From this,

$$
\begin{aligned}
v_h(L) &= v_h(\overline{x}) + \int_{\overline{x}}^{L} v_h'(x)\,dx \\
&= \frac{1}{L}\int_0^L v_h(x)\,dx + \int_{\overline{x}}^{L} v_h'(x)\,dx \\
&\leq \frac{1}{L}\|1\|_0\,\|v_h\|_0 + \int_0^L |v_h'(x)|\,dx \\
&\leq \frac{1}{L}\|1\|_0\,\|v_h\|_0 + \|1\|_0\|v_h'\|_0 \\
&\leq \left(\frac{1}{\sqrt{L}} + \sqrt{L}\right)^{\frac{1}{2}} \|v_h\|_1\,,
\end{aligned}
$$

where we have used that $\|1\|_0 = \sqrt{L}$. We have thus proved that, if f is square-integrable and L is finite, then $\ell(\cdot)$ is continuous in the H^1-norm, i.e.,

$$|\ell(v_h)| \leq m\,\|v_h\|_1\,, \qquad \forall\, v_h \in \mathscr{W}_h\,,$$

where $m = \|f\|_0 + k(L)\,d_L\left(\frac{1}{\sqrt{L}} + \sqrt{L}\right)^{\frac{1}{2}}$. **It is again important to notice that both M and m do not depend on the mesh.**

We have thus checked all the hypothesis of Thm. 3.1 and can thus infer the following theorem:

Theorem 3.2. *If the finite element space $\mathscr{W}_h$ consists of* **continuous functions** *that in addition are C^1 in each closed element, then the finite element solution u_h defined by* (3.15) *exists, is unique, and satisfies*

$$\|u - u_h\|_1 \leq C \min_{w_h \in \mathscr{S}_h} \|u - w_h\|_1 \tag{3.20}$$

with the constant C independent of the adopted mesh.

3.3.2 Convergence of the Finite Element Solution

We now know that (3.20) holds for the finite element solution of (3.13). The remaining question is: How does the right-hand side of (3.20) behave as the mesh is refined?

Let us consider the so-called P_k **Lagrange elements** in one dimension. Some of these elements were introduced in Chapter 1. The basis functions are C^0-continuous across the element boundaries, and inside each element K the basis functions span the space $P_k(K)$. The "simplest" finite element (cf. 1.4.1) space corresponds to the P_1 Lagrange element.

The P_k Lagrange elements (P_k-elements for short in what follows) are the most popular finite elements for second order elliptic problems in 1D.

A well known result of numerical analysis is the following **interpolation estimate**:

Theorem 3.3. *Let u be a C^{k+1} function in $\Omega = [0, L]$ such that $u(0) = g_0$ and let $\mathscr{W}_h$ be the P_k Lagrange 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 that depends on k but not on h or u such that*

$$E_1(\mathscr{S}_h, u) = \min_{w_h \in \mathscr{S}_h} \| u - w_h \|_1 \leq C_I\, h^k \, \| u^{(k+1)} \|_0 \, , \tag{3.21}$$

and

$$E_0(\mathscr{S}_h, u) = \min_{w_h \in \mathscr{S}_h} \| u - w_h \|_0 \leq C_I\, h^{k+1} \, \| u^{(k+1)} \|_0 \, , \tag{3.22}$$

where $u^{(k+1)} = d^{k+1} u / d x^{k+1}$. The minima of $\| u - w_h \|$ in the H^1 and L^2 norms, denoted by E_1 and E_0 above, will be referred to as **interpolation errors** *of the finite element space $\mathscr{S}_h$ in the respective norms.*

From the theorem we observe that:

- The order of the interpolation error is $\mathscr{O}(h^k)$ in the H^1-norm, assuming that $\| u^{(k+1)} \|_0 < \infty$.
- If the norm does not contain the derivative, the interpolation error for the function decreases more rapidly with h. The interpolation error is $\mathscr{O}(h^{k+1})$ in the L^2-norm.
- A function u whose $(k+1)$-th derivative is identically zero can be interpolated exactly. These are all polynomial functions of degree $\leq k$ over Ω.
- The magnitude of the interpolation error does not depend on how large the derivatives of order $\leq k$ are. Only the $(k+1)$-th derivative matters.

