Treatment of essential boundary conditions

- For solving with non-homogeneous Dirichlet boundary conditions $u|_{\partial\Omega} = g$, we need an extension $\widetilde{g} \in H^1(\Omega)$ with $\widetilde{g}|_{\partial\Omega} = g$.
- Let V_N be the finite element space on Ω and let $V_{N,0}$ be the finite element space which contains all basis functions associated with boundary points.
- Compute $\widetilde{g}_N \in V_N$ as the interpolant which satisfies

$$\widetilde{g}_N(\mathbf{x}_i) = \begin{cases} g(\mathbf{x}_i), & \text{if } \mathbf{x}_i \text{ is a boundary node} \\ 0, & \text{otherwise} \end{cases}$$

- The Galerkin formulation reads

$$\text{seek } u_{0,N} \in V_{N,0} \text{ such that } a(u_{0,N} + \widetilde{g}_N, v_N) = \ell(v_N) \text{ for all } v_N \in V_{N,0}$$

which is equivalent to

$$\boxed{\text{seek } u_{0,N} \in V_{N,0} \text{ such that } a(u_{0,N}, v_N) = \ell(v_N) - a(\widetilde{g}_N, v_N) \text{ for all } v_N \in V_{N,0}}$$

Stationary heat conduction with radiation condition

Boundary value problem. Seek $u : \Omega \to \mathbb{R}$ such that

$$-\operatorname{div}\big(\sigma\nabla u\big) = f \text{ in } \Omega, \quad -\sigma\frac{\partial u}{\partial \mathbf{n}} = q|u-u_0|(u-u_0)^3 \text{ on } \partial\Omega$$

- Test the PDE with a smooth function $v \in C^\infty(\overline{\Omega})$ and integrate over the domain:

$$-\int_\Omega \operatorname{div}(\sigma\nabla u)v\,\mathrm{d}\mathbf{x} = \int_\Omega \sigma\nabla u\nabla v\,\mathrm{d}\mathbf{x} - \int_{\partial\Omega} \underbrace{\sigma\frac{\partial u}{\partial \mathbf{n}}}_{=-q|u-u_0|(u-u_0)^3} v\,\mathrm{d}s = \int_\Omega f v\,\mathrm{d}\mathbf{x}.$$

Variational formulation. Seek $u \in H^1(\Omega)$ such that

$$\int_\Omega \sigma\nabla u\nabla v\,\mathrm{d}\mathbf{x} + \int_{\partial\Omega} q|u-u_0|(u-u_0)^3 v\,\mathrm{d}s = \int_\Omega f v\,\mathrm{d}\mathbf{x}$$

for all $v \in H^1(\Omega)$.

- Related non-quadratic minimization problem:

$$J(v) = \frac{1}{2}\int_\Omega \sigma\nabla v\nabla v\,\mathrm{d}\mathbf{x} + \frac{1}{5}\int_{\partial\Omega} q|v-u_0|^5\,\mathrm{d}s - \int_\Omega f v\,\mathrm{d}\mathbf{x} \to \inf_{v\in H^1(\Omega)}$$

Convex minimization

Let V be an Hilbert space and $J : V \to \mathbb{R}$ differentiable:

non-quadratic minimization problem		non-linear variational problem
$u = \operatorname{argmin}_{v \in V} J(v)$	$\longleftrightarrow$	$J'(u)[v] = 0$ for all $v \in V$

Definition. The non-linear functional $J : V \to \mathbb{R}$ is called strictly convex if

$$J\big(tu + (1-t)v\big) < tJ(u) + (1-t)J(v) \text{ for all } u, v \in V \text{ and } t \in (0,1).$$

It is called coercive if

$$\frac{J(v)}{\|v\|_V} \to \infty \text{ as } \|v\|_V \to \infty$$

Theorem. Let V be a Hilbert space and let $J : V \to \mathbb{R}$ be a non-linear functional which is strictly convex and coercive. Then, there exists a unique solution $u \in V$ such that

$$J(u) \le J(v) \text{ for all } v \in V.$$

Discretization

- Gradient of the functional J:

$$J'(u)[v] = \lim_{\varepsilon\to 0}\frac{J(u+\varepsilon v)-J(u)}{\varepsilon}$$
$$= \int_\Omega \sigma\nabla u\nabla v\,\mathrm{d}\mathbf{x} + \int_{\partial\Omega} q|u-u_0|(u-u_0)^3 v\,\mathrm{d}s - \int_\Omega f v\,\mathrm{d}\mathbf{x}$$

- The necessary optimality condition $J'(u)[v] = 0$ for all $v\in H^1(\Omega)$ coincides with the variational formulation.

