Lagrange elements. For P_1 elements it is known that a weaker condition than mesh regularity, called *maximum angle condition* (maximum angle $\leq \gamma < \pi$ for all meshes, irrespective of h), is sufficient for E_1 converge to zero with order h [1, 6].

- **Accurate approximation of the Dirichlet boundary condition.** Notice that, remembering that Ω is a polygon, the Dirichlet condition $u = g$ is imposed along straight segments. Let s be the arc-length tangential coordinate along one of the segments. If g, viewed as a function of s, is not continuous and piecewise polynomial, then there is no function w_h in $\mathscr{W}_h$ that satisfies $w_h = g$ on $\partial\Omega_D$. This would mean that $\mathscr{S}_h$ is empty!! For this reason the space $\mathscr{S}_h$ must be defined with an interpolation (or some sort of approximation) of g, that we will denote by g_h. The subscript h indicates that g_h in general depends on the mesh. The definition becomes

  $$\mathscr{S}_h = \left\{ w_h \in \mathscr{W}_h \mid w_h = g_h \text{ on } \partial\Omega_D \right\}\ . \tag{3.63}$$

  For E_1 to be of order h^k it is **necessary** that g_h is close enough to g, more specifically that

  $$\|g - g_h\|_{L^2(\partial\Omega_D)} \leq C_D\, h^k \tag{3.64}$$

  for some $C_D > 0$. Fortunately, if g_h is the P_k **Lagrange interpolant** of g (both viewed as functions of s) **and g is smooth enough**, then (3.64) is automatically satisfied.

In fact, for the P_k Lagrange interpolant it holds that

(a) $\|g - g_h\|_{L^2(\partial\Omega_D)}$ is of order h^{k+1} if $g''(s)$ is in $L^2(\partial\Omega_D)$. If g has no singularities, this requirement is equivalent to g being C^1.

(b) $\|g - g_h\|_{L^2(\partial\Omega_D)}$ is of order h^k if $g'(s)$ is in $L^2(\partial\Omega_D)$. If g has no singularities, this requirement is equivalent to g being C^0.

The proof can be easily adapted from that of (3.29), since $\partial\Omega_D$ is a one-dimensional manifold (possibly consisting of several disjoint parts).

With this we arrive at

Theorem 3.6. Convergence of P_k Lagrange finite element approximation for elliptic second order problems in 2D. *Let u be the solution to (3.46)-(3.48), supposed smooth enough for its $(k+1)$-th derivatives to be in $L^2(\Omega)$. Let the coefficients satisfy*

$$0 < k_{\min} \le k(x) \le k_{\max} < +\infty\,, \qquad in\,\Omega\,, \tag{3.65}$$

$$0 \le c_{\min} \le c(x) \le c_{\max} < +\infty\,, \qquad in\,\Omega\,, \tag{3.66}$$

$$f \in L^2(\Omega)\,, \tag{3.67}$$

$$H \in L^2(\partial\Omega_N)\,. \tag{3.68}$$

Let u_h be the solution to (3.49)-(3.53), with $\mathscr{W}_h$ being the P_k **Lagrange finite element space** *associated with a* **regular triangulation** *$\mathscr{T}_h$.*

Assume that (3.52) has been replaced by (3.63), where g_h is the P_k **Lagrange interpolant** *of g along $\partial\Omega_D$.*

Then, there exists a constant $C(u) > 0$, dependent on u but independent of the mesh size h, such that

$$\|u - u_h\|_1 \le C(u)\,h^k\,. \tag{3.69}$$

Proof. **(a)** The proof when the interpolation error of the Dirichlet condition is zero, that is, when $g_h = g$, is immediate. Of course for this to happen the boundary data g must be a polynomial of degree $\le k$. The estimate (3.69) follows directly from (3.60) and (3.61).

