

Eksamensoppgave 4 - INF5620

Finite elements for a 1D wave equation

Henrik Andersen Sveinsson

December 16, 2013

1 Oppgavetekst

We consider the 1D wave equation problem on $\Omega = [0, L]$:

$$u_{tt} = c^2 u_{xx} + f, \quad u(x, 0) = I(x), \quad u_t(x, 0) = 0, \quad u_x(0) = u_x(L) = 0. \quad (1)$$

1.1 a

Explain how the initial condition can be approximated by the finite element method using the principles of least squares, projection (Galerkin), and interpolation (collocation).

1.2 b

Discretize in time by a centered difference: $u_{tt}(x_i, t_n) \approx [D_t D_t u]_i^n$. Derive a variational formulation of the time-discrete wave equation problem using the Galerkin method. Derive formulas for the element matrix and vector corresponding to the term with u_{xx} in the PDE.

1.3 c

Show how the element matrix associated with the u_{xx} term is computed for P1 elements. Explain the assembly principle and what the resulting global matrix look like when all cells have equal length.

1.4 d

Set up the discrete equations for this wave problem on operator form (assume P1 elements). Briefly explain the idea of an analysis of the scheme based on exact solution of the discrete equations. State the main results. Compare the main results with those of the finite difference method (Problem 3: Analysis of wave problems).

2 Hvordan legge opp oppgaven?

Jeg kommer til å bruke mest tid på oppgave b og c, siden disse på sett og vis illustrerer hele elementmetoden, og jeg har fått mye innsikt av å jobbe gruyndig med disse.

3 Tilnærming av initialbetingelsen

3.1 Interpolasjon (Kolokasjon)

Ideen med interpolasjon er at den tilnærmede funksjonen skal ha samme verdi som den opprinnelige funksjonen på et endelig sett med punkter $x_i \in \mathcal{I}_s$:

$$u(x_i) = \sum_{j \in \mathcal{I}_s} c_j \psi_j(x_i) = f(x_i), i \in \mathcal{I}_s \quad (2)$$

Vi kan like godt skrive at $\psi_j(x_i) = A_{i,j}$ og $f(x_i) = b_i$, slik at vi får et likningssystem:

$$\sum_{j \in \mathcal{I}_s} A_{i,j} c_j = b_i, i \in \mathcal{I}_s \quad (3)$$

Dermed har vi en matriselikning $\mathbf{A}\mathbf{c} = \mathbf{b}$ som løses for å finne de ukjente koeffisientene c_j . Det finnes sett med funksjoner som er slik at $\psi_i(x_j) = \delta_{i,j}$, slik at matrisen A blir diagonal.

3.2 Galerkin (Projeksjon)

Utgangspunktet er at vi vil tilnærme en funksjon f som:

$$f(x) \approx u = \sum_{j \in \mathcal{I}_s} c_j \psi_j \quad (4)$$

Der $u \in V$ er den tilnærmede funksjonen. Prinsippet som skal brukes, er at forskjellen mellom f og u skal være ortogonal på V som inneholder funksjonene vi ekspanderer f i. Dette viser seg å være ekvivalent med minste kvadraters metode.

Vi kan sette opp:

$$e = f - \sum_{j \in \mathcal{I}_s} c_j \psi_j \quad (5)$$

At e er ortogonal på V er det samme som at $(e, v) = 0 \forall v \in V$. For vårt spesifikke valg av basisfunksjoner ψ_j , får vi:

$$(e, \psi_i) = 0 \forall i \in \mathcal{I}_s \quad (6)$$

Vi setter så inn at $e = f - \sum_{j \in \mathcal{I}_s} c_j \psi_j$. Det gir at

$$(f - \sum_{j \in \mathcal{I}_s} c_j \psi_j, \psi_i), i \in \mathcal{I}_s \quad (7)$$

Vi ser at dette blir

$$\sum_{j \in \mathcal{I}_s} (\psi_j, \psi_i) c_j = (f, \psi_i), i \in \mathcal{I}_s \quad (8)$$

Dette er et lineært likningssett $A\mathbf{c} = \mathbf{b}$, som kan løses for \mathbf{c} .

$$A_{i,j} = (\psi_i, \psi_j) \quad (9)$$

$$b_i = (f, \psi_i) \quad (10)$$

For initialbetingelsen spesielt, så setter vi $b_i = I(x_i), i \in \mathcal{I}_s$.