Galerkin discretization. Let $V_N = \operatorname{span}\{\phi_1,\ldots,\phi_N\}\subset H^1(\Omega)$. Seek $u_N = \sum_{i=1}^N u_i\phi_i \in V_N$ such that

$$\mathbf{A}_N(\mathbf{u}_N) = \mathbf{f}_N$$

where

$$\mathbf{A}_N(\mathbf{u}_N) = \left[\int_\Omega \sigma\nabla u_N\nabla\phi_j\,\mathrm{d}\mathbf{x} + \int_{\partial\Omega} q|u_N-u_0|(u_N-u_0)^3\phi_j\,\mathrm{d}s\right]_{j=1}^N$$
$$\mathbf{f}_N = \left[\int_\Omega f\phi_i\,\mathrm{d}\mathbf{x}\right]_{j=1}^N.$$

Iterative solution

- Use the Newton method

$$x^{(k+1)} = x^{(k)} - \big(f'(x^{(k)})\big)^{-1} f(x^{(k)}), \quad k = 0,1,2,\ldots$$

  for solving the system of non-linear equations.

- Hessian of the functional J:

$$J''(u)[v,w] = \lim_{\varepsilon\to 0} \frac{J'(u+\varepsilon v)[w] - J'(u)[w]}{\varepsilon}$$
$$= \int_\Omega \sigma \nabla v \nabla w \,\mathrm{d}\mathbf{x} + 4\int_{\partial\Omega} q|u-u_0|(u-u_0)^2 vw \,\mathrm{d}s.$$

Newton method. Having computed $u_N^{(k)} = \sum_{i=1}^N u_i^{(k)} \phi_i \in V_N$, one step of the Newton method consists of seeking the next iterate $u_N^{(k+1)} = \sum_{i=1}^N u_i^{(k+1)} \phi_i \in V_N$ according to

$$\text{solve} \quad \mathbf{B}_N(\mathbf{u}_N^{(k)})\mathbf{d} = \mathbf{A}_N(\mathbf{u}_N^{(k)}) - \mathbf{f}_N$$
$$\text{update} \quad \mathbf{u}_N^{(k+1)} = \mathbf{u}_N^{(k)} - \mathbf{d}$$

where

$$\mathbf{B}_N(\mathbf{u}_N^{(k)}) = \left[\int_\Omega \sigma\nabla\phi_i\nabla\phi_j \,\mathrm{d}\mathbf{x} + 4\int_{\partial\Omega} q|u_N^{(k)} - u_0|(u_N^{(k)} - u_0)^2 \phi_i\phi_j \,\mathrm{d}s\right]_{j=1}^N.$$

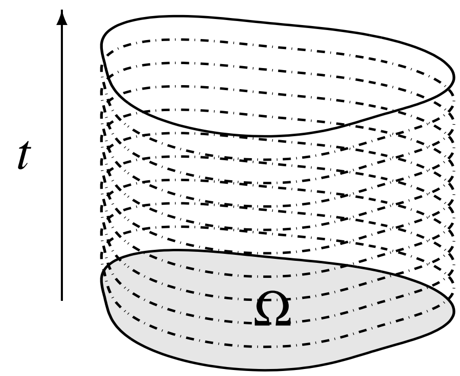
Transient heat conduction

Conservation of energy. For all control volumes $V \subset \Omega$, it holds

$$\frac{\partial}{\partial t}\int_V \rho u \,\mathrm{d}\mathbf{x} + \int_{\partial V} \mathbf{j}\cdot\mathbf{n}\,\mathrm{d}s = \int_V f\,\mathrm{d}\mathbf{x},$$

where f is the time dependent heat source or sink and ρ is the heat capacity (unit $[\rho] =$ J/K), satisfying

$$0 < \underline{\rho} \le \rho(\mathbf{x}) \le \overline{\rho} < \infty \text{ for all } \mathbf{x} \in \Omega.$$



