Chapter 4: The Finite Element Method

Michael Hanke

Mathematical Models, Analysis and Simulation, Part I

A Model Problem

Read: Strang, p 229–244

$$-u'' = f(x), \quad 0 < x < 1$$
$$u(0) = u(1) = 0 \qquad \textbf{(D)}$$

Applications:

- axial deformation of an elastic bar
- conduction of heat in a bar
- many others

This formulation is the starting point for *finite difference methods*.

Q: Are there alternatives?

Principle of Virtual Work

In equilibrium, the virtual work vanishes for all possible virtual displacements.

Multiply by a *test function* v (virtual displacement) and integrate:

$$\int_0^1 -u''v dx = \int_0^1 fv dx$$

Since $v(0) = v(1) = 0$, integration by parts yields:

$$\int_0^1 fv dx = \int_0^1 u'v' dx - [u'v]_0^1$$
$$= \int_0^1 u'v' dx$$

We obtain the *weak* or *variational* formulation:

Find u with $u(0) = u(1) = 0$ such that

$$\int_0^1 u'v' dx = \int_0^1 fv dx \text{ for all admissible } v \qquad \textbf{(V)}$$

Principle of Minimum Energy

In equilibrium, the energy of a system attains a minimum.

Energy:

$$P(u) = \int_0^1 \left[\frac{1}{2}u'^2 - fu\right] dx$$

The *minimization* formulation:

Find u with $u(0) = u(1) = 0$ such that

$$P(u) \le P(v) \text{ for all admissible } v \qquad \textbf{(M)}$$

Note: **D** is the Euler-Lagrange equation for **M**.

Notes on These Formulations

- **M** $\Rightarrow$ **V**
- Solutions of **V** and **M** need only be once differentiable.
- If u is twice differentiable, **D** and **V** are equivalent.

$$\mathbf{M} \Longrightarrow \mathbf{V} \Longleftarrow \mathbf{D}$$

Hence, the variational formulation is the most general one.

Sobolev Spaces

Q: What are admissible functions?

They must be once differentiable (in a generalized sense) and fulfill the (essential) boundary conditions.

$$V := H_0^1(0,1) := \{v | v(0) = v(1) = 0, \int_0^1 (v'^2 + v^2) dx < \infty\}$$

This is a special case of *Sobolev spaces* $H^p(\Omega)$: Let Ω be a domain in $\mathbb{R}^n$,

$$H^p(\Omega) := \{v | \int_\Omega \left(\left(v^{(p)}\right)^2 + \cdots + v'^2 + v^2 \right) dx < \infty\}$$

These spaces are examples of a complete inner product space.

Notation:

$$\|v\|_p = \left(\int_\Omega \left(\left(v^{(p)}\right)^2 + \cdots + v'^2 + v^2 \right) dx \right)^{1/2}$$

Ritz and Galerkin Methods

- Choose a (convenient) finite dimensional subspace $V_h \subset V$.
- Choose a basis of V_h.
- *Ritz method:* Start from **M**. Determine $u_h \in V_h$ as the minimizer of $P(v_h)$ where v_h is taken from V_h.
- *Galerkin method:* Start from **V**. Determine $u_h \in V_h$ such that **V** is fulfilled for all $v_h \in V_h$.

Theorem. *The Ritz and Galerkin procedures are equivalent.*

Q: How to choose V_h?

Criteria include:

- good approximation quality,
- efficient numerical algorithms,
- stable computations.

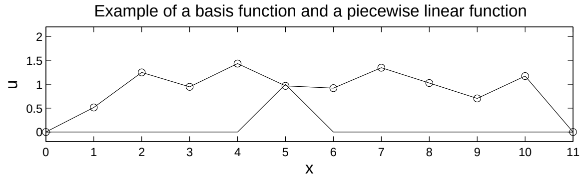
Finite Element Method: Example

Consider the introductory example. Subdivide $[0,1]$ into $N+1$ subintervals (not necessarily equidistant):

$$0 = x_0 < x_1 < \cdots < x_N < x_{N+1} = 1.$$

- V_h: set of all piecewise linear functions with corners at the grid points x_i.

Example of a basis function and a piecewise linear function



- Basis functions: Hat functions

$$\phi_i(x) = \begin{cases} \frac{x - x_{i-1}}{x_i - x_{i-1}}, & x_{i-1} \le x \le x_i, \\ \frac{x_i - x}{x_{i+1} - x_i}, & x_i \le x \le x_{i+1}, \\ 0, & \text{elsewhere} \end{cases}$$

- Ansatz

$$\boxed{u_h(x) = \sum_{j=1}^{N} u_j \phi_j(x)}$$

Example (cont.)