3.3 Least Squares

For minste kvadrater ønsker vi å minimere kvadratet av residualen integrert over domenet:

$$(R, R)_{min} \quad (11)$$

Når vi skriver løsningen vår som en lineærkombinasjon av funksjoner i et funksjonsrom er dette ekvivalent med

$$\left(R, \frac{\partial R}{\partial c_i}\right) \forall i \in \mathcal{I}_s \quad (12)$$

4 Tidsdiskretisering og variasjonell formulering

Vi diskretiserer $u_{tt} = c^2 u_{xx} + f$ med sentrert differanse i tid:

$$\frac{u^{n+1} - 2u^n + u^{n-1}}{\Delta t^2} = c^2 u_{xx} + f \quad (13)$$

$$\mathcal{L}(u) = -u^{n+1} + 2u^n - u^{n-1} + c^2 \Delta t^2 (u_{xx} + f) = 0 \quad (14)$$

Vi har altså en diskret likning i tid, og en kontinuerlig i rom. Vi ønsker nå å tilnærme den romlige løsningen med elementfunksjoner. Da kommer ikke løsningen til å tilfredsstille likningen over, men vi kan lage løsningen slik at den minimerer feilen, nemlig ved å kreve at feilen står ortogonalt på rommet vi utspenner løsningen i.

Vi tenker nå at $u = \sum_{j \in \mathcal{I}_s} c_j \psi_j$, altså at u er utspent i et rom $V = \text{span}\{\psi_1, \dots, \psi_N\}$

Vi setter altså at:

$$R = \mathcal{L}(u) \quad (15)$$

og krever at

$$(R, v) = 0, \forall v \in V \quad (16)$$

som er variasjonsformuleringen.

Vi skriver nå om likningen til ny tidsnotasjon: $u \equiv u^{n+1}$, $u_1 \equiv u^n$, $u_2 \equiv u^{n-1}$.

Da ser variasjonsformuleringen vår ut som dette:

$$(u, v) = 2(u_1, v) - (u_2, v) + c^2 \Delta t^2 [(u_{1,xx}, v) + (f, v)] \quad (17)$$

Så er vi ikke så glade i å derivere løsningene våre to ganger, så vi delvis integrerer den dobbelderiverte, noe som gir oss en naturlig måte å takle grensebetingelser på.

Delvis integrasjonen gir

$$\int_0^L u_{xx} v \, dx = [u_x v]_0^L - \int_0^L u_x v_x \, dx \quad (18)$$

Vi ser at $[u_x v]_0^L = 0$ siden grensebetingelsene er $u_x(0) = u_x(L) = 0$. Dermed blir varisjonsformuleringen:

$$(u, v) = 2(u_1, v) - (u_2, v) + c^2 \Delta t^2 [(u_{1,x}, v_x) + (f, v)] \quad (19)$$

Vi setter så inn den faktiske ekspansjonen av u :

$$\sum_{j \in \mathcal{I}_s} (\psi_j, \psi_i) c_j = 2 \sum_{j \in \mathcal{I}_s} (\psi_j, \psi_i) c_{1,j} - \sum_{j \in \mathcal{I}_s} (\psi_j, \psi_i) c_{2,j} + c^2 \Delta t^2 \left[\sum_{j \in \mathcal{I}_s} (\psi'_j, \psi'_i) c_{1,j} + (f, \psi_i) \right] \quad (20)$$

Vi ser så nærmere på leddet som kommer fra u_{xx} , nemlig $\sum_{j \in \mathcal{I}_s} (\psi'_j, \psi'_i) c_{1,j}$. Dette leddet er i prinsippet en matrise og en vektor, slik at

$$A_{i,j} = \int_0^L \psi'_j(x) \psi'_i(x) \, dx \quad (21)$$

Dette er formelen for elementene i den globale matrisen. Så går vi over til enkelt-elementer av typen $P1$. For ordens skyld skifter vi til elementbasisfunksjoner og integrerer over et referanseelement

$$\tilde{A}_{i,j} = \int_0^L \tilde{\varphi}'_j(X) \frac{dX}{dx} \tilde{\varphi}'_i(X) \frac{dX}{dx} \frac{dx}{dX} dX \quad (22)$$

Her har vi tatt med jacobideterminanten $\frac{dx}{dX}$ og kjerneregelen på de deriverte elementfunksjonene.