Combining (3.21) with (3.20) we arrive at the estimate that we were looking for.

Corollary 3.1. *Under the hypotheses of Theorems 3.2 and 3.3, when h is small enough the finite element solution u_h built with P_k-elements satisfies, for some constant C depending on k but not on u or h,*

$$\| u - u_h \|_1 \leq C\, h^k \, \| u^{(k+1)} \|_0 \, . \tag{3.23}$$

Notice that the function and its k first derivatives need to be continuous for the theorem to apply, and the $(k+1)$-th derivative needs to be square-integrable.

We have thus proved that, using P_k-elements, u_h indeed converges to u in the H^1-norm and that the order of convergence is $\mathscr{O}(h^k)$.

We conclude this section with a partial proof of Theorem 3.3.

Proof. For brevity, we restrict the proof to the case $k = 1$. Consider a one-dimensional finite element mesh of P_1-elements over a domain $\Omega = [0, L]$, and a smooth function $u \colon [0, L] \to \mathbb{R}$. The mesh has n_{el} elements with domain $\Omega_e = [x_e, x_{e+1}]$ and size $h_e = x_{e+1} - x_e > 0$ for $e = 1, \ldots, n_{\mathrm{el}}$, with $x_1 = 0$ and $x_{n_{\mathrm{el}}+1} = L$. The number of nodes is $m = n_{\mathrm{el}} + 1$.

Let $\{N_1, \ldots, N_m\}$ be the hat functions over this mesh and $W_h = \mathrm{span}\{N_1, N_2, \ldots, N_m\}$ the corresponding finite element space.

Definition 3.2. *We define the* **Lagrange finite element interpolant** *(or* **Lagrange interpolant** *for short) $\mathscr{I} u$ of u as*

$$\mathscr{I} u = \sum_{a=1}^{m} u(x_a) N_a. \tag{3.24}$$

More generally, the Lagrange interpolant $\mathscr{I} u$ is the unique function of $\mathscr{W}_h$ that has as components the values of u at the nodes of the mesh.

Notice that $\mathscr{I} u \in \mathscr{S}_h$, since $\mathscr{I} u \in \mathscr{W}_h$ and

$$\mathscr{I} u(0) = \sum_a u(x_a) N_a(0) = u(x_1) N_1(0) = u(0) N_1(0) = u(0) = g_0 \,.$$

We will estimate an upper bound for the *interpolation* errors $\|u - \mathscr{I} u\|_0$ and $\|u - \mathscr{I} u\|_1$ over $[0, L]$. To this end, we will split the errors as a sum of local errors over each element, namely,

$$\begin{aligned} \|u - \mathscr{I} u\|_0^2 &= \int_0^L |u - \mathscr{I} u|^2 \, dx \\ &= \sum_{e=1}^{n_{\mathrm{el}}} \|u - \mathscr{I} u\|_{0,\Omega_e}^2 \end{aligned} \tag{3.25}$$

and

$$\begin{aligned} \|u - \mathscr{I} u\|_1^2 &= \int_0^L \left(|u - \mathscr{I} u|^2 + |u' - \mathscr{I} u'|^2\right) dx \\ &= \|u - \mathscr{I} u\|_0^2 + \sum_{e=1}^{n_{\mathrm{el}}} \|u' - \mathscr{I} u'\|_{0,\Omega_e}^2 \,. \end{aligned} \tag{3.26}$$

Let's see then how the error over an element can be obtained. Over element e consider the error function between u and $\mathscr{I} u$,

$$\eta(x) = u(x) - \mathscr{I} u(x) = u(x) - \left[u(x_e) N_1^e(x) + u(x_{e+1}) N_2^e(x)\right].$$

It satisfies that

(a) $\eta(x_e) = \eta(x_{e+1}) = 0$,

(b) for $x \in (x_e, x_{e+1})$ its derivative is

$$\eta'(x) = u'(x) - \mathscr{I} u'(x) = u'(x) - \frac{u(x_{e+1}) - u(x_e)}{h_e} \text{ and}$$