- *Ritz:* Insert

$$u_h(x) = \sum_{j=1}^{N} u_j \phi_j(x), \quad u_h'(x) = \sum_{j=1}^{N} u_j \phi_j'(x)$$

into $P(v)$:

$$P(u_h) = \sum_{j,k=1}^{N} \frac{1}{2} u_j u_k \underbrace{\int_0^1 \phi_j' \phi_k' dx}_{a_{jk}} - \sum_{j=1}^{N} u_j \underbrace{\int_0^1 \phi_j f dx}_{f_j}$$

$$= \frac{1}{2}\mathbf{u}^T \mathbf{A} \mathbf{u} - \mathbf{u}^T \mathbf{f}.$$

- – vector of unknowns: $\mathbf{u} = (u_1, \ldots, u_N)^T$, where $u_i = u_h(x_i)$
  - stiffness matrix $\mathbf{A}$
  - load vector $\mathbf{f}$

Example (cont.)

- properties of the stiffness matrix
  - It is *symmetric*: $a_{jk} = a_{kj}$.
  - It is *positive semi-definite*:

$$\mathbf{u}^T \mathbf{A} \mathbf{u} = \int_0^1 u_h'^2 dx \geq 0.$$

  - It is *positive definite*:

$$\mathbf{u}^T \mathbf{A} \mathbf{u} = 0 \Leftrightarrow u_h'(x) \equiv 0 \qquad \Leftrightarrow u_j = 0 \text{ for all } j$$

  - *Definiteness* depends on the boundary conditions!
- $P(u_h)$ reduces to a quadratic functional. Derivation with respect to the unknowns yields

$$\boxed{\mathbf{A}\mathbf{u} = \mathbf{f}}$$

- How to integrate: numerical quadrature
- Exercise: Do the same analysis for the Galerkin method!

Error Estimation

A *reliable and efficient* method requires an error estimate and a method to adapt the discretization to the problem at hand to produce a *prescribed error* with *minimal resources*.

Some notation:

- Left-hand side of **V**: $a(u,v) := \int_0^1 u'v'dx$.
  Exercise: Show that $a(u,v)$ is a scalar product on V!
- Right-hand side of **V**: $L(v) = \int_0^1 fvdx$
- **M**: $P(v) = \frac{1}{2}a(v,v) - L(v)$
- Exact solution: $a(u,v) = L(v)$ for all $v \in V$
- Galerkin: $a(u_h, v_h) = L(v_h)$ for all $v_h \in V_h$
- The error: $e_h = u - u_h$.

Since $V_h \subset V$:

$$\begin{array}{rl} a(u,v_h) = L(v_h) & \text{for all } v_h \in V_h \\ -a(u_h,v_h) = -L(v_h) & \text{for all } v_h \in V_h \\ \hline a(e_h,v_h) = 0 & \text{for all } v_h \in V_h. \end{array}$$

This is called *Galerkin orthogonality*.

Error Estimation (cont.)

With any interpolant $\Pi_h u$ of u in V_h:

$$a(e_h,e_h) = a(e_h, u-u_h) = a(e_h, u - \Pi_h u + \Pi_h u - u_h) = a(e_h, u - \Pi_h u)$$

since $a(e_h, \Pi_h u - u_h) = 0$.

Cauchy-Schwarz inequality:

$$a(e_h, u - \Pi_h u)^2 \leq a(e_h,e_h) a(u - \Pi_h u, u - \Pi_h u).$$

Finally

$$\boxed{a(e_h,e_h) \leq a(u - \Pi_h u, u - \Pi_h u)}$$

The right-hand term is computable if u is two times continuously differentiable.

Error Estimation (cont.)

Linear interpolation on $I = [x_i, x_{i+1}]$ gives:

$$u'(x) - (\Pi_h u)'(x) = u'(x) - \frac{u(x_{i+1}) - u(x_i)}{x_{i+1} - x_i} = u'(x) - u'(\xi)$$

where $\xi \in (x_i, x_{i+1})$.

$$\begin{aligned}\int_I \big(u' - (\Pi_h u)'\big)^2 dx &= \int_I \big(u'(x) - u'(\xi)\big)^2 dx \\ &= \int_I \Big(\int_\xi^x u''(s) ds\Big)^2 dx \\ &\le \int_I \Big(\int_\xi^x u''^2(s) ds \cdot \int_\xi^x 1^2 ds\Big) dx \\ &\le (x_{i+1} - x_i)^2 \int_I u''^2(s) ds.\end{aligned}$$

Error Estimation (cont.)