- Gauss' theorem yields the local form of energy conservation: $\rho\frac{\partial u}{\partial t} + \operatorname{div}\mathbf{j} = f \text{ in } \Omega$
- Fourier's law & conservation of energy yields the PDE: $\rho\frac{\partial u}{\partial t} - \operatorname{div}(\sigma\nabla u) = f \text{ in } \Omega$
- If $\sigma \equiv const$ and $\rho = const$, then rescaling gives heat equation: $\frac{\partial u}{\partial t} - \Delta u = f \text{ in } \Omega$
- To be supplemented by boundary conditions on $\partial\Omega$ and initial condition at time $t = 0$:

$$u(0,\cdot) = u_0 \text{ in } \Omega.$$

Variational formulation

Transient heat conduction. Seek $u : [0,T] \times \Omega \to \mathbb{R}$ such that

$$\rho \frac{\partial u}{\partial t} - \operatorname{div}(\sigma \nabla u) = f \text{ in } \Omega$$

with boundary conditions on $\partial\Omega$ and initial condition at time $t = 0$:

$$u(0,\cdot) = u_0 \text{ in } \Omega.$$

- maximum principle: if $f(t,\mathbf{x}) \equiv 0$ for all $t \in [0,T]$ and $\mathbf{x} \in \Omega$, then

$$\inf_{\substack{\tau\in[0,t]\\ \mathbf{z}\in\partial\Omega}} u(\tau,\mathbf{z}) \leq u(t,\mathbf{x}) \leq \sup_{\substack{\tau\in[0,t]\\ \mathbf{z}\in\partial\Omega}} u(\tau,\mathbf{z}) \text{ for all } t \in [0,T] \text{ and } \mathbf{x} \in \Omega$$

- For simplicity, assume homogenous Dirichlet boundary conditions in the following.

Variational formulation in space. For $t \in (0,T)$, seek $u(t) \in H_0^1(\Omega)$ such that

$$\frac{\partial}{\partial t} \int_\Omega \rho u(t) v \, \mathrm{d}\mathbf{x} + \int_\Omega \sigma \nabla u(t) \nabla v \, \mathrm{d}\mathbf{x} = \int_\Omega f(t) v \, \mathrm{d}\mathbf{x} \text{ for all } v \in H_0^1(\Omega).$$

Semi-discretization in space by finite elements

Galerkin discretization. Let $V_N = \operatorname{span}\{\phi_1,\ldots,\phi_N\} \subset H_0^1(\Omega)$. For $t \in (0,T)$, seek $u_N(t) = \sum_{i=1}^N u_i(t)\phi_i \in V_N$ such that

$$\mathbf{M}_N \frac{\partial}{\partial t}\mathbf{u}_N(t) + \mathbf{A}_N \mathbf{u}_N(t) = \mathbf{f}_N(t), \quad \mathbf{u}_N(0) = \mathbf{u}_{N,0},$$

where

$$\mathbf{M}_N = \left[\int_\Omega \rho\phi_i\phi_j \,\mathrm{d}\mathbf{x}\right]_{j,i=1}^N, \quad \mathbf{A}_N = \left[\int_\Omega \sigma\nabla\phi_i\nabla\phi_j \,\mathrm{d}\mathbf{x}\right]_{j,i=1}^N$$

$$\mathbf{f}_N(t) = \left[\int_\Omega f(t)\phi_i \,\mathrm{d}\mathbf{x}\right]_{j=1}^N.$$

- It holds due to the first Poincaré-Friedrichs inequality

$$\lambda_{\min}(\mathbf{M}^{-1/2}\mathbf{A}_N\mathbf{M}^{-1/2}) = \inf_{\mathbf{x}\in\mathbb{R}^N} \frac{\mathbf{x}^\top\mathbf{A}_N\mathbf{x}}{\mathbf{x}^\top\mathbf{M}_N\mathbf{x}} = \inf_{v_N\in V_N} \frac{a(v_N,v_N)}{\|v_N\|_0^2} \geq \underline{c}$$

- But, due to the inverse estimate $\|v_N\|_1 \leq \frac{c}{h}\|v_N\|_0$ (this is sharp!), it also holds

$$\lambda_{\max}(\mathbf{M}^{-1/2}\mathbf{A}_N\mathbf{M}^{-1/2}) = \sup_{\mathbf{x}\in\mathbb{R}^N} \frac{\mathbf{x}^\top\mathbf{A}_N\mathbf{x}}{\mathbf{x}^\top\mathbf{M}_N\mathbf{x}} = \sup_{v_N\in V_N} \frac{a(v_N,v_N)}{\|v_N\|_0^2} \geq \frac{\overline{c}}{h^2}$$

$\rightsquigarrow$ stiff system of linear ordinary differential equations

Time discretization

- explicit Euler method:

$$\mathbf{M}_N \frac{\mathbf{u}_N(t_{i+1}) - \mathbf{u}_N(t_i)}{t_{i+1} - t_i} + \mathbf{A}_N \mathbf{u}_N(t_i) = \mathbf{f}_N(t_i)$$