(c) its second derivative coincides with that of u, namely,

$$\eta''(x) = u''(x). \tag{3.27}$$

Since u is assumed smooth, because of (a) and of Rolle's theorem there exists $\xi \in (x_e, x_{e+1})$ such that $\eta'(\xi) = 0$. As a consequence,

$$\begin{aligned} |\eta'(x)| &= \left|\int_\xi^x \eta''(y)\,dy\right| \\ &= \left|\int_\xi^x u''(y)\,dy\right| \\ &\leq \left[\int_\xi^x u''(y)^2\,dy\right]^{\frac{1}{2}} \left[\int_\xi^x 1^2\,dy\right]^{\frac{1}{2}} \\ &\leq h_e^{\frac{1}{2}} \|u''\|_{0,\Omega_e} \end{aligned} \tag{3.28}$$

for all $x \in \Omega_e$. Notice now that

$$|\eta(x)| = \left|\int_{x_e}^x \eta'(y)\,dy\right| \leq h_e \max_{y\in\Omega_e} |\eta'(y)| \leq h_e^{\frac{3}{2}} \|u''\|_{0,\Omega_e}$$

which implies that

$$\|u - \mathscr{I}u\|_{0,\Omega_e} = \left[\int_{x_e}^{x_{e+1}} \eta(x)^2\,dx\right]^{\frac{1}{2}} \leq h_e^2 \|u''\|_{0,\Omega_e}\,.$$

Summing over the elements we get

$$\begin{aligned} \|u - \mathscr{I}u\|_0^2 = \sum_{e=1}^{n_{\mathrm{el}}} \|u - \mathscr{I}u\|_{0,\Omega_e}^2 \leq \sum_e h_e^4 \|u''\|_{0,\Omega_e}^2 \\ \leq h^4 \sum_e \|u''\|_{0,\Omega_e}^2 = h^4 \|u''\|_0^2. \end{aligned} \tag{3.29}$$

and thus, since $\mathscr{I}u$ belongs to $\mathscr{W}_h$, (3.22) is proved with $C_I = 1$.

To prove (3.21) we use (3.28) to see that

$$\begin{aligned} \|u' - \mathscr{I}u'\|_{0,\Omega_e} &= \left[\int_{x_e}^{x_{e+1}} \eta'(x)^2\,dx\right]^{\frac{1}{2}} \\ &\leq h_e^{\frac{1}{2}} \|u''\|_{0,\Omega_e} \left[\int_{x_e}^{x_{e+1}} 1^2\,dx\right]^{\frac{1}{2}} \\ &= h_e \|u''\|_{0,\Omega_e}\,. \end{aligned}$$

Then,

$$\sum_{e=1}^{n_{\mathrm{el}}} \|u' - \mathscr{I}u'\|_{0,\Omega_e}^2 \leq \sum_e h_e^2 \|u''\|_{0,\Omega_e}^2 \leq \left(\max_e h_e^2\right) \sum_e \|u''\|_{0,\Omega_e}^2 = h^2 \|u''\|_0^2\,. \tag{3.30}$$

Inserting (3.30) and (3.29) into (3.26) we get

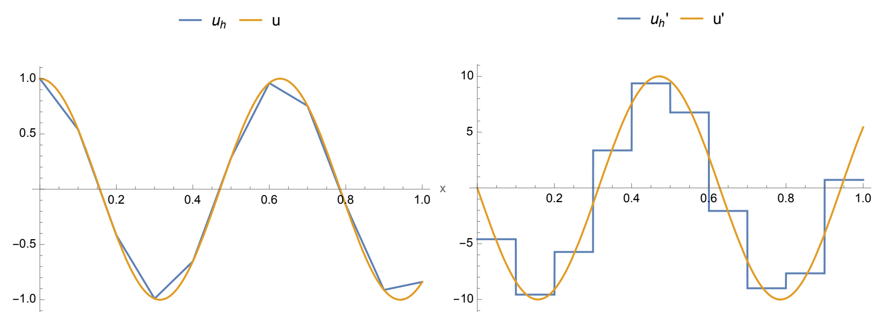
$$\|u - \mathscr{I}u\|_1 \leq \left(h^4 + h^2\right)^{\frac{1}{2}} \|u''\|_0\,.$$

At this point we must only assume that the mesh is fine enough, more specifically that $h \le 1$ so that $h^4 \le h^2$ and thus $\left(h^4 + h^2\right)^{\frac{1}{2}} \le \sqrt{2}\,h^2$, to finalize the proof of (3.21) with $C_I = \sqrt{2}$.