Theorem.

$$\|u' - u_h'\|^2 \le \sum_{i=0}^N (x_{i+1} - x_i)^2 \int_{x_i}^{x_{i+1}} u''^2(s) ds$$

A standard theorem says (Friedrich's inequality): There is a constant C such that

$$\|v\| \le C \|\nabla v\|$$

for all $v \in H_0^1(\Omega)$. Hence,

$$\|e_h\| \le C\|e_h'\| \le Ch\|u''\|.$$

Note: In the present case, one can even show:

- Second order convergence: $\|e_h\| = O(h^2)$.
- Pointwise convergence: $\max_{x\in[0,1]} |e_h(x)| \le C_1 h^2$.

Adaptive Algorithms

The error estimate above is an *a-priori* one: It uses only qualitative assumption on the given data.

An *a-posteriori* error estimate uses the actual discrete solution u_h to approximate u''. There are different ways available of doing this. Note that the error is *localized*.

Adaptive algorithm:

1. Construct an initial grid
2. Discretize by FEM
3. Compute the approximation u_h
4. Compute an *a-posteriori* error estimate
5. User selected error criterion met?
   - Yes: We are done.
   - No: Select subintervals with large error and subdivide them.

Adptive Algorithms (cont.)

For the last step, a number of different strategies are available:

- Refine the worst elements.
- Equidistribution of the error.

Note The success of the algorithm depends on the regularity of the solution. For problems with singularities, the refinement process may never terminate.

The Program ADFEM

This program implements an adaptive algorithm for the problem

$$-\frac{d}{dx}\left(d(x)\frac{du}{dx}\right)+c(x)\frac{du}{dx}+a(x)u=f(x)$$

$$x=x_{\min}: u=g_0 \text{ or } d(0)\frac{du}{dx}+k_0u=g_0$$

$$x=x_{\max}: u=g_1 \text{ or } d(1)\frac{du}{dx}+k_1u=g_1$$

Assumptions: $d(x)\geq d_0>0$

Error control:

- in L^2 norm
- in energy norm $\|e_h\|_E:=\sqrt{a(e_h,e_h)}$ (if $c=0,a\geq 0$)
- pointwise error

More explicit:

$$\|v\|_E^2=\int_{x_{\min}}^{x_{\max}}(dv'^2+av^2)dx$$

A 2D Model Problem

Read: Strang, p 293–309

$$-\nabla\cdot(c(x)\nabla u)+r(x)u=f(x), x\in\Omega\subset\mathbb{R}^2$$
$$\partial\Omega=\Gamma_D\cup\Gamma_N \qquad \text{(D)}$$
$$\text{on } \Gamma_D: u=g_1, \text{ on } \Gamma_N: \frac{\partial u}{\partial n}=g_2$$

with $c(x)\geq c_0>0$, $r\geq 0$ and $\Gamma_D\neq\emptyset$.

The variational formulation and the minimization formulation will be constructed by the principle of virtual work and the principle of minimum energy, respectively: Let v be a test function with $v(x)=0$ on Γ_D.

$$\begin{aligned}\int_\Omega fv dx &= \int_\Omega(-\nabla(c(x)\nabla u)+r(x)u)v dx\\ &=\int_\Omega c(x)\nabla u\cdot\nabla v-\int_{\partial\Omega}n\cdot(c(x)\nabla u)v d\Gamma+\int_\Omega r(x)uv dx\\ &=\underbrace{\int_\Omega(c(x)\nabla u\cdot\nabla v+r(x)uv)dx}_{a(u,v)}-\int_{\Gamma_N}c(x)g_2(x)v d\Gamma.\end{aligned}$$

A 2D Problem (cont.)

Remember:

$$H^1(\Omega)=\{v|\int_\Omega(|\nabla v|^2+v^2)dx<\infty\}$$

Define

$$a(u,v)=\int_\Omega(c(x)\nabla u\cdot\nabla v+r(x)uv)dx, \quad u,v\in H^1(\Omega)$$

$$L(v)=\int_\Omega fv dx+\int_{\Gamma_N}c(x)g_2(x)v d\Gamma, v\in H^1(\Omega)$$

Let now $V_g:=\{v\in H^1(\Omega)|v=g \text{ on } \Gamma_D\}$.

Variational formulation: Find $u\in V_{g_1}$ such that

$$a(u,v)=L(v) \text{ for all } v\in V_0. \qquad \text{(V)}$$

Minimization formulation: Find $u\in V_{g_1}$ such that

$$P(u)=\min_{v\in V_{g_1}}P(v) \text{ with } P(v)=\frac{1}{2}a(v,v)-L(v) \qquad \text{(M)}$$

Theorem. ***M** and **V** are equivalent.*

Exercise: Prove this!

The Galerkin Method

We are following the lines of the one-dimensional example:

1. Choose a finite set of *trial functions* or, basis functions $\phi_1(x),\dots,\phi_N(x)$.
2. Admit approximations to u of the form $u_h(x)=u_1\phi_1(x)+\cdots+u_N\phi_N(x)$.
3. determine the N unknown numbers $\mathbf{u}=(u_1,\dots,u_N)^T$ from **V**, using N different test functions $\phi_k(x)$.

$$L(\phi_j)=a(u_h,\phi_j)=a(\sum_{k=1}^N u_k\phi_k,\phi_j)$$

$$\underbrace{L(\phi_j)}_{f_j}=\sum_{k=1}^N\underbrace{a(\phi_k,\phi_j)}_{a_{jk}}u_k$$

The coefficients can be determined from

$$\boxed{\mathbf{Au}=\mathbf{f}}$$

with the *stiffness matrix* $\mathbf{A}=(a_{jk})$ and the *load vector* $\mathbf{f}=(f_1,\dots,f_N)^T$.

Exercise: Show that the Ritz approach leads to the same system.

A Finite Element Example: $P1$ Triangles

Bottlenecks of the Galerkin method:

- The computation of $\mathbf{A}$ is expensive. Every element is a 2D integral.
- Since the number of degrees of freedom N is large, a high-dimensional linear system must be solved.

Wishes:

- Choose basis functions which are flexible enough to approximate the solution accurately with a small number N of trial functions.
- Try to make $\mathbf{A}$ sparse. That means, use an “almost” orthogonal basis.
- The condition number should not be too large.

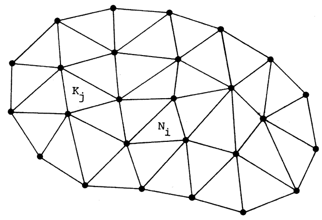
Idea borrowed from 1D: Choose piecewise polynomial trial functions which vanish “almost” everywhere on Ω.

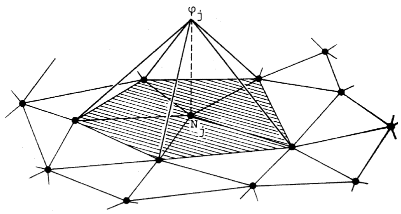
Note: Later we will use other choices, too. (Pseudo spectral method)

$P1$ Triangles (cont.)

- Approximate Ω as the union of non-overlapping triangles T_k whose corners form the set of nodes x_i.
- V_h is defined as the set of continuous functions whose restriction to one triangle is a first degree polynomial.
- Choose basis functions (cf the 1D case!)

$$\phi_i(x_j) = \delta_{ij}$$





Properties

Consequences:

- The stiffness matrix is sparse.
- There is a very efficient algorithm for computing $\mathbf{A}$ and $\mathbf{f}$ (*assembly*).
- The condition number is $\mathrm{cond}(\mathbf{A}) = O(h^2)$.

Theorem. *Under the given assumption on the coefficients (and some regularity assumptions on Ω and the triangulation), the solution to the Galerkin equation exists and is unique. If u is sufficiently smooth,*

$$\|e\|_1 \le Ch.$$

Under additional assumptions on the data,

$$\|e\| \le Ch^2.$$

Stability Estimate

The key properties of a and L for the theorem to hold are

1. $a(v,v) \ge \alpha \|v\|_1^2 \forall v \in V_0$
2. $|a(u,v)| \le C\|u\|_1 \|v\|_1 \forall u,v \in V_0$
3. $|L(v)| \le M\|v\|_1 \forall v \in V_0$

As a consequence of (1), we obtain a stability estimate:

$$\begin{array}{ccc} a(u,u) & = & L(u) \\ \downarrow & & \downarrow \\ \alpha\|u\|^2 \le \alpha\|u\|_1^2 & \le & (f,u) \le \|f\| \cdot \|u\| \end{array}$$

Consequently,

$$\boxed{\|u\| \le \frac{1}{\alpha}\|f\|.}$$