1.4 The Finite Element Method

By now we have learned about how to construct a variational equation from the differential equation and boundary conditions of a problem, and to formulate a variational method to obtain an approximation of the solution. This last step relied on the construction of a trial and a test space in which to seek the approximate solution. We next look at how Finite Elements provide a systematic way to construct such spaces.

A Finite Element Method (FEM) is obtained by seeking the solution of a variational method in trial and test spaces constructed with finite elements.

We will also describe *how* to compute the stiffness matrix and load vector for a finite element method. The way this computation is performed, called **assembly**, is a distinctive feature and a virtue of the Finite Element method, since it can be done very efficiently in a computer.

1.4.1 The Simplest C^0 Finite Element Space

We show a first example of the construction of a variational method with the simplest finite element space of continuous functions. We do this for the problem in Example 1.55, so that we can contrast the use of variational methods with and without finite element spaces. In this example, we seek an approximation to the problem of finding $u\colon [0,1] \to \mathbb{R}$ such that

$$-u''(x) = 1 \qquad x \in (0,1) \tag{1.96a}$$

$$u(0) = 2 \tag{1.96b}$$

$$u'(L) = 0, \tag{1.96c}$$

using the following variational equation that u satisfies,

$$\int_0^1 u'(x)v'(x)\,dx = \int_0^1 v(x)\,dx \tag{1.97}$$

for all $v \in \mathcal{V} = \{w\colon [0,1] \to \mathbb{R} \text{ smooth } \mid w(0) = 0\}$.
Steps:

1. **Build the mesh of the domain.** Let the domain of the problem be the interval $\Omega = [c,d]$. We partition the domain into $n_{\text{el}} \in \mathbb{N}$ intervals by selecting $\{x_i\}_{i=1,\ldots,n_{\text{el}}+1}$ such that

$$c = x_1 < \ldots < x_{n_{\text{el}}+1} = d. \tag{1.98}$$

   Each point x_i is a **vertex**, and i is its **vertex number**. Interval $[x_i, x_{i+1}]$ is called *element* i, for $i = 1,\ldots,n_{\text{el}}$. Strictly speaking, this is the domain of the element, but it is common to refer to the domain of the element simply as "element." The collection of nodes and elements is the **mesh**; we shall give a more complete definition of the mesh when we look at 2D problems.

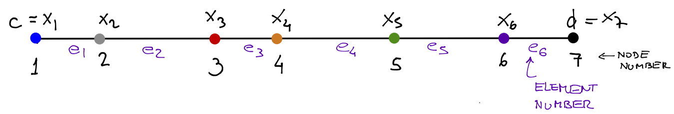


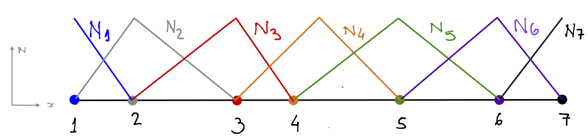
Figure 1.9

For our example, we choose a uniform mesh, so $c = 0, d = 1$, and $x_a = (a-1)/n_{\text{el}}$ for $i = 1, \ldots, n_{\text{el}} + 1$.

2. **Build basis functions.** For this example, we build the so-called continuous "piecewise affine" elements, or hat functions. These functions have domain $[c, d]$, and for $a = 1, \ldots, n_{\text{el}} + 1$ are defined as

$$N_a(x) = \begin{cases} 0 & x < x_{a-1} \\ \frac{x - x_{a-1}}{x_a - x_{a-1}} & x_{a-1} \leq x < x_a \\ 1 & x = x_a \\ \frac{x_{a+1} - x}{x_{a+1} - x_a} & x_a < x \leq x_{a+1} \\ 0 & x_{a+1} < x \end{cases} \tag{1.99}$$
$$= \max\left[0, \min\left(\frac{x - x_{a-1}}{x_a - x_{a-1}}, \frac{x - x_{a+1}}{x_a - x_{a+1}}\right)\right].$$

Notice that when $a = 1$, $x \in [c, d]$ implies that we only have the case $x \geq x_a$ [8]. Similarly, when $a = n_{\text{el}} + 1$, $x \in [c, d]$ implies that we only have the case $x \leq x_a$[9]. These functions are plotted below.



The Finite Element space is $\mathscr{W}_h = \text{span}(N_1, \ldots, N_{n_{\text{el}}+1}\})$, so for this example, $m = n_{\text{el}} + 1$. Some properties that will be generalized later to many other shape functions:

(a) You can check that all of the add up to 1, i.e., $\sum_{a=1}^{n_{\text{el}}+1} N_a(x) = 1$ for $x \in [c, d]$.