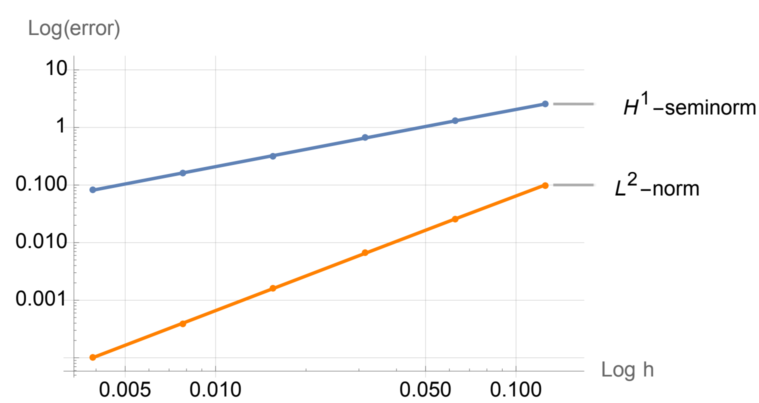
Since we have obtained two possible values (1 and $\sqrt{2}$) for C_I, we conclude that the theorem holds with the greater value, i.e., $C_I = \sqrt{2}$.

□

Example 3.1 Consider a mesh of $n_{\text{el}} \in \mathbb{N}$ equal-length P^1-elements over a domain $\Omega = [0,1]$, with $h = 1/n_{\text{el}}$, and $u(x) = \cos(10x) \in C^2([0,1])$. The interpolant and its first derivative are plotted below for $n_{\text{el}} = 11$.



The L^2-norms of the error $u - \mathscr{I}u$ and its derivative $u' - Iu'$ (termed the H^1-seminorm) are plotted next, in a log-log scale. Why a log-log plot? Because if we compute the log on both sides of a relation of the form error$\sim Ch^r$, we get that log(error) $\sim \log C + r \log h$, so log(error) decreases linearly with $\log h$, and the slope of the line is equal to the order r.



Notice that the L^2-error decreases 2 units in the vertical axis per unit in the horizontal axis, a slope equal to 2, in agreement with the expected order of convergence in (3.29). Similarly, the H^1-seminorm decreases 1 unit in the vertical axis per unit in the horizontal one, indicating a slope equal to

1, in agreement with the first order of convergence in (3.30). The square of the H^1-norm is the sum of the squares of the two curves, and when plotted, it essentially overlaps with the curve of the H^1-seminorm, so it was not plotted. A quick inspection reveals that at the largest value of h, the L^2-norm of the error is already more than 10 times smaller than the H^1-seminorm, so when the two squared values are added, the error on the value of the function contributes less than 1% to the sum. The H^1-error is dominated by the error in the derivative.

3.3.3 Consequences of the Convergence in the H^1-norm

Let us go back to the estimate proved in the previous section:

$$\|u - u_h\|_1 \le C\, h^k \, \|u^{(k+1)}\|_0 \, . \tag{3.31}$$

From it, we can derive several useful consequences:

1. For any $x \in \Omega$,

$$|u(x) - u_h(x)| \le \left| \int_0^x (u'(s) - u_h'(s))\, ds \right| \le \|u' - u_h'\|_0 \, L^{\frac{1}{2}} \le C(u)\, h^k \, . \tag{3.32}$$

   In other words, u_h converges **uniformly** in Ω to u. At any point x, $u_h(x)$ converges to $u(x)$ with order *at least* equal to $\mathcal{O}(h^k)$. In fact, the order is in general greater than k. This tells us that $u_h(x)$ is a **convergent approximation** of $u(x)$.