$\rightsquigarrow$ first order accurate

- implicit Euler method:

$$\mathbf{M}_N \frac{\mathbf{u}_N(t_{i+1}) - \mathbf{u}_N(t_i)}{t_{i+1} - t_i} + \mathbf{A}_N \mathbf{u}_N(t_{i+1}) = \mathbf{f}_N(t_{i+1})$$

$\rightsquigarrow$ first order accurate

- Theta-method $= (1-\theta) \times$ explicit Euler method $+\ \theta \times$ implicit Euler method:

$$\mathbf{M}_N \frac{\mathbf{u}_N(t_{i+1}) - \mathbf{u}_N(t_i)}{t_{i+1} - t_i} + \mathbf{A}_N \{(1-\theta)\mathbf{u}_N(t_i) + \theta \mathbf{u}_N(t_{i+1})\} = (1-\theta)\mathbf{f}_N(t_i) + \theta \mathbf{f}_N(t_{i+1})$$

- Crank-Nicolson method is recovered for $\theta = 1/2$: $\rightsquigarrow$ second-order accurate

$$\left\{\mathbf{M}_N + \frac{1}{2}\mathbf{A}_N\right\}\mathbf{u}_N(t_{i+1}) = \left\{\mathbf{M}_N + \frac{1}{2}\mathbf{A}_N\right\}\mathbf{u}_N(t_i) + \frac{t_{i+1} - t_i}{2}\{\mathbf{f}_N(t_i) + \mathbf{f}_N(t_{i+1})\}$$

- The theta-method is stable for $\theta \geq 1/2$ and even unconditionally stable for $\theta > 1/2$.

Theta scheme in practice

```
% set parameters
T = 1;              % stopping time
N = 100;            % number of time steps
theta = 0.6;        % parameter for the theta scheme

% initialisation
[P,F,B] = mesh;             % load mesh
g = initial(P,F,B);         % compute initial data
f = rhs(P,F,B);             % compute right-hand side
A = stiffness(P,F,B);       % compute system matrix
M = mass(P,F,B);            % compute mass matrix

% compute system matrices
S1 = M+T*theta/N*A;
S2 = M-T*(1-theta)/N*A;

% time stepping
u = zeros(size(P,1),N+1);
u(:,1) = g;
for i = 1:N
  u(:,i+1) = S1\(S2*u(:,i)+k*f);
end;
```

Higher-order implicit Runge-Kutta methods

- Radau(3) method:

$$\left\{\mathbf{M}_N + \frac{5}{12}\Delta t_i \mathbf{A}_N\right\}\mathbf{v}_1 - \frac{1}{12}\Delta_i \mathbf{A}_N \mathbf{v}_2 = \mathbf{f}\left(t_i + \frac{1}{3}\Delta t_i\right) - \mathbf{A}_N \mathbf{u}_N(t_i)$$

$$\frac{3}{4}\Delta t_i \mathbf{A}_N \mathbf{v}_1 + \left\{\mathbf{M}_N + \frac{1}{4}\Delta t_i \mathbf{A}_N\right\}\mathbf{v}_2 = \mathbf{f}(t_i + \Delta t_i) - \mathbf{A}_N \mathbf{u}_N(t_i)$$

$$\mathbf{u}_N(t_i + \Delta t_i) = \mathbf{u}_N(t_i) + \frac{3}{4}\mathbf{v}_1 + \frac{1}{4}\mathbf{v}_2$$

⇝ third order accurate

- SDIRK(2) method (singly diagonally implicit Runge-Kutta method):

$$\left\{\mathbf{M}_N + \lambda \Delta t_i \mathbf{A}_N\right\}\mathbf{v}_1 = \mathbf{f}(t_i + \lambda \Delta t_i) - \mathbf{A}_N \mathbf{u}_N(t_i)$$

$$(1-\lambda)\Delta t_i \mathbf{A}_N \mathbf{v}_1 + \left\{\mathbf{M}_N + \lambda \Delta t_i \mathbf{A}_N\right\}\mathbf{v}_2 = \mathbf{f}(t_i + \Delta t_i) - \mathbf{A}_N \mathbf{u}_N(t_i)$$

$$\mathbf{u}_N(t_i + \Delta t_i) = \mathbf{u}_N(t_i) + (1-\lambda)\mathbf{v}_1 + \lambda \mathbf{v}_2$$

where $\lambda = 1 - \frac{1}{\sqrt{2}}$

⇝ second-order accurate