A simple way to see this, is to notice that in each element e, the two non-zero functions N_e and N_{e+1} are affine, and hence their sum is affine. But $N_e + N_{e+1}$ is equal to 1 at $x = x_e$ and $x = x_{e+1}$, and hence their sum is the only affine function that is equal to 1 at both locations: this is the constant function equal to 1.

[8] We do not know what x_{a-1} is in this case, nor do we need it.
[9] We do not know what x_{a+1} is in this case, but again, we do not need it.

(b) Notice that $N_b(x_a) = \delta_{ba}$. This is a particular version of a more general property we will see later in the class, and it has the following neat consequence. A function $w_h \in \mathscr{W}_h$ can be written as $w_h = w_1 N_1 + w_2 N_2 + \ldots + w_{n_{\text{el}}+1} N_{n_{\text{el}}+1}$, where $w_1, \ldots, w_{n_{\text{el}}+1}$ are the components of w_h in the basis. At the same time, $w_a = w_h(x_a)$ for $a = 1, \ldots, n_{\text{el}} + 1$, that is, the component w_a is the value of the function w_h at x_a. This follows because

$$w_h(x_a) = w_1 \underbrace{N_1(x_a)}_{=\delta_{1a}} + \ldots + w_b \underbrace{N_b(x_a)}_{=\delta_{ba}} + \ldots + w_{n_{\text{el}}+1} \underbrace{N(x_a)}_{=\delta_{(n_{\text{el}}+1)a}} = w_a,$$

so it is a special property of the basis we chose for $\mathscr{W}_h$.

Had we chosen the basis $\{N_1 + N_2, N_2, \ldots, N_{n_{\text{el}}+1}\}$, for example, then $w_h(x_2) = w_1 + w_2$, and in this case w_2 does not necessarily coincide with the value of w_h at x_2.

To indicate that the degrees of freedom $\{w_1, \ldots, w_{n_{\text{el}}+1}\}$ of the function w_h in the basis $\{N_1, \ldots, N_{n_{\text{el}}+1}\}$ are precisely the values of w_h at each vertex x_a, we say that there is a **node** of the finite element space at each vertex of this mesh, and graphically depict it with a filled disk at the vertex; see Fig. 1.9.

(c) Notice that $N_a(x) \neq 0$ only in a small part of the domain. This is normally referred to by saying that the basis functions have “compact support.” In the Finite Element context, this (generally) means that basis functions are non-zero in at most one element and its neighbors.

In this case we defined the space $\mathscr{W}_h$ as the span of a set of basis functions. Alternatively, it could have been defined as

$$\mathscr{W}_h = \{w_h \colon [c, d] \to \mathbb{R} \text{ continuous} \mid w_h \text{ is affine on each element } e\}. \tag{1.100}$$

This space would often be referred to as the "space of piecewise affine functions over $[c, d]$," with the tacit understanding that functions would be affine over each element. Both definitions are equivalent; it is simple to see that functions in $\text{span}(\{N_1, \ldots, N_{n_{\text{el}}+1}\})$ are piecewise affine, and that any piecewise affine function can be expressed as a linear combination of functions in the basis $\{N_1, \ldots, N_{n_{\text{el}}+1}\}$.

3. **Build $\mathscr{V}_h$ and $\mathscr{S}_h$.** Collect essential boundary conditions and impose them on functions in $\mathscr{W}_h$ to obtain $\mathscr{S}_h$. The space $\mathscr{V}_h$ follows as the direction of $\mathscr{S}_h$. In our example,

$$\mathscr{S}_h = \{u_h \in \mathscr{W}_h \mid u_h(0) = 2\},$$
$$\mathscr{V}_h = \{v_h \in \mathscr{W}_h \mid v_h(0) = 0\}.$$

To find a basis for $\mathscr{V}_h$, notice that for any $v_h = v_1 N_1 + \ldots, + v_{n_{\text{el}}+1} N_{n_{\text{el}}+1} \in \mathscr{W}_h$, $v_h(0) = 0$ if and only if $v_1 = 0$. Similarly, any function $u_h \in \mathscr{W}_h$ satisfies that