2. Assume that the problem represents a one-dimensional elastic problem, such as the vertical loading of a column by its own weight. In this case we would have $c(x) = 0$, $g_0 = 0$, $f(x) = -\rho A(x)\, g$ and $d_L = 0$, and u would represent the vertical displacement of the section at height x. The symbol ρ denotes de density, $A(x)$ the cross-sectional area at x, and g the acceleration of gravity. The elastic energy of such bar is given by

$$U(u) = \frac{1}{2} a(u, u), \tag{3.33}$$

   and it is important that $U(u_h)$ converges to $U(u)$ as $h \to 0$. This is readily obtained using the symmetry of $a(\cdot, \cdot)$,

$$\begin{aligned} 2U(u) - 2U(u_h) &= a(u,u) - a(u_h, u_h) - a(u, u_h) + a(u_h, u) \\ &= a(u - u_h, u - u_h) \\ &\le M \|u - u_h\|_1^2 \\ &\le M C^2\, h^{2k} \, \|u^{(k+1)}\|_0^2 \, , \end{aligned}$$

   which indeed guarantees convergence of the energy with order $\mathcal{O}(h^{2k})$.

3. Any integral of the form

$$I(u) = \int_\Omega g(u(x), u'(x))\, dx\,, \tag{3.34}$$

where g is twice continuously differentiable in its two variables, is approximated by $I(u_h)$ with order at least k but in general higher, as high as $\mathcal{O}(h^{2k})$.

Consider the example above of a column under its own weight. The change in gravitational energy of the column due to its deformation $u(x)$ is given by

$$G(u) = \int_\Omega \rho\, g\, A(x)\, u(x)\, dx\,. \tag{3.35}$$

This quantity is approximated by $G(u_h)$, in fact

$$G(u) - G(u_h) = \mathcal{O}(h^{2k})\,.$$

4. The convergence of $\|u - u_h\|_0$ to zero is also immediate from (3.31), since

$$\|u - u_h\|_0 \le \|u - u_h\|_1 \le C\, h^k\, \|u^{(k+1)}\|_0\,.$$

Notice that this proves convergence with order $\mathcal{O}(h^k)$. However, we know from (3.22) that it is possible to approximate u in the L^2-norm with order $\mathcal{O}(h^{k+1})$. The estimation above is thus **suboptimal**. But it turns out to be just a flaw of our demonstration. Using an elegant argument known as **Aubin-Nitsche trick** it can be proved that

$$\|u - u_h\|_0 \le C\, h^{k+1}\, \|u^{(k+1)}\|_0\,. \tag{3.36}$$

The finite element solution u_h converges with optimal order in the L^2-norm.

3.3.4 An Example of Numerical Convergence

Let us use the code developed in Chapter 1 to solve a particular case of Problem 1.1 for which we know the exact solution. This will allow us to confirm the predictions of our mathematical estimates.

A popular way to assess convergence is by constructing a **manufactured solution**. Here we will select the domain $[0,1]$, the coefficients

$$k(x) = 1, \qquad b(x) = 0, \qquad c(x) = 1$$

and the exact solution

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

The trick is simply to set the source term $f(x)$ and the boundary conditions g_0 and d_L so that u solves the problem.

Noticing that

$$(ku')' = 2\alpha\left(\cos(\alpha x^2) - 2\alpha x^2 \sin(\alpha x^2)\right)$$

we compute that to satisfy the differential equation we need to set

$$f(x) = -2\alpha\left(\cos(\alpha x^2) - 2\alpha x^2 \sin(\alpha x^2)\right) + \sin(\alpha x^2),$$

$g_0 = u(x=0) = 0$, and

$$d_L = u'(x=L) = 2\alpha x \cos(\alpha x^2)\big|_{x=1} = 2\alpha\,\cos(\alpha)\ .$$

We select $\alpha = 20$.

With these data, we can readily use our code. Let us begin with a uniform mesh of 40 P_1 Lagrange elements. In Figure 3.1 we show the exact solution u, the finite element solution u_h and the P_1 Lagrange interpolant $\mathscr{I}u$, together with their derivatives. The interpolant $\mathscr{I}u$ coincides with u at the nodes. In part (c) of the figure we can see the interpolation error $u - \mathscr{I}u$, which has a parabola-like shape inside each element and is larger at elements with larger $|u''|$.