$P1$ -funksjoene på hver element er

$$\tilde{\phi}_0(X) = \frac{1}{2}(1 - X), \quad \tilde{\phi}_1(X) = \frac{1}{2}(1 + X), \quad (23)$$

For $P1$ -elementer $[-1, 1]$ er $\frac{dx}{dX} = h$.

Det betyr at en elementmatrise blir:

$$\tilde{A}^{(e)} = \frac{1}{h} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad (24)$$

5 Assemblering

Elementene vi jobber med har overlappende noder. Dersom vi antar at vi har regulær nummerering av nodene setter vi nå sammen den globale matrisen ved å la hjørnene på elementmatrisene overlappe langs hoveddiagonalen.

Den resulterende matrisen er $[-1 \ 2 \ -1]$ repeterende, med noen endringer ved randbetingelsene. Vi vil typisk bruke en randfunksjon for å inkorporere randbetingelser, og denne innføres ved å endre likningssystemet etterpå. Vi kan gjøre endringen symmetrisk ved å gjøre noen radoperasjoner på matrisen etterpå.

6 Sammenlikning med finite difference

Hovedtanken når vi sammenlikner med finite difference er å plukke ut den i'te raden i matriselikningen vi løser og skrive det om med differanseoperatorer. Vi forholder oss fortsatt til likningen

$$\sum_{j \in \mathcal{I}_s} (\psi_j, \psi_i) c_j = 2 \sum_{j \in \mathcal{I}_s} (\psi_j, \psi_i) c_{1,j} - \sum_{j \in \mathcal{I}_s} (\psi_j, \psi_i) c_{2,j} + c^2 \Delta t^2 \left[\sum_{j \in \mathcal{I}_s} (\psi'_j, \psi'_i) c_{1,j} + (f, \psi_i) \right] \quad (25)$$

For å gjøre ting mer håndterbart skriver vi den på matriseform. Vi definerer oss følgende:

- $K = \{\sum (\psi'_i, \psi'_j)\}$
- $M = \{\sum (\psi_i, \psi_j)\}$
- $\mathbf{f} = \{(f, \psi_i)\}$

$$M\mathbf{c} = 2M\mathbf{c}_1 - M\mathbf{c}_2 + c^2 \Delta t^2 (K\mathbf{c}_1 + \mathbf{f}) \quad (26)$$

Vi ønsker nå å skrive dette tilbake til et format som er sammenliknbart med endelige differanser, for å se hvilke modifikasjoner vi har gjort i forhold til den metoden.

Vi vet at K -matrisen anvendt på forrige sett med P1-koeffisienter tilsvarer en standard sentrert differanseoperator.

$$Kc_i^n = [D_x D_x u]_i^n \quad (27)$$

Så vet vi tilsvarende om M -matrisen

$$Mc_i^n = [u - \frac{1}{6} h^2 D_x D_x u]_i^n \quad (28)$$

Dette lar oss skrive hele den diskrete likningen som

$$\left[u - \frac{1}{6} h^2 D_x D_x u \right]_i^{n+1} = 2 \left[u - \frac{1}{6} h^2 D_x D_x u \right]_i^n - \left[u - \frac{1}{6} h^2 D_x D_x u \right]_i^{n-1} + c^2 \Delta t^2 [D_x D_x u]_i^n \quad (29)$$

Vi ser bort ifra driveleddet f akkurat nå, for enkelhets skyld. Dersom vi nå deler på Δt^2 , og flytter de tre like parentesene til venstre side, ser vi at vi har en sentrert differanse i tid av hele parentesen:

$$D_t D_t \left[u - \frac{1}{6} h^2 D_x D_x u \right]_i^n = c^2 [D_x D_x u]_i^n \quad (30)$$

6.1 Stabilitet og fasefeil

Vi vil se på løsningene av de diskrete likningene. Da kommer vi til å få ledd som setter krav på parametrene vi setter inn. Vi vil også se på forskjellen mellom bølgehastighet i den kontinuerlige og numeriske løsningen, for å analysere rippler i løsningen. Det viser seg at for at dette skjemaet skal være stabilt må

$$C \leq \frac{1}{\sqrt{3}} \quad (31)$$

Dette kan sammenliknes med kriteriet i den vanlige finite difference approsjasjonen der $C \leq 1$

Vi mister den fantastiske egenskapen at vi kan sette Couranttallet til 1 og få en eksakt løsning.

Vi kan også se på $\frac{\tilde{c}}{c}$ for finite element og finite difference, og se at i finite element har en annerledes dispersjonsrelasjon enn finite difference.