(b) When $g \neq g_h$ the proof requires a little more work. Let us assume for simplicity that $\partial\Omega_D = \partial\Omega$. Let θ be the *exact solution* of (3.46)-(3.48) when $f = 0$ and the boundary data is $g - g_h$. We can decompose $u = u_h^* + \theta$, u_h^* defined as the *exact solution* corresponding to boundary data g_h. We know that $\|u_h^* - u_h\|_1 \le C\,h^k$ because of (a) above. A stability inequality from the theory of elliptic PDEs establishes that

$$\|\theta\|_1 \le C_2\,\|\theta\|_{L^2(\partial\Omega)} = C_2\,\|g - g_h\|_{L^2(\partial\Omega)}\,,$$

which combined with (3.64) yields $\|\theta\|_1 \le C_3\,h^k$.

□

Thus, if the meshing algorithm is regular and the boundary conditions are regular and well approximated, the numerical solution u_h converges to the exact solution u in the sense of the H^1-norm.

Any quantity that is continuous in the H^1-norm will converge as well. For example, the H^1-norm itself will satisfy $\|u_h\|_1 \overset{h\to 0}{\rightarrow} \|u\|_1$.

3.5.3 Numerical example: The uniformly heated square rod with a hot lid

Consider u to be the solution of $-\Delta u = 1$ in the unit square domain $\Omega = (0,1) \times (0,1)$, with Dirichlet boundary conditions $u = 0$ over the left, bottom and right sides, and $u = \delta$ over the top side (the lid).

Under a thermal interpretation, u is the temperature field of a uniformly heated square bar with three of its sides in contact with a thermostat at temperature $u = 0$ and the remaining side in contact with a thermostat at $u = \delta$.

The problem seems physically possible, though perfect thermal contact with perfect thermostats are ideal boundary conditions that cannot be realized in practice.

We would like to predict its solution, approximately of course. For this purpose we build a sequence of finite element meshes M1, ..., M5, with mesh size approximately $h = 0.25, 0.25/2, \ldots, 0.25/16$. On each mesh we compute the finite element solution u_h defined by (3.49)-(3.53), with

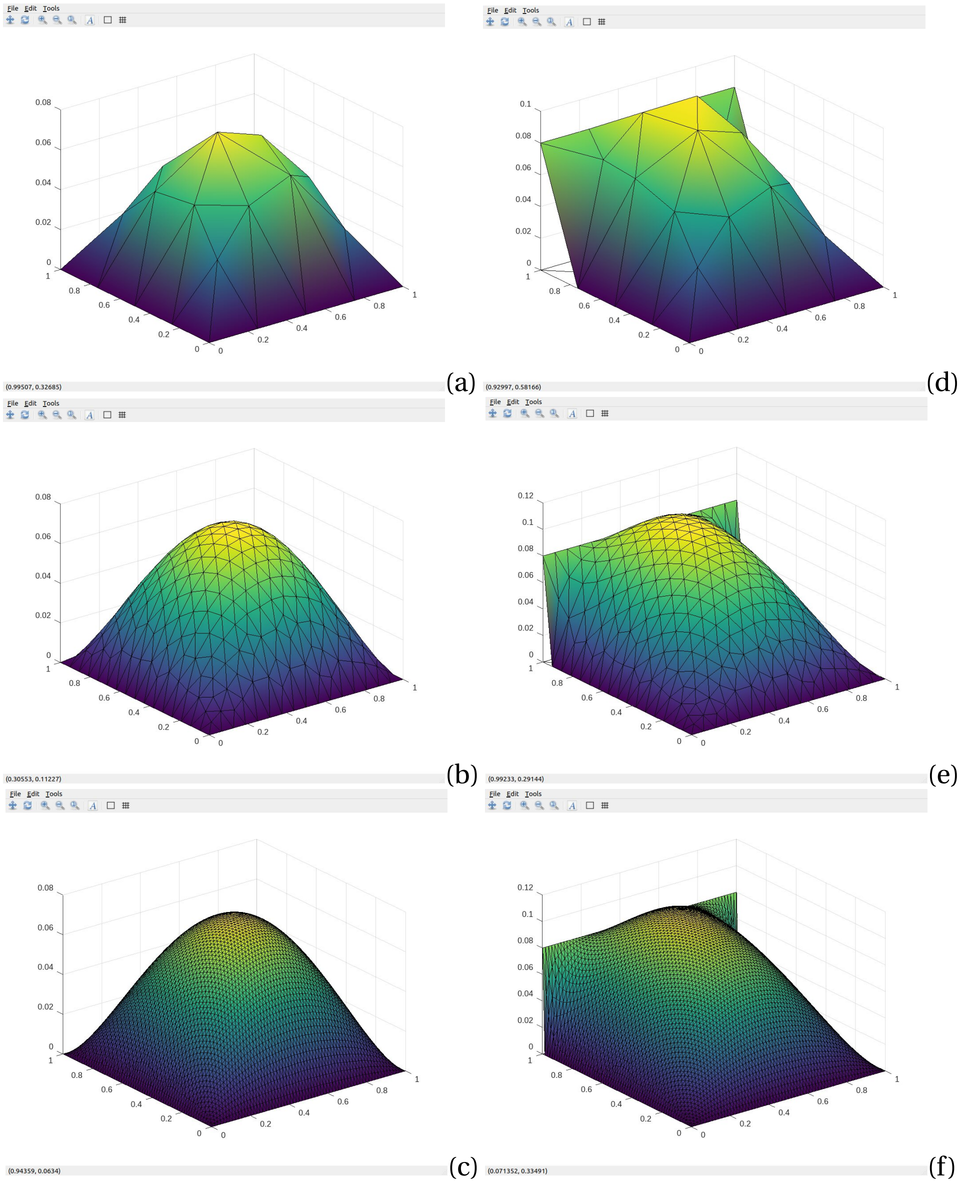
$$K = 1, \;\; c = 0, \;\; f = 1, \;\; \partial\Omega_D = \partial\Omega$$

and the function g equal to δ on the upper side and zero elsewhere. We use the P_1 Lagrange finite element code of Chapter 2.

We will discuss two values of δ, namely $\delta = 0$ and $\delta = 0.08$. The corresponding numerical solutions for meshes $M1$, $M3$ and $M5$ are shown in Fig. 3.5.

By direct inspection of Fig. 3.5 we can argue that the numerical solutions seem to converge to some smooth function u which, if the method makes sense, must be the exact solution of the PDE.

Numerical solutions are computed so as to estimate some quantity of interest $Q(u)$ by its approximation $Q(u_h)$. The method is useful (for this quantity) if $\lim_{h\to 0} Q(u_h) = Q(u)$. For this to happen, it is first necessary that the sequence $Q(u_h)$ converges to something, and this is what we are going to check.



(a) (d) (b) (e) (c) (f)

Figure 3.5 Example.

Let us consider the following quantities of interest:

$$\begin{aligned}
Q_1(u_h) &= \int_\Omega |u_h(x)|\,d\Omega, && (3.70)\\
Q_2(u_h) &= \left[\int_\Omega |u_h(x)|^2\,d\Omega\right]^{\frac{1}{2}}, && (3.71)\\
Q_3(u_h) &= \sup_{x\in\Omega} |u_h(x)|, && (3.72)\\
Q_4(u_h) &= \int_\Omega \|\nabla u_h(x)\|\,d\Omega, && (3.73)\\
Q_5(u_h) &= \left[\int_\Omega \|\nabla u_h(x)\|^2\,d\Omega\right]^{\frac{1}{2}}, && (3.74)\\
Q_6(u_h) &= \sup_{x\in\Omega} \|\nabla u_h(x)\|. && (3.75)
\end{aligned}$$

Quantities $Q_1 - Q_3$ correspond to $\|u_h\|_{L^p(\Omega)}$, with $p = 1, 2$ and ∞. Quantities $Q_4 - Q_6$ are analogous for $\|\nabla u_h\|_{L^p(\Omega)}$.

The values obtained for the five meshes, in the case $\delta = 0$, were the following:

$\delta = 0$	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6
M1 ($h = 1/4$)	0.03242	0.001497	0.07488	0.17069	0.18006	0.24431
M2 ($h = 1/8$)	0.03431	0.001641	0.07259	0.17311	0.18524	0.29084
M3 ($h = 1/16$)	0.03491	0.001685	0.07339	0.17412	0.18685	0.31111
M4 ($h = 1/32$)	0.03508	0.001698	0.07363	0.17442	0.18730	0.32606
M5 ($h = 1/64$)	0.03512	0.001701	0.07365	0.17450	0.18742	0.33106
convergence?	YES	YES	YES	YES	YES	YES

In the last row we show the result of a simplified analysis of the sequence of numbers to predict whether it is converging to something or not. It assumes that $Q(h)$ behaves as $q + ch^\beta$ and uses the computed values for the smallest 3 values of h to infer q, c and β. If $\beta > 0$ we say that the sequence is converging.

When $\delta = 0$ the boundary condition is continuous along the boundary. It is observed that all the computed quantities converge, suggesting that Q_1-Q_6 are well defined for the exact solution u.

Of course not all quantities converge with equal speed. It is evident from the table that Q_6 converges much more slowly than all the others.

Let us now turn to the case $\delta = 0.08$. In this case the boundary condition is discontinuous along $\partial\Omega$. **It can be proved that the exact solution to this problem exists and is unique, but it does not belong to** $H^1(\Omega)$, namely, $\|u\|_1 = +\infty$.

From the practical viewpoint, nothing happens. We simply change the imposed value of u_h at some boundary nodes. This does not change the system matrix and thus the linear system is well posed and u_h perfectly defined.

From the theoretical viewpoint, on the other hand, we cannot apply Céa's lemma 3.1 and thus we do not know whether u_h converges or not to u, or in which norm. Most importantly, the quantities of interest may exhibit different behaviors. The table below shows $Q_1 - Q_6$ for several meshes to take a look at what happens.

$\delta = 0.08$	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6
M1 ($h = 1/4$)	5.458e-02	6.240e-02	9.962e-02	1.933e-01	2.197e-01	4.224e-01
M2 ($h = 1/8$)	5.494e-02	6.273e-02	1.003e-01	2.050e-01	2.361e-01	7.366e-01
M3 ($h = 1/16$)	5.510e-02	6.288e-02	1.007e-01	2.096e-01	2.490e-01	1.407e+00
M4 ($h = 1/32$)	5.513e-02	6.292e-02	1.008e-01	2.116e-01	2.604e-01	2.764e+00
M5 ($h = 1/64$)	5.514e-02	6.293e-02	1.008e-01	2.125e-01	2.711e-01	5.513e+00
convergence?	YES	YES	YES	YES	**NO**	**NO**

The quantities $Q_1 - Q_4$ are convergent as in the previous case. This is numerical evidence that u belongs to $L^p(\Omega)$ for $p = 1$, 2 and ∞ and that ∇u belongs to $L^1(\Omega)$. As expected since $Q_5(u) = +\infty$, $Q_5(u_h)$ slowly diverges as $h \to 0$. The values obtained for any h are thus meaningless. One should not confuse the small changes in Q_5 with "mesh convergence" and erroneously infer that "$Q_5(u)$ must be something close to 0.3". The quantity Q_6 diverges severely.

3.6 Summary

- A **finite element method** is defined by a bilinear form $a(\cdot,\cdot)$, a linear form $\ell(\cdot)$, a finite element space $\mathscr{W}_h$ and a treatment of the essential boundary conditions (by interpolation, typically).

- The finite element solution u_h is defined by (3.2). The coercivity condition (3.6), which can easily be checked, guarantees that the linear system is well posed and thus u_h exists and is unique. The condition is not necessary, in fact a weaker version of coercivity is sufficient for existence and uniqueness.

- For u_h to approximate the exact solution u it is **necessary** that the method is **consistent**. This means that the residual $a(u, v_h) - \ell(v_h)$ is identically zero or at least sufficiently small. This condition is usually easy to check and allows to identify the required continuity of functions in $\mathscr{W}_h$ across inter-element boundaries.

- Consider a sequence of meshes $\mathscr{T}_h$ for the problem domain, with $h \to 0$. The corresponding sequence $\{u_h\}$ of numerical solutions then converges to u in the sense of a norm $\|\cdot\|$. For this to happen, **continuity and coercivity** conditions must hold with respect to $\|\cdot\|$.

- All quantities of interest that are continuous with respect to $\|\cdot\|$ then automatically **converge**, and it thus makes sense to consider $Q(u_h)$ as an approximation to $Q(u)$. When h is small enough $|Q(u) - Q(u_h)|$ typically behave as $c\,h^{\beta}$, β being the order of the approximation.