The approximation error $u - u_h$ (also shown in Figure 3.1(c)) can be seen as the sum of the interpolation error $u - \mathscr{I}u$ plus a numerical error $\mathscr{I}u - u_h$ that belongs to $\mathcal{V}_h$. Céa's lemma (Thm. 3.1) tells us something far from obvious: That the numerical error is bounded, up to a constant, by the interpolation error. In this case, the error $u' - u_h'$ is very close to $u' - (\mathscr{I}u)'$, so that the approximation error comes mostly from interpolation.

The errors in u and u', for a mesh with twice the number of elements, are shown in Figure 3.2. Notice from the vertical scales that the error in the function roughly gets divided by four with respect to the 40-element case, and the error in the derivative gets divided by two.

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

Quantity	$h = 1/40$	$h = 1/80$	$h = 1/160$	$h = 1/320$	Order
$\lVert u - u_h \rVert_0$	0.3429	0.08327	0.02067	5.158E-3	$\mathscr{O}(h^2)$
$\lVert u - \mathscr{I}u \rVert_0$	0.0277	7.022E-3	1.762E-3	4.409E-4	$\mathscr{O}(h^2)$
$\lVert u' - u_h' \rVert_0$	3.6159	1.7920	0.8947	0.4477	$\mathscr{O}(h)$
$\lVert u' - \mathscr{I}u' \rVert_0$	3.5174	1.7805	0.8932	0.4475	$\mathscr{O}(h)$

As expected, many other quantities also converge as well. Here are some examples:

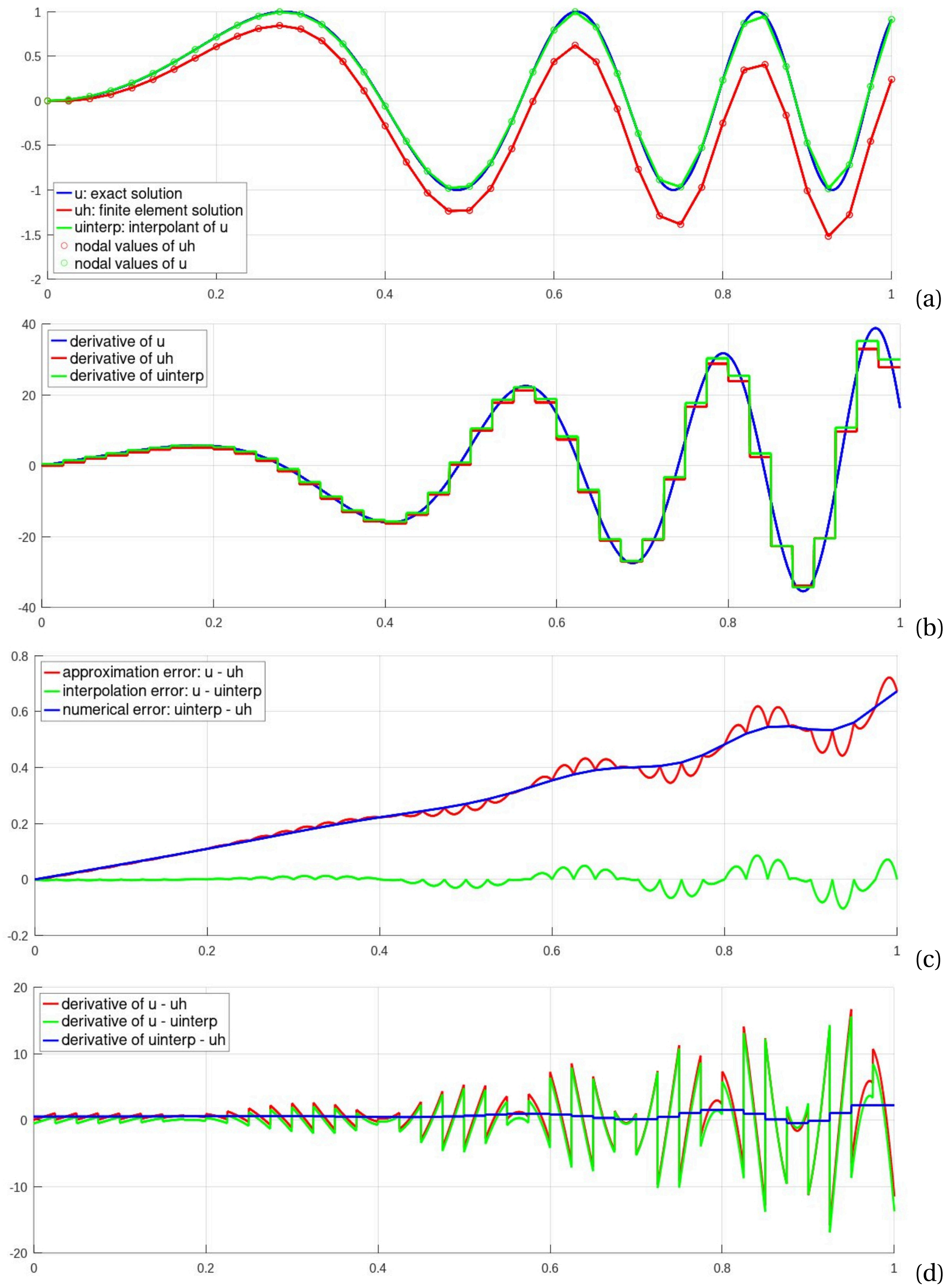


Figure 3.1 Example of numerical convergence. Results on a mesh of 40 elements, all of size $h = 1/40$. (a) Exact solution u, finite element solution u_h and interpolant of exact solution $\mathscr{I}u$. (b) Derivatives of u, u_h and $\mathscr{I}u$. (c) Approximation error $u - u_h$ and interpolation error $u - \mathscr{I}u$. (d) Errors in the derivative.

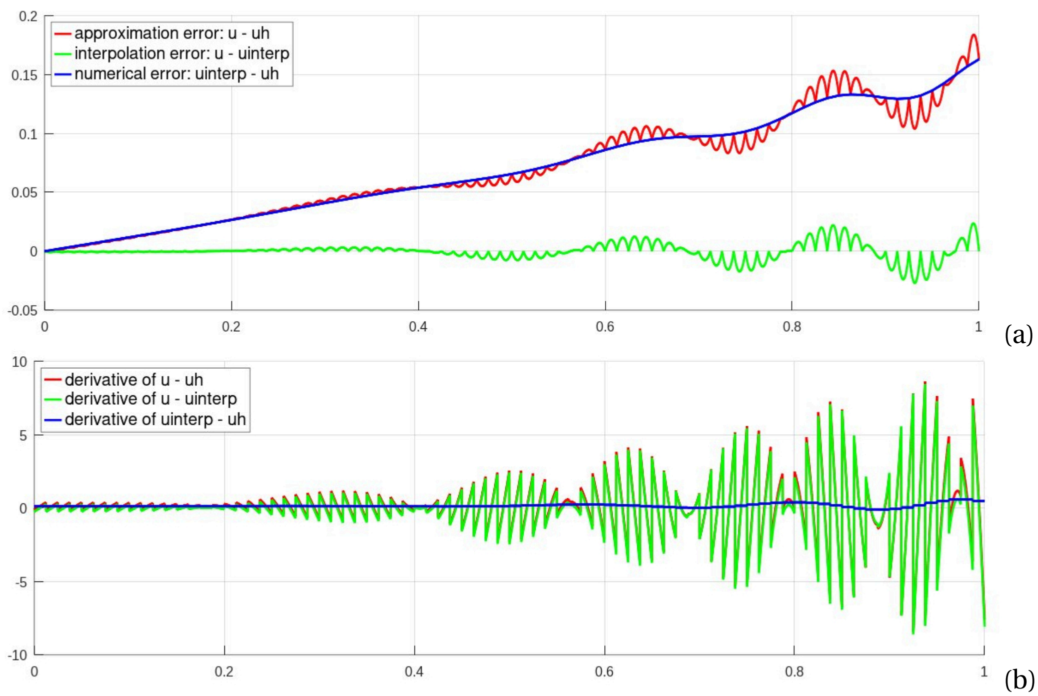


Figure 3.2 Example of numerical convergence. Results on a mesh of 80 elements, all of size $h = 1/80$. (a) Approximation error $u - u_h$ and interpolation error $u - \mathscr{I}u$. (b) Errors in the derivative. Notice the change in the vertical scales with respect to the previous figure.