$u_h(0) = 2$ if and only if $u_1 = 2$. Then, in terms of the basis functions in $\mathcal{W}_h$, these spaces can be described by

$$\begin{aligned}
\mathcal{V}_h &= \{v_2 N_2 + \ldots + v_{n_{\text{el}}+1} N_{n_{\text{el}}+1} \mid v_2, \ldots, v_{n_{\text{el}}+1} \in \mathbb{R}\} \\
&= \operatorname{span}\left(\{N_2, \ldots, N_{n_{\text{el}}+1}\}\right). \\
\mathcal{S}_h &= \{u_h \in \mathcal{W}_h \mid u_1 = 2\} \\
&= \{2N_1 + v_h \mid v_h \in \mathcal{V}_h\}.
\end{aligned}$$

The index set η_a that identifies the basis functions for $\mathcal{V}_h$, and its complement η_g, are

$$\begin{aligned}
\eta_a &= \{2, \ldots, n_{\text{el}} + 1\} \\
\eta_g &= \{1\}.
\end{aligned}$$

Finally, we need to identify the components $\overline{u}_a$ for $a \in \eta_g$ to impose the fact that $u_h \in \mathcal{S}_h$. In this case, based on the description of $\mathcal{S}_h$ above, it is $\overline{u}_1 = 2$.

As an exercise, we can also identify a function $\overline{u}_h \in \mathcal{S}_h$. For example, we can choose $\overline{u}_h = 2N_1$, which gives $\overline{u}_1 = 2$. Alternatively, we can set $\overline{u}_h = 2 = 2\sum_{a \in \eta} N_a$, which is also in $\mathcal{S}_h$ because it is in $\mathcal{W}_h$ and $\overline{u}_h(0) = 2$.

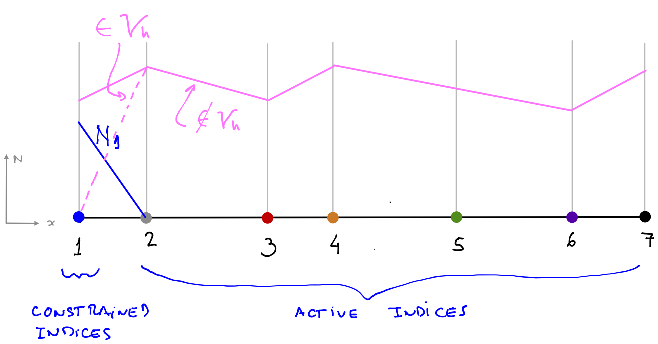


Figure 1.10 If $w_1 \neq 0$, then $w_h \notin \mathcal{V}_h$.

Notice that the constraint imposed on u_1 stems from the essential boundary condition (EBC). Should more EBC be present, more constraints should be imposed on functions in $\mathcal{W}_h$ to belong to $\mathcal{V}_h$ and $\mathcal{S}_h$. The EBC will also determine components of $\overline{u}_a$ for $a \in \eta_g$.

Not every constraint can be imposed by simply selecting a subset of a set of basis functions for $\mathcal{W}_h$, as we have assumed so far. For example, had we chosen the basis $\{N_1, N_1 + N_2, N_3, \ldots, N_{n_{\text{el}}+1}\}$ for $\mathcal{W}_h$, then $w_h(0) = w_1 + w_2$, and the condition for $w_h \in \mathcal{W}_h$ to belong to $\mathcal{V}_h$ is to have $w_1 + w_2 = 0$, see Fig. 1.11. For example, the function $w_h = N_1 + (N_2 - N_1)$ has $w_1 = 1$ and $w_2 = 1$ and is in $\mathcal{V}_h$. Therefore, just setting either w_1 or w_2 (or both) to zero does not lead to a basis for $\mathcal{V}_h$. It is not possible in this case to extract a subset of

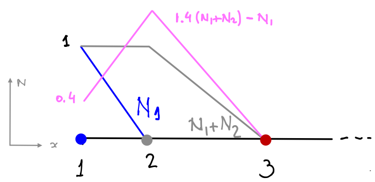


Figure 1.11 The function $w_h = 1.4(N_1 + N_2) - N_1$ has $w_1 = -1$ and $w_2 = 1.4$ in the basis $\{N_1, N_1 + N_2, \ldots\}$, and $w_h(0) = w_1 + w_2 = 0.4 \neq w_1$. Therefore, to impose the EBC at $x = 0$ in this basis it is not enough to solely set the value of w_1.

$\{N_1, N_1 + N_2, N_3, \ldots, N_{n_{\mathrm{el}}+1}\}$ to serve as a basis for $\mathscr{V}_h$[10]. As a result, it is not possible to define η_a or η_g.

In the finite element method, this type of situations need a different treatment (e.g., with Lagrange multipliers or Nitsche's method), and the most commonly used finite element bases are constructed so that essential boundary conditions can be imposed by setting the values of some components, such as $v_1 = 0$ here. In other words, in the finite element method it is common for the basis for $\mathscr{V}_h$ to be a subset of the basis for $\mathscr{W}_h$. This is going to be the case for the examples we will see.

Summarizing, in this step we:

(a) Identify active and constrained index sets, η_a and η_g (this assumes that the basis for $\mathscr{V}_h$ is a subset of the basis for $\mathscr{W}_h$).

(b) Build $\mathscr{V}_h = \operatorname{span}\left(\{N_a \in \{N_1, \ldots, N_m\} \mid a \in \eta_a\}\right)$.

(c) Identify $\overline{u}_h \in \mathscr{W}_h$ so that $\mathscr{S}_h = \{v + \overline{u}_h \mid v \in \mathscr{V}_h\} = \left\{v + \sum_{a \in \eta_g} \overline{u}_a N_a \mid v \in \mathscr{V}_h\right\}$.

4. **Compute** K **and** F**.** We proceed as we did earlier and compute the stiffness matrix and load vector. We compute its entries according to (1.78).

For our example, we can set $h = 1/n_{\mathrm{el}}$ with $n_{\mathrm{el}} = 5$ and hence $m = 6$,

$$\ell(N_a) = \int_0^1 1 \cdot N_a(x)\, dx = \begin{cases} \dfrac{h}{2} & a \in \{1, m\} \\ h & a \in \{2, \ldots, m-1\}. \end{cases}$$

$$a(N_b, N_a) = \int_0^1 N_b'(x) N_a'(x)\, dx = \begin{cases} 0 & |a-b| > 1 \\ -\dfrac{1}{h} & |a-b| = 1 \\ \dfrac{2}{h} & a = b \in \{2, \ldots, m-1\}. \\ \dfrac{1}{h} & a = b \in \{1, m\}. \end{cases}$$

[10] Of course, $\mathscr{V}_h$ has a basis, but it is not a subset of the chosen basis for $\mathscr{W}_h$.

The only index in η_g is 1. Therefore, according to (1.78c),

$$\begin{aligned}
K_{21} &= a(N_1, N_2) = -\frac{1}{h},\\
K_{12} &= \delta_{12} = 0,\\
K_{11} &= \delta_{11} = 1,\\
K_{22} &= a(N_2, N_2) = \frac{2}{h},\\
K_{23} &= a(N_3, N_2) = -\frac{1}{h},\\
K_{24} &= a(N_4, N_2) = 0,\\
K_{66} &= a(N_6, N_6) = \frac{1}{h},\\
F_1 &= \overline{u}_1 = 2,\\
F_5 &= \ell(N_5) = h,\\
F_6 &= \ell(N_6) = \frac{h}{2}.
\end{aligned}$$

We have not replaced $h = 1/m = 1/5$ yet, for clarity. In this case, the stiffness matrix and load vector are, now replacing $h = 1/5$,

$$K = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -5 & 10 & -5 & 0 & 0 & 0 \\ 0 & -5 & 10 & -5 & 0 & 0 \\ 0 & 0 & -5 & 10 & -5 & 0 \\ 0 & 0 & 0 & -5 & 10 & -5 \\ 0 & 0 & 0 & 0 & -5 & 5 \end{bmatrix} \qquad F = \begin{bmatrix} 2 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.1 \end{bmatrix}. \tag{1.101}$$

5. **Solve and Compute the Finite Element Solution.** We now solve the system $KU = F$, and then build the finite element solution as $u_h(x) = \sum_{a=1}^{m} u_a N_a(x)$.

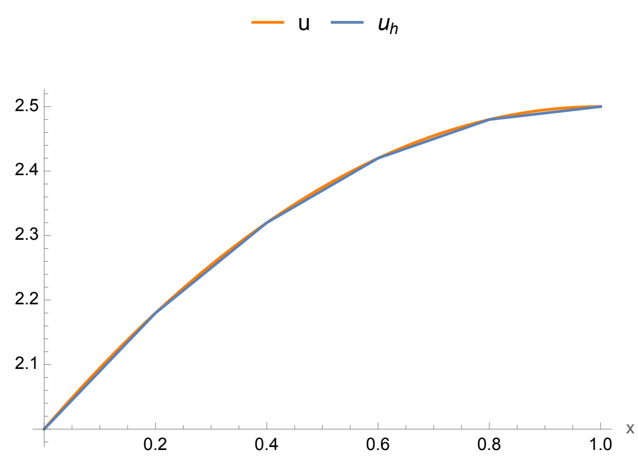
   For our example,

$$U = \begin{bmatrix} 2 \\ 2.18 \\ 2.32 \\ 2.42 \\ 2.48 \\ 2.5 \end{bmatrix}$$

   and hence

$$u_h(x) = 2N_1(x) + 2.18N_2(x) + 2.32N_3(x) + 2.42N_4(x) + 2.48N_5(x) + 2.5N_6(x).$$

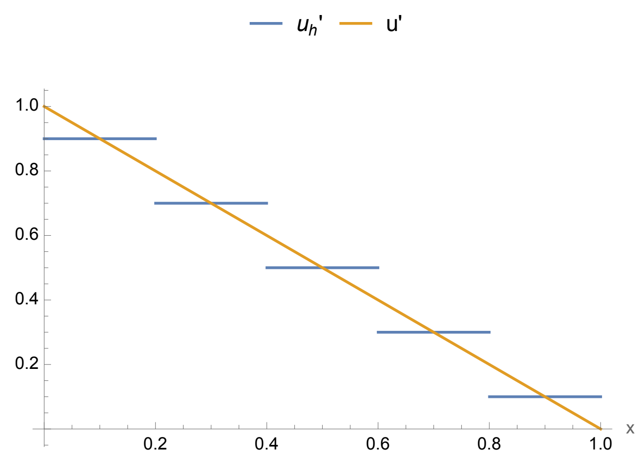
   This function is plotted below, together with the exact solution $u(x) = 2 + x - x^2/2$.



The derivative of u_h follows as

$$u_h'(x) = 2N_1'(x) + 2.18N_2'(x) + 2.32N_3'(x) + 2.42N_4'(x) + 2.48N_5'(x) + 2.5N_6'(x).$$

The derivative of the finite element approximation is a piecewise constant function, a fact reflected in its graph, shown below.



In staring at this graph, you may come to the realization that the finite element approximation cannot even be evaluated as a candidate solution to the differential equation of the problem, (1.96a), since neither the first nor the second derivatives are defined at the vertices of the mesh. Yet, we did obtain a good approximation of the exact solution in this way. This observation highlights why finite element methods are constructed from a variational equation of the problem: (we will see that) it is simpler to construct functions if milder smoothness requirements are imposed.

We conclude by showing the stiffness matrix of this problem for a constant h and any m, in the case in which $\mathcal{V}_h = \mathcal{W}_h$ (no EBC). This is a matrix also found in finite differences in the same problem, and hence it is a commonly found matrix in elementary numerical analysis textbooks. According to step 4, the matrix is

$$K = \frac{1}{h}\begin{bmatrix} 1 & -1 & 0 & 0 & \dots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix}.$$

About Consistency. In creating the variational method for this example we selected a space $\mathcal{V}_h$ that contains functions that are not smooth (the first derivative is generally discontinuous). Therefore, $\mathcal{V}_h$ is *not* a subset of the test space $\mathcal{V}$ in the variational equation we started from, (1.97). Since $\mathcal{V}_h \not\subseteq \mathcal{V}$, we cannot immediately tell that the method is consistent. Instead, we have to check its consistency, namely, we need to check if $F(u, v_h) = 0$ for all $v_h \in \mathcal{V}_h$

This is true in this case, and the method is consistent. To see this, we can use the integration by parts formula for piecewise smooth functions (1.45) in Theorem 1.2. Since any $v_h \in \mathcal{V}_h$ is a continuous function by construction, then the same integration-by-parts formula used for smooth functions holds. We proceed as we do when finding the Euler-Lagrange equations, we integrate by parts the left hand side of (1.97) to eliminate derivatives over the test function v_h:

$$\begin{aligned} F(u, v_h) &= \int_0^1 u'(x) v_h'(x)\, dx - \int_0^1 v_h(x)\, dx = \\ &= \underbrace{u'(1)\, v_h(1)}_{=0,\text{ due to (1.96c)}} - \underbrace{u'(0)\, v_h(0)}_{=0,\, v_h \in \mathcal{V}_h} - \int_0^1 u''(x)\, v_h(x)\, dx - \int_0^1 v_h(x) dx \\ &= -\int_0^1 \underbrace{(u''(x) + 1)}_{=0,\text{ due to (1.96a)}}\, v_h(x)\, dx \\ &= 0. \end{aligned}$$

Therefore, we can state that

$$F(u, v) = 0 \qquad \forall v \in \mathcal{V} + \mathcal{V}_h,$$

where $\mathcal{V} + \mathcal{V}_h = \{w = v + v_h \mid v \in \mathcal{V}, v_h \in \mathcal{V}_h\}$.