

TMA4212

Numerical Solution of Differential Equations by Difference Methods

Course Notes

Author:

Trym Sæther
Physics and Mathematics

Semester:

Spring 2024
6th semester

Norwegian University of Science and Technology

Department of Mathematical Sciences

Contents

Chapter 1

Matematiske Forutsetninger

1.1 Lineær algebra

1.1.1 Jordan form

Jordan-matrise-formen er en kanonisk form som enhver kompleks matrise kan bringes til ved en similaritetstransformasjon.

Denne formen er spesielt nyttig når matrisen ikke er diagonaliserbar, og gir oss en elegant måte å forstå matrisens struktur og oppførsel.

Definition 1. Jordan-form

La $A \in \mathbb{C}^{n \times n}$ være en matrise. Det finnes en invertibel matrise P slik at $P^{-1}AP = J$ hvor J er på Jordan-form, dvs.

$$J = \begin{pmatrix} J_1 & & \\ & J_2 & \\ & & \ddots \\ & & & J_k \end{pmatrix}$$

hvor hver J_i er en Jordan-blokk med formen

$$J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}$$

der λ_i er en egenverdi av A .

Remark 1. Jordan-blokker

Jordan-blokkene representerer hvordan A virker på generaliserte egenvektorer.

En Jordan-blokk av størrelse m med egenverdi λ tilsvarer en kjede av m generaliserte egenvektorer v_1, v_2, \dots, v_m som oppfyller

$$(A - \lambda I)v_1 = 0, \quad (A - \lambda I)v_2 = v_1, \quad \dots, \quad (A - \lambda I)v_m = v_{m-1}$$

Dette viser hvordan matrisen "kobler sammen" vektorer i sin virkning.

Jordan-form

| | | |
|---|---|---|
| 3 | 1 | |
| 0 | 3 | |
| | | 2 |

Figure 1.1: Visualisering av Jordan-form med to Jordan-blokker. De blå linjene indikerer grensene mellom de forskjellige Jordan-blokkene.

Example 1. Jordan-form av en 3×3 matrise

La

$$A = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Denne matrisen har egenverdi $\lambda_1 = 4$ med algebraisk og geometrisk multiplisitet 1, egenverdi $\lambda_2 = 2$ med algebraisk multiplisitet 2 og geometrisk multiplisitet 1. Vi kan beregne transformasjonsmatrisen P og få Jordanformen:

$$J = P^{-1}AP = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

med en 1×1 Jordan-blokk for $\lambda_1 = 4$ og en 2×2 Jordan-blokk for $\lambda_2 = 2$.

Geometrisk tolkning av Jordan-form

En Jordan-blokk med egenverdi λ representerer en transformasjon med to effekter:

- **Skalering:** Egenvektorer strekkes/komprimeres med faktor λ
- **Skjæring:** Generaliserte egenvektorer forskyves i retning av egenvektorer med lavere rang, i tillegg til å skaleres

Denne tolkningen gjør det lettere å forstå hvordan en ikke-diagonaliserbar matrise oppfører seg geometrisk.

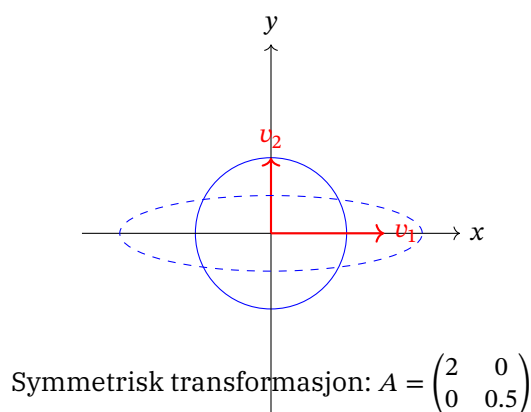


Figure 1.3: Virkningen av en symmetrisk matrise på en sirkel. Transformasjonen strekker og komprimerer kun langs ortogonale akser (egenvektorene), uten rotasjon eller skjæring.

Definition 3. Positiv definit matrise

En symmetrisk matrise $A \in \mathbb{R}^{n \times n}$ er positiv definit hvis

$$x^T A x > 0 \quad \text{for alle } x \in \mathbb{R}^n, x \neq 0$$

A er positiv semidefinit hvis $x^T A x \geq 0$ for alle $x \in \mathbb{R}^n$.

Intuitivt betyr dette at en positiv definit matrise "peker" enhver vektor i en retning som gir en positiv projeksjon på den opprinnelige vektoren.

Theorem 2. Ekvivalente betingelser for positiv definitthet

Følgende betingelser er ekvivalente for en symmetrisk matrise $A \in \mathbb{R}^{n \times n}$:

1. A er positiv definit.
2. Alle egenverdiene til A er positive.
3. Alle hovedminorene til A er positive (Sylvesters kriterium).
4. $A = B^T B$ for en matrise B med full kolonnerang.
5. Det finnes en unik positiv definit matrise C slik at $C^2 = A$ (den positive kvadratroten).

Disse ekvivalente betingelsene gir oss flere praktiske måter å sjekke om en matrise er positiv definit.

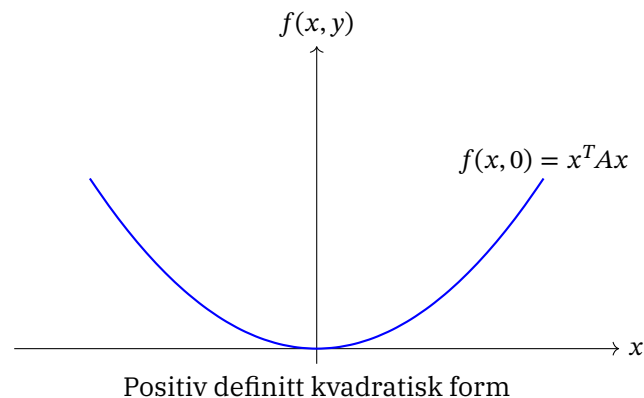


Figure 1.4: Illustrasjon av en positiv definit kvadratisk form i én dimensjon. Grafen er en parabel som er konkv oppover med globalt minimum i origo, noe som visualiserer at $x^T A x > 0$ for alle $x \neq 0$.

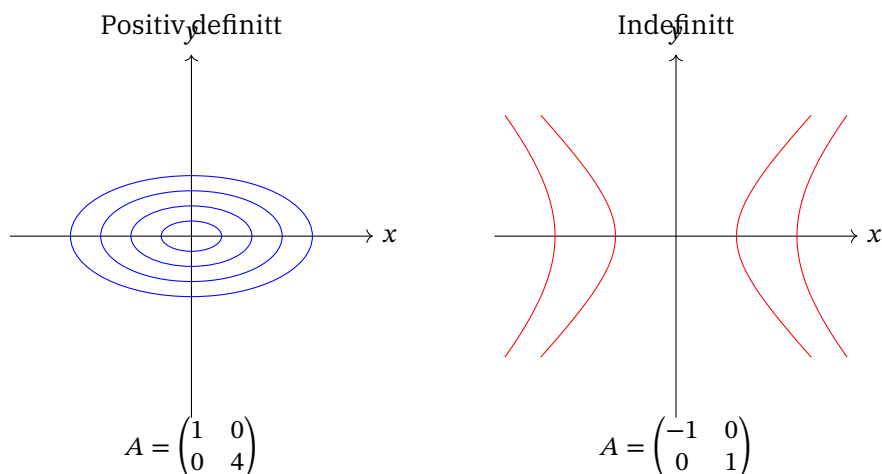


Figure 1.5: Konturer av kvadratiske former i 2D. Venstre: positiv definit matrise gir elliptiske konturer, som representerer nivåkurver av en "bolle"-formet flate. Høyre: indefinit matrise gir hyperbolske konturer, som tilsvarer en "sal"-formet flate.

1.1.4 Gershgorin's teorem

Gershgorin's teorem gir en elegant og praktisk metode for å lokalisere egenverdiene til en matrise ved å konstruere sirkler i det komplekse planet.

Theorem 3. Gershgorin's sirkelteorem

La $A = [a_{ij}] \in \mathbb{C}^{n \times n}$. For hver $i \in \{1, 2, \dots, n\}$, definer Gershgorin-sirkelen

$$D_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\}$$

hvor $r_i = \sum_{j=1, j \neq i}^n |a_{ij}|$ er summen av absoluttverdiene til de ikke-diagonale elementene i rad i .

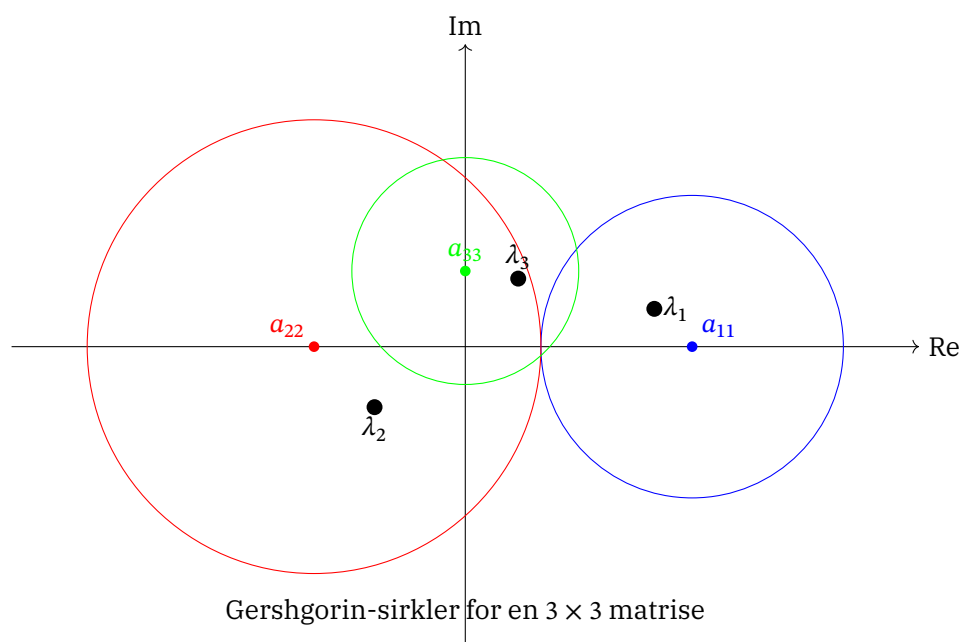


Figure 1.6: Gershgorin-sirkler for en 3×3 matrise. Hver sirkel er sentrert i et diagonalt element a_{ii} og har radius lik summen av absoluttverdiene til de ikke-diagonale elementene i rad i . Egenverdiene $\lambda_1, \lambda_2, \lambda_3$ må ligge innenfor minst én av sirklene.

Enhver egenverdi av A ligger i minst én av Gershgorin-sirklene, dvs. i unionen

$$\bigcup_{i=1}^n D_i$$

Videre, hvis k av sirklene danner et sammenhengende område som er separert fra de andre $n - k$ sirklene, da inneholder dette området nøyaktig k egenverdier av A (telt med multiplisitet).

Intuitivt forteller dette teoremet oss at diagonalelementene i en matrise gir en god indikasjon på hvor egenverdiene ligger, og at de ikke-diagonale elementene bestemmer hvor langt egenverdiene kan avvike fra diagonalelementene.

Example 2. Anvendelse av Gershgorin's teorem

For matrisen

$$A = \begin{pmatrix} 3 & 1 & 0 \\ 0.5 & 5 & 2 \\ 0 & 1 & 4 \end{pmatrix}$$

kan vi for hver rad i beregne radiusene r_i og sentrene a_{ii} :

$$r_1 = |1| + |0| = 1,$$

$$a_{11} = 3$$

$$r_2 = |0.5| + |2| = 2.5,$$

$$a_{22} = 5$$

$$r_3 = |1| + |0| = 1,$$

$$a_{33} = 4$$

slik at vi får de tre Gersgorin-sirklene:

$$D_1 = \{z \in \mathbb{C} : |z - 3| \leq 1\}$$

$$D_2 = \{z \in \mathbb{C} : |z - 5| \leq 2.5\}$$

$$D_3 = \{z \in \mathbb{C} : |z - 4| \leq 1\}$$

Dette betyr at enhver egenverdi av A må ligge i intervallet $[2, 7.5]$ på den reelle aksen.

1.1.5 Vektor- og matrisenormer

Normer gir oss en måte å måle "størrelsen" av vektorer og matriser, og er grunnleggende verktøy i numerisk analyse for å analysere feil og konvergens.

Vektornormer

Definition 4. Vektornorm

En vektornorm på \mathbb{R}^n er en funksjon $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$ som tilfredsstiller:

1. $\|x\| \geq 0$ for alle $x \in \mathbb{R}^n$, og $\|x\| = 0$ hvis og bare hvis $x = 0$ (positivitet)
2. $\|\alpha x\| = |\alpha| \cdot \|x\|$ for alle $x \in \mathbb{R}^n$ og $\alpha \in \mathbb{R}$ (homogenitet)
3. $\|x + y\| \leq \|x\| + \|y\|$ for alle $x, y \in \mathbb{R}^n$ (trekantulikheten)

De mest vanlige vektornormene er p -normene:

$$\|x\|_1 = \sum_{i=1}^n |x_i| \quad (1\text{-norm eller Manhattan-norm})$$

$$\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2} \quad (2\text{-norm eller euklidsk norm})$$

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i| \quad (\infty\text{-norm eller maksimum-norm})$$

Hver norm gir en forskjellig måte å måle avstand på. 1-normen svarer til "taxikjører-avstand" (langs gatenett), 2-normen til "fugleflukt-avstand", og ∞ -normen til "maksimal koordinatforskjell".

Matrisenormer

Definition 5. Matrisenorm

En matrisenorm på $\mathbb{R}^{m \times n}$ er en funksjon $\|\cdot\| : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ som tilfredsstiller:

1. $\|A\| \geq 0$ for alle $A \in \mathbb{R}^{m \times n}$, og $\|A\| = 0$ hvis og bare hvis $A = 0$ (positivitet)
2. $\|\alpha A\| = |\alpha| \cdot \|A\|$ for alle $A \in \mathbb{R}^{m \times n}$ og $\alpha \in \mathbb{R}$ (homogenitet)
3. $\|A + B\| \leq \|A\| + \|B\|$ for alle $A, B \in \mathbb{R}^{m \times n}$ (trekantulikheten)

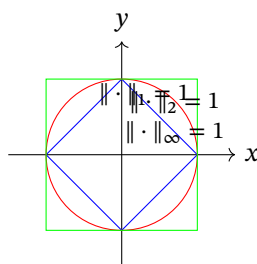


Figure 1.7: Enhetskuler for forskjellige p -normer i \mathbb{R}^2 . Enhver vektor på disse kurvene har norm lik 1. Blå: 1-norm (diamant). Rød: 2-norm (sirkel). Grønn: ∞ -norm (kvadrat).

En praktisk matrisenorm er Frobenius-normen, som måler matrisen som om den var en lang vektor:

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} = \sqrt{\text{tr}(A^T A)}$$

1.1.6 Konsistente og underordnede matrisenormer

Definition 6. Konsistent matrisenorm

En matrisenorm $\|\cdot\|$ på $\mathbb{R}^{n \times n}$ kalles konsistent (eller sub-multiplikativ) hvis

$$\|AB\| \leq \|A\| \cdot \|B\|$$

for alle $A, B \in \mathbb{R}^{n \times n}$.

Denne egenskapen er viktig fordi den lar oss "ta ut" normen når vi har produkter av matriser, noe som er grunnleggende for mange konvergensbevis.

Definition 7. Underordnet matrisenorm

Gitt en vektornorm $\|\cdot\|_v$ på \mathbb{R}^n , defineres den underordnede matrisenormen $\|\cdot\|_M$ på $\mathbb{R}^{n \times n}$ som

$$\|A\|_M = \max_{\|x\|_v=1} \|Ax\|_v = \max_{x \neq 0} \frac{\|Ax\|_v}{\|x\|_v}$$

Intuitivt måler den underordnede normen den maksimale strekningen som matrisen A kan påføre enhver enhetsvektor.

De vanligste underordnede matrisenormene er:

$$\begin{aligned} \|A\|_1 &= \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| && \text{(maksimal kolonnesum)} \\ \|A\|_2 &= \sqrt{\lambda_{\max}(A^T A)} && \text{(største singulærverdi)} \\ \|A\|_\infty &= \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| && \text{(maksimal radsum)} \end{aligned}$$

Disse normene er praktiske fordi de er relativt enkle å beregne og har direkte tolkninger.

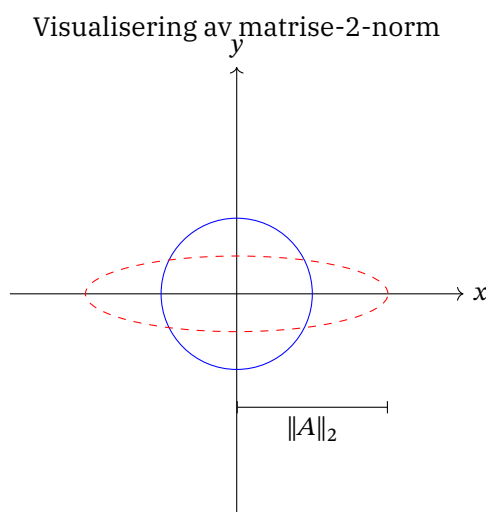


Figure 1.8: Illustrasjon av 2-matrisenormen for en matrise A . Den blå sirkelen er enhetskulen i 2-norm, og den røde ellipsen er resultatet av å anvende A på denne enhetskulen. Normen $\|A\|_2 = 2$ er den maksimale strekningen som A påfører.

Theorem 4. Egenskaper ved underordnede normer

1. Enhver underordnet matrisenorm er konsistent.
2. For enhver matrise $A \in \mathbb{R}^{n \times n}$ og vektor $x \in \mathbb{R}^n$, gjelder

$$\|Ax\|_v \leq \|A\|_M \cdot \|x\|_v$$

1.1.7 Matrisenormer og spektralradius

Definition 8. Spektralradius

Spektralradiusen til en matrise $A \in \mathbb{C}^{n \times n}$ er definert som

$$\rho(A) = \max\{|\lambda| : \lambda \text{ er en egenverdi av } A\}$$

altså den største absoluttverdien av egenverdiene til A .

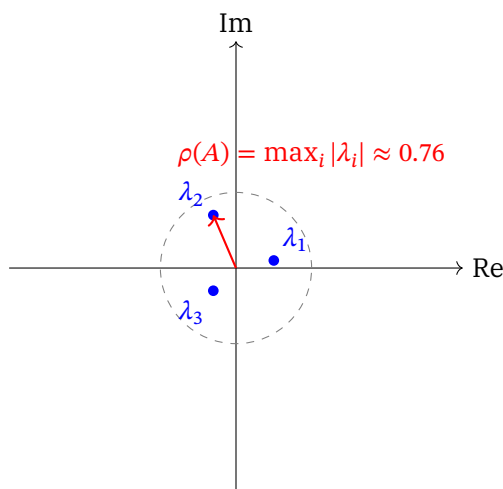
Spektralradiusen er et fundamentalt mål på hvordan en matrise oppfører seg ved gjentatt anvendelse.

Theorem 5. Spektralradius og matrisenormer

La $A \in \mathbb{C}^{n \times n}$ og $\|\cdot\|$ være en konsistent matrisenorm. Da gjelder:

1. $\rho(A) \leq \|A\|$
2. For enhver $\varepsilon > 0$ finnes det en konsistent matrisenorm $\|\cdot\|_\varepsilon$ slik at $\|A\|_\varepsilon \leq \rho(A) + \varepsilon$

Dette betyr at spektralradiusen er den "best mulige nedre grense" for enhver konsistent matrisenorm.



Spektralradius og konvergens: $\rho(A) < 1 \implies \lim_{k \rightarrow \infty} A^k = 0$

Figure 1.9: Spektralradius og konvergens. Når alle egenverdiene ligger innenfor enhetssirkelen ($\rho(A) < 1$), konvergerer sekvensen A^k mot nullmatrisen når $k \rightarrow \infty$. Dette er grunnleggende for stabilitetsanalyse av iterative metoder.

Theorem 6. Konvergens av matrisepotenser

La $A \in \mathbb{C}^{n \times n}$. Da gjelder følgende:

1. $\lim_{k \rightarrow \infty} A^k = 0$ hvis og bare hvis $\rho(A) < 1$
2. For enhver konsistent matrisenorm $\|\cdot\|$, hvis $\|A\| < 1$, så er $\lim_{k \rightarrow \infty} A^k = 0$

Dette resultatet er ekstremt viktig i numerisk analyse, spesielt for å analysere konvergens av iterative metoder.

1.2 Normer og indre produkt

1.2.1 Normer

Normer i funksjonsrom utvider konseptet med vektornormer til uendeligdimensjonale rom. De måler "størrelsen" eller "lengden" av funksjoner.

1.2.2 Indre produkt

Definition 9. Indreprodukt av funksjoner i L^2 (Skalarprodukt)

Hilbertrommet L^2 er gitt ved operasjonen $(L_2, \langle \cdot, \cdot \rangle_2)$ hvor L_2 er settet av alle funksjoner som oppfyller:

$$\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty$$

og vi har for $v, w \in L_2$

$$\langle v, w \rangle_w = \int_{\Omega} v(x)w(x) dx$$

hvor $\Omega = [a, b]$, er et endelig intervall.

Dette indre produktet gir oss en måte å måle ”overlappen” eller ”likheten” mellom to funksjoner. Intuitivt, jo større indreprodukt, desto mer ”like” er funksjonene. Hvis indreproduktet er null, kalles funksjonene ortogonale, noe som betyr at de ikke har noen ”overlapp” i denne forstand.

Et viktig eksempel er Fourier-analysens bruk av indreprodukt for å måle hvor mye en funksjon ”inneholder” av ulike frekvenskomponenter.

1.3 Funksjonsrom

Funksjonsrom er samlinger av funksjoner med bestemte egenskaper. De danner grunnlaget for å analysere differensialligninger i en generalisert ramme.

Funksjonsrommet V Rommet av funksjoner som tilfredsstiller randbetingelsene.

$$V = \{v \in C^2(\Omega) \mid v(a) = \alpha, v(b) = \beta\}$$

Dette rommet representerer alle mulige funksjoner som kan være løsninger på problemet vårt.

Testfunksjoner $v(x)$ Funksjoner i V som brukes til å formulere den svake formen. Testfunksjoner lar oss ”teste” en mulig løsning ved å integrere mot disse funksjonene.

Basisfunksjoner $\phi_i(x)$ Lokale funksjoner som spenner ut løsningsrommet.

- Har kompakt støtte (er null utenfor et lite område)
- Oppfyller $\phi_i(x_j) = \delta_{ij}$ (Kronecker delta)

Basisfunksjoner er byggesteinene vi bruker til å konstruere numeriske løsninger. I endelig element-metoden er disse typisk ”hatt-funksjoner” eller polynomer som er positive i små områder og null ellers.

Lemma 1. Fundamental lemma of calculus of variations La $\Omega \subset \mathbb{R}^n$ være et åpent område med regulær rand og $f : \Omega \rightarrow \mathbb{R}$ være en kontinuerlig funksjon. Hvis

$$\int_{\Omega} \omega(x) \varphi(x) dx = 0 \tag{1.1}$$

holder for alle $\varphi \in C_c^\infty(\Omega)$ (dvs. uendelig differensierbare funksjoner med kompakt støtte i Ω), da er $\omega(x) = 0$ for alle $x \in \Omega$.

Dette lemmaet er grunnleggende for variasjonelle formuleringer av differensialligninger. Det forteller oss at hvis integralet av et produkt er null for alle testfunksjoner, så må funksjonen selv være null nesten overalt.

1.3.1 Energifunksjoner

La $F : V \mapsto \mathbb{R}$ være en funksjon som tar inn en funksjon v og gir ut et reelt tall. Energien til en funksjon v er gitt ved:

$$F(v) = \frac{1}{2} \langle v, v \rangle - \langle f, v \rangle$$

I mange fysiske problemer representerer dette faktisk den fysiske energien i systemet. Minimering av denne energifunksjonen gir oss løsningen på det tilsvarende differensialligningsmessige problemet, noe som er bakgrunnen for variasjonelle formuleringer.

1.4 Sobolev-rom $H^1(\Omega)$

Sobolev-rom er funksjonsrom som er spesielt tilpasset for å jobbe med partiell-deriverte differensialligninger.

$$H^1(\Omega) = \{v \in L^2(\Omega) : \nabla v \in L^2(\Omega)\} \quad (1.2)$$

Dette er rommet av kvadratisk integrerbare funksjoner som også har kvadratisk integrerbare deriverte. I motsetning til C^1 -funksjoner, trenger Sobolev-funksjoner ikke være kontinuerlig differensierbare i vanlig forstand - de kan ha "svake" deriverte som eksisterer i en mer generalisert betydning.

Definition 10. Sobolev-rom $\mathcal{H}_0^1(\Omega)$

La $u : \overset{\subseteq \mathbb{R}^d}{\Omega} \rightarrow \mathbb{R}$ være en funksjon, med $u, \nabla u \in L^2(\Omega)$.

$$\|u\|_{\mathcal{H}^1(\Omega)}^2 = \|u\|_{L^2(\Omega)}^2 + \|\nabla u\|_{L^2(\Omega)}^2 = \int_{\Omega} |u|^2 dx + \int_{\Omega} |\nabla u|^2 dx \quad (1.3)$$

$$\mathcal{H}_0^1(\Omega) = \{u \in \mathcal{H}^1(\Omega) \mid u = 0 \text{ på } \partial\Omega\} \quad (1.4)$$

Med andre ord, en funksjon u er i $\mathcal{H}_0^1(\Omega)$ hvis:

- u og ∇u er kvadratisk integrerbare i $H^1(\Omega)$

$$\|u\|_{L^2(\Omega)}^2 = \int_{\Omega} |u|^2 dx < \infty \quad \text{og} \quad \|\nabla u\|_{L^2(\Omega)}^2 = \int_{\Omega} |\nabla u|^2 dx < \infty \quad (1.5)$$

- u er lik null på randen $\partial\Omega$ av domenet Ω

$$u(x) = 0 \quad \forall x \in \partial\Omega \quad (1.6)$$

Rommet $H_0^1(\Omega)$ er spesielt viktig for differensialligninger med Dirichlet-grensebetingelser, der løsningene er spesifisert på randen.

$$V := H_0^1(\Omega) = \{v \in H^1(\Omega) : v = 0 \text{ på } \partial\Omega\} \quad (1.7)$$

1.5 Svake deriverte og Variasjonell formulering

1.5.1 Svake deriverte

Svake deriverte er en generalisering av vanlige deriverte som tillater oss å jobbe med funksjoner som ikke nødvendigvis er glatte i vanlig forstand.

Idéen bak svake deriverte er å definere den deriverte av en funksjon basert på hvordan den oppfører seg i integraler mot testfunksjoner, snarere enn ved grenseverdier av differenskvotienter.

Dette konseptet er essensielt for å håndtere problemer hvor løsninger kan ha diskontinuiteter eller singulariteter, noe som er vanlig i mange fysiske problemer.

1.5.2 Variasjonell formulering

Variasjonell formulering er en teknikk for å omformulere en differensialligning som et variasjonspproblem, der løsningen minimerer en viss energifunksjonell eller tilfredsstiller en svak form av den originale ligningen.

For en typisk annen ordens differensialligning:

$$-\nabla \cdot (a(x)\nabla u) + c(x)u = f(x) \quad \text{i } \Omega \quad (1.8)$$

med Dirichlet-grensebetingelser $u = 0$ på $\partial\Omega$, er den svake formuleringen:

Finn $u \in H_0^1(\Omega)$ slik at

$$\int_{\Omega} a(x)\nabla u \cdot \nabla v \, dx + \int_{\Omega} c(x)uv \, dx = \int_{\Omega} f(x)v \, dx \quad \forall v \in H_0^1(\Omega) \quad (1.9)$$

Denne formuleringen krever mindre glatthet av løsningen u og er utgangspunktet for finite element-metoden.

1.5.3 Bilinær form

Definition 11. Bilineær form

En bilinær form $a : V \times V \rightarrow \mathbb{R}$ er en funksjon som er lineær i begge argumentene, dvs. for alle $u, v, w \in V$ og $\alpha, \beta \in \mathbb{R}$ gjelder:

$$a(\alpha u + \beta v, w) = \alpha a(u, w) + \beta a(v, w)$$

$$a(u, \alpha v + \beta w) = \alpha a(u, v) + \beta a(u, w).$$

For eksempel i Poisson-ligningen blir den bilinære formen:

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, dx \quad (1.10)$$

Bilinære former er byggeklossene i variasjonelle formuleringer og representerer typisk energien eller arbeidet i det fysiske systemet.

1.5.4 Lax-Milgram teoremet

Lax-Milgram-teoremet er et fundamentalt resultat som garanterer eksistens og entydighet av løsninger til variasjonelle problemer under visse betingelser.

Theorem 7. Lax-Milgram

La V være et Hilbert-rom, og la $a : V \times V \rightarrow \mathbb{R}$ være en bilinær form som er *kontinuerlig* og *koersiv*. Anta at $F : V \rightarrow \mathbb{R}$ er en begrenset lineær funksjonell. Da finnes det en entydig løsning $u \in V$ slik at

$$a(u, v) = F(v) \quad \forall v \in V.$$

For at teoremet skal gjelde, må bilinærformen $a(\cdot, \cdot)$ oppfylle to viktige egenskaper:

Kontinuitet Det finnes en konstant $M > 0$ slik at

$$|a(u, v)| \leq M \|u\|_V \|v\|_V \quad \text{for alle } u, v \in V.$$

Dette betyr at små endringer i inngangsfunksjonene resulterer i små endringer i bilinærformen.

Koersivitet Det finnes en konstant $\alpha > 0$ slik at

$$a(v, v) \geq \alpha \|v\|_V^2 \quad \text{for alle } v \in V.$$

Dette er en slags "positiv definitthet" for bilinærformen og sikrer at problemet er velposert.

Når F er en begrenset lineær funksjonell på V (altså $F \in V'$), garanterer teoremet at:

- Det finnes en entydig løsning $u \in V$ til variasjonsproblemet

$$a(u, v) = F(v) \quad \text{for alle } v \in V.$$

- Løsningen oppfylle den stabile estimeringen

$$\|u\|_V \leq \frac{1}{\alpha} \|F\|_{V'}.$$

Dette betyr at løsningen avhenger kontinuerlig av dataene, en viktig egenskap for numeriske metoder.

Galerkins ortogonalitet

Definition 12. Galerkins Ortogonalitet

$$a(u - u_h, v) = 0 \quad \forall v \in V_h$$

hvor u_h er den approksimerte løsningen, u er den eksakte løsningen, og V_h er det endelige elementrommet.

Galerkins ortogonalitet betyr at feilen $u - u_h$ er "ortogonal" til hele approksimasjonsrommet V_h med hensyn til bilinærformen $a(\cdot, \cdot)$. Dette er en fundamental egenskap ved Galerkins metode, og betyr at approksimasjonen u_h er optimal i en viss forstand.

1.5.5 Céas lemma

Céas lemma gir en øvre grense for feilen i Galerkin-approximasjoner og er et viktig verktøy for å analysere konvergens av endelig element-metoder.

Lemma 2. Céas lemma

$$\|u - u_h\| \leq \frac{M}{\alpha} \inf_{v \in V_h} \|u - v\|$$

hvor M er kontinuitetskonstanten for bilinærformen $a(\cdot, \cdot)$, og α er koersivitetskonstanten.

Dette lemmaet forteller oss at feilen i den numeriske løsningen u_h er begrenset av den best mulige approximasjonen av den eksakte løsningen u i rommet V_h , multiplisert med forholdet $\frac{M}{\alpha}$.

Dette er et kraftig resultat fordi det lar oss fokusere på approximasjonsegenskapene til rommet V_h når vi analyserer konvergens av endelig element-metoder. For eksempel, hvis vi vet at V_h består av stykkevis lineære funksjoner, kan vi bruke klassisk approximasjonsteori til å vise at $\inf_{v \in V_h} \|u - v\| = O(h^2)$ når u er glatt nok, og dermed at $\|u - u_h\| = O(h^2)$.

Chapter 2

Finite Difference Method (FDM)

FDM er en numerisk metode for å løse ODE og PDE ved å diskretisere domenet og tilnærme den deriverte med differanser.

2.1 Teori

2.1.1 Taylors formel

Theorem 8. Taylor-rekke for to variabler

La $u_m^n = u(x_m, t_n) \approx u(x, t)$ være approksimasjonen av den eksakte løsningen u i punktet (x_m, t_n) .

$$u_{m+1}^{n+1} = \sum_{i=0}^p \frac{1}{i!} (h\partial_x + k\partial_t)^i u(x, t) + r_{p+1}(x, t) \quad (\text{taylor})$$

$$r_p = \frac{1}{(p+1)!} (h\partial_x + k\partial_t)^{p+1} u(x, t) \quad (\text{rest})$$

der h, k er steglengdene i (x, t) og p er ordren til Taylor-rekken.

Remark. Uniform Gridstørrelse

Hvis vi har en uniform gridstørrelse \mathbb{G} i x og t , kan vi skrive:

$$\mathbb{G} = \{(x_m, t_n) \mid x_m = x_0 + mh, t_n = t_0 + nk \text{ for } m = 0, 1, \dots, M \text{ og } n = 0, 1, \dots, N\}$$

der h er steglengden i x og k er steglengden i t .

2.2 Motivasjon

Gitt en vilkårlig differensiallikning (ODE), som f.eks. $\frac{d^2 f(x)}{dx^2} = f(x) - x^3$ med randbetingelser: $f(x_0) = \alpha$ og $f(x_N) = \beta$. Denne ODE har allerede en eksakt løsning: $f(x) = c_0 e^x + c_1 e^{-x} + x^3 + 6x$. Men, hva om vi ikke har en eksakt løsning? Da kan vi bruke FDM til å tilnærme løsningen.

2.3 FDM-oppskrift

1. **Diskretisering:** Del intervallet $x \in [\alpha, \beta]$ inn i N like store delintervaller med $x_i = \alpha + ih$ for $i = 0, 1, \dots, N$ og $h = 2/N$
2. **Taylor-utvikling:** Rekkeutvikle $f(x \pm h)$ rundt x .
3. **Differanse metoder:** Finn en lineærkombinasjon av $f(x \pm h)$ som tilnærmer den deriverte.

$$C_{-1}f(x-h) + C_0f(x) + C_1f(x+h) = \frac{d^2f(x)}{dx^2} + O(h) \quad (\text{Central})$$

$$C_0f(x) + C_1f(x+h) + C_2f(x+2h) = \frac{d^2f(x)}{dx^2} + O(h^2) \quad (\text{Forward})$$

$$C_{-2}f(x-2h) + C_{-1}f(x-h) + C_0f(x) = \frac{d^2f(x)}{dx^2} + O(h) \quad (\text{Backward})$$

4. **Diskretiser og løs ODE:** Anvend en av disse metodene, f.eks. ?? på ?? og løs ligningssettet $T_h \vec{f} = \vec{b}$, hvor T_h er en tri-diagonal.

2.4 Differanseoperatorer

Differanseoperatorer er diskrete tilnærminger av deriverte og kan brukes til å konstruere differanseskjemaer for numerisk løsning av ODE og PDE.

Differanseoperatorer er fundamentale byggeklosser for diskretisering av deriverte, og danner grunnlaget for numeriske metoder for differensiallikninger.

2.4.1 Fremover differanseoperator

Definition 13. Fremover differanse

$$\Delta_h u(x) = u(x+h) - u(x) \quad (2.1)$$

der $h > 0$ er steglengden.

Remark

Fremover differanse tilnærmer den første deriverte med nøyaktighet $O(h)$:

$$\frac{\Delta_h u(x)}{h} = u'(x) + \frac{h}{2} u''(x) + O(h^2) \quad (2.2)$$

2.4.2 Bakover differanseoperator

Definition 14. Bakover differanse

Bakover differanseoperatoren er definert som:

$$\nabla_h u(x) = \frac{1}{h} (u(x) - u(x-h)) \quad (\text{bakover})$$

$$r_1(x) = \frac{h}{2} u''(x) + O(h^2) \quad (\text{rest})$$

der $h > 0$ er steglengden og r_p er restleddet.

Remark

Bakover differanse tilnærmer også den første deriverte med nøyaktighet $O(h)$, men bruker tidligere punkter:

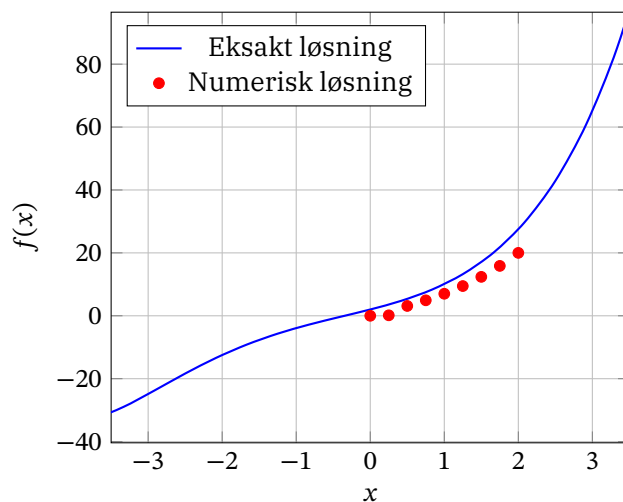
$$\frac{\nabla_h u(x)}{h} = u'(x) - \frac{h}{2} u''(x) + O(h^2) \quad (2.3)$$

Example 3

$f''(x) = f(x) - x^3$ med $[f(0), f(2)] = [0, 20]$ og $N = 5$. Det gir 1D FDM-ligningssettet med sentral differanseoperator:

$$\overbrace{\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \frac{1}{h^2} & -\frac{2}{h^2} - 1 & \frac{1}{h^2} & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \frac{1}{h^2} \\ 0 & \cdots & 0 & 0 & 1 \end{bmatrix}}^{L_h} \overbrace{\begin{bmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_{N-1}) \\ f(x_N) \end{bmatrix}}^{\vec{f}} = \overbrace{\begin{bmatrix} \alpha \\ h^2(x_1^3 - x_1) \\ \vdots \\ h^2(x_{N-1}^3 - x_{N-1}) \\ \beta \end{bmatrix}}^{\vec{b}}$$

$$L_h \vec{f} = \vec{b} \quad (\text{SUS!})$$



2.4.3 Sentral differanseoperator

Definition 15. Sentral differanse

$$\delta_h u(x) = u(x + h/2) - u(x - h/2) \quad (2.4)$$

Remark

Sentral differanse gir en mer nøyaktig tilnærming av den første deriverte:

$$\frac{\delta_h u(x)}{h} = u'(x) + O(h^2) \quad (2.5)$$

Dette gjør den til et foretrukket valg når høyere presisjon er nødvendig.

2.4.4 Andre ordens sentral differanseoperator

Definition 16. Andre ordens sentral differanse

$$\delta_h^2 u(x) = u(x + h) - 2u(x) + u(x - h) \quad (2.6)$$

Proposition 1. Tilnærming til andre deriverte

Andre ordens sentral differanseoperator gir en approksimering av den andre deriverte med nøyaktighet $O(h^2)$:

$$\frac{\delta_h^2 u(x)}{h^2} = u''(x) + O(h^2) \quad (2.7)$$

Remark

Denne operatoren er spesielt nyttig for numerisk løsning av elliptiske og paraboliske partielle differensiallikninger, som varmeledningslikningen og Poissons likning.

2.4.5 Gjennomsnittoperator

Definition 17. Gjennomsnittoperator

$$\mu_h u(x) = \frac{1}{2} (u(x + h/2) + u(x - h/2)) \quad (2.8)$$

Remark

Gjennomsnittoperatoren kan brukes til å konstruere differanseskjemaer med høyere nøyaktighet ved å kombinere fremover og bakover differanser. Den kan også brukes til å lage differanseskjemaer med høyere nøyaktighet ved å kombinere fremover og bakover differanser.

2.4.6 Skiftoperator

Definition 18. Skiftoperator

$$E_h u(x) = u(x + h) \quad (2.9)$$

Remark

Skiftoperatoren kan brukes til å uttrykke alle andre differanseoperatorer, for eksempel:

$$\Delta_h = E_h - I \quad (2.10)$$

$$\nabla_h = I - E_{-h} \quad (2.11)$$

$$\delta_h^2 = E_h - 2I + E_{-h} \quad (2.12)$$

der I er identitetsoperatoren, $Iu(x) = u(x)$.

Theorem 9. Differanseoperatorenes relasjoner

Differanseoperatorene er relatert til hverandre på følgende måter:

$$\delta_h^2 = \Delta_h \nabla_h = \nabla_h \Delta_h \quad (2.13)$$

$$\delta_h = \mu_h \Delta_h = \mu_h \nabla_h E_h \quad (2.14)$$

$$E_h = I + \Delta_h \quad (2.15)$$

der I er identitetsoperatoren, $Iu(x) = u(x)$.

Remark

Fremover- og bakover-differanseoperatorer er asymmetriske og gir første ordens nøyaktighet, mens sentral differanse er symmetrisk og gir andre ordens nøyaktighet ved tilnærming av første deriverte.

Example 4. Anvendelse av differanseoperatorer

For funksjonen $u(x) = x^2$ ved $x = 1$ med $h = 0.1$:

$$\Delta_h u(1) = (1 + 0.1)^2 - 1^2 = 1.21 - 1 = 0.21$$

$$\nabla_h u(1) = 1^2 - (1 - 0.1)^2 = 1 - 0.81 = 0.19$$

$$\delta_h^2 u(1) = (1 + 0.1)^2 - 2(1^2) + (1 - 0.1)^2 = 1.21 - 2 + 0.81 = 0.02$$

Dette stemmer overens med andre deriverte $u''(x) = 2$ siden $\frac{\delta_h^2 u(1)}{h^2} = \frac{0.02}{0.01} = 2$.

Proposition 2. Taylor-utvikling av differanseoperatorer

iifferanseoperatorenes tilnærminger til deriverte kan uttrykkes ved Taylor-rekker:

$$\begin{aligned}\frac{\Delta_h u(x)}{h} &= u'(x) + \frac{h}{2} u''(x) + O(h^2) \\ \frac{\nabla_h u(x)}{h} &= u'(x) - \frac{h}{2} u''(x) + O(h^2) \\ \frac{\delta_h u(x)}{h} &= u'(x) + O(h^2) \\ \frac{\delta_h^2 u(x)}{h^2} &= u''(x) + O(h^2)\end{aligned}$$

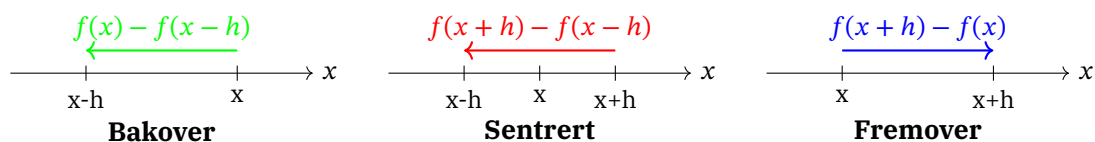
Stensiler

Figure 2.1: Finite Difference Stencils: Backward, Central, and Forward Difference

Som figuren viser, benytter **fremover**-differanser hovedsakelig punkter *til høyre* for x_i , **bakover**-differanser benytter hovedsakelig punkter *til venstre* for x_i , mens **sentrerte**-differanser er symmetriske rundt x_i .

Endelig differanseskjema

Endelige differanseformler gir diskrete tilnærminger av deriverte ved å bruke funksjonsverdier i strategisk utvalgte punkter (ofte kalt *stencil*). Kolonnene merket **-3** til **+3** (eller flere, avhengig av formelens bredde) oppgir *dimensjonsløse koeffisienter* som multipliserer funksjonsverdiene i x_{i+k} .

For å beregne den deriverte numerisk, må du så dividere med h^n , der n er den deriverteordenen:

$$\left. \frac{d^n f}{dx^n} \right|_{x_i} \approx \frac{1}{h^n} \sum_{k=-m}^m (c_k f(x_{i+k}))$$

- m er halvparten av stencilens bredde (f.eks. 3 for en 7-punkts stencil).
- c_k er dimensjonsløse koeffisienter.

Table 2.1: Finite Difference Coefficients for Various Orders of Accuracy

| Order | Type | Coefficients at Position | | | | | | | Error |
|--------|----------|--------------------------|------|------|------|-------|------|----|----------|
| | | -3 | -2 | -1 | 0 | +1 | +2 | +3 | |
| First | Forward | | | | -1 | 1 | | | $O(h)$ |
| | Central | | | -1/2 | 0 | 1/2 | | | $O(h^2)$ |
| | Backward | | | -1 | 1 | | | | $O(h)$ |
| Higher | Forward | | | | -3/2 | 2 | -1/2 | | $O(h^2)$ |
| | Central | -1/12 | 2/3 | 0 | -2/3 | 1/12 | | | $O(h^4)$ |
| | Backward | | 1/2 | -2 | 3/2 | | | | $O(h^2)$ |
| Second | Forward | | | | 1 | -2 | 1 | | $O(h^2)$ |
| | Central | 1/12 | -2/3 | 0 | 2/3 | -1/12 | | | $O(h^4)$ |
| | Backward | | | 1 | -2 | 1 | | | $O(h^2)$ |

Example 5

Hvis vi ønsker å bruke en fremover differanseformel for den første deriverte med $O(h^3)$ nøyaktighet, kan vi bruke koeffisientene fra tabellen: $\{-11, 18, -9, 2\}$.

Dette gir oss:

$$\begin{aligned}\frac{d}{dx}f(x) &\approx \frac{1}{h}(-11f(x) + 18f(x+h) - 9f(x+2h) + 2f(x+3h)) \\ &= \frac{-11f(x) + 18f(x+h) - 9f(x+2h) + 2f(x+3h)}{h}\end{aligned}$$

For en fremover differanseformel av den andre deriverte med $O(h^2)$ nøyaktighet kan vi bruke koeffisientene: $\{2, -5, 4, -1\}$.

Dette gir oss:

$$\begin{aligned}\frac{d^2}{dx^2}f(x) &\approx \frac{1}{h^2}(2f(x) - 5f(x+h) + 4f(x+2h) - f(x+3h)) \\ &= \frac{2f(x) - 5f(x+h) + 4f(x+2h) - f(x+3h)}{h^2}\end{aligned}$$

Tips for bruk:

- Kontroller alltid *tegnkonvensjon og indeksering* $f(x_{i+k}) \leftrightarrow$ koeffisient under kolonne k .
- Etter at koeffisientene er multiplisert med funksjonsverdiene, *deler med h^n* for en n -te deriverte.
- Fremover/bakover brukes ofte nær kantene av et domene for å unngå ute-av-rekkeviddepunkter.
- Sentrerte differanser gir vanligvis høyere nøyaktighet enn fremover- eller bakoverdifferanser, siden de utnytter symmetrien rundt punktet ved å bruke naboverdier likt på begge sider.

2.5 CFL-betingelsen (Courant-Friedrichs-Lewy)

CFL-betingelsen er et nødvendig kriterium for stabilitet i tidsdiskretiseringer av hyperbolske og parabolske [Partiell differensiallikning](#). Betingelsen sikrer at det numeriske domenet av avhengighet inkluderer det fysiske domenet av avhengighet.

Definition 19. Courant-tall

Courant-tallet er definert som:

$$C = \frac{u\Delta t}{\Delta x} \quad (2.16)$$

der u er bølgehastigheten, Δt er tidssteget, og Δx er romsteget.

2.5.1 CFL-betingelser for ulike ligninger og skjemaer

Table 2.2: CFL-betingelser for ulike typer differensialligninger og numeriske skjemaer

| Ligning | Skjema | CFL-betingelse | Kommentar | Stabilitet |
|----------------------|--------------------|----------------------------------|---|------------------|
| Adveksjon | FTCS | $C \leq 1$ | $C = \frac{u\Delta t}{\Delta x}$ | Betinget stabil |
| | Upwind (fremover) | $C \leq 1$ | $C = \frac{\Delta x}{ u \Delta t}$ | Betinget stabil |
| | Lax-Wendroff | $C \leq 1$ | $C = \frac{\Delta x}{ u \Delta t}$ | Betinget stabil |
| | Implisitt | Ingen | - | Ubetinget stabil |
| Diffusjon | Eksplisitt | $r \leq \frac{1}{2}$ | $r = \frac{D\Delta t}{(\Delta x)^2}$ | Betinget stabil |
| | Implisitt (BTCS) | Ingen | - | Ubetinget stabil |
| | Crank-Nicolson | Ingen | - | Ubetinget stabil |
| | DuFort-Frankel | $r < \frac{1}{2\mu}$ | $\mu = \sin^2 \frac{\pi \Delta x}{\lambda_{min}}$ | Betinget stabil |
| Bølge | Eksplisitt | $C \leq 1$ | $C = \frac{c\Delta t}{\Delta x}$ | Betinget stabil |
| | Implisitt | Ingen | - | Ubetinget stabil |
| | Leapfrog | $C \leq 1$ | $C = \frac{c\Delta t}{\Delta x}$ | Betinget stabil |
| Konveksjon-diffusjon | Eksplisitt | $C \leq 1, r \leq \frac{1}{2}$ | Begge må oppfylles | Betinget stabil |
| | FTCS | $r \leq \frac{1}{2}, P_e \leq 2$ | $P_e = \frac{ u \Delta x}{D}$ | Betinget stabil |
| | Upwind | $C \leq 1$ | - | Betinget stabil |
| 2D Adveksjon | Eksplisitt | $C_{2D} \leq 1$ | $C_{2D} = \Delta t \left(\frac{ u_x }{\Delta x} + \frac{ u_y }{\Delta y} \right)$ | Betinget stabil |
| | Alternativt vilkår | $C_{alt} \leq 1$ | $C_{alt} = \Delta t \sqrt{\frac{u_x^2}{\Delta x^2} + \frac{u_y^2}{\Delta y^2}}$ | Betinget stabil |
| 2D Diffusjon | Eksplisitt | $r_{2D} \leq \frac{1}{4}$ | $r_{2D} = D\Delta t \left(\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta y)^2} \right)$ | Betinget stabil |
| | Alternativt vilkår | $r_{alt} \leq \frac{1}{2d}$ | $d = \text{antall dimensjoner}$ | Betinget stabil |

2.5.2 Fysisk tolkning av CFL-betingelsen

- $C < 1$: Informasjon reiser mindre enn én celle per tidssteg (sikker)

- $C = 1$: Informasjon reiser nøyaktig én celle per tidssteg (grensetilfelle)
- $C > 1$: Informasjon reiser mer enn én celle per tidssteg (ustabilt)

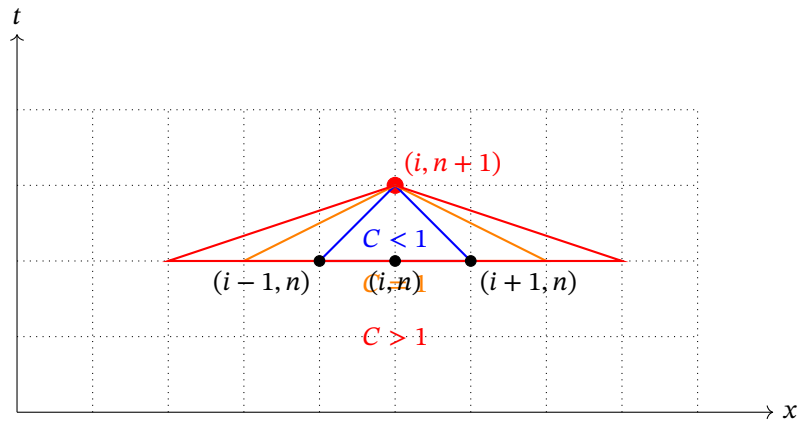


Figure 2.2: Illustrasjon av CFL-betingelsen med numeriske og fysiske domener av avhengighet

2.5.3 Praktiske tilnærminger

1. **Adaptivt tidssteg:** Beregn tillatt Δt for hvert steg basert på gjeldende Δx og fysiske parametre.
2. **Global reduksjon:** Multipliser teoretisk maksimum Δt med en sikkerhetsfaktor (typisk 0.8 – 0.9).
3. **Alternative metoder:** Vurder implisitte metoder for problemer med strenge CFL-betingelser.

Example 6. Beregning av maksimalt tidssteg

For en adveksjonsligning med $u = 100$ m/s og $\Delta x = 0.01$ m:

$$\Delta t \leq \frac{\Delta x}{|u|}$$

$$\Delta t \leq \frac{0.01}{100} = 0.0001 \text{ s}$$

Med en sikkerhetsfaktor på 0.9 vil det praktiske tidssteget bli $\Delta t = 9 \times 10^{-5}$ s.

Chapter 3

Elementmetoden (FEM)

3.1 Introduksjon

Finite Element Metode (FEM) er en numerisk metode for å løse partielle differensiallikninger. Metoden går ut på å dele opp domenet i små, enkle geometriske elementer (som en linje, trekant, firkant osv.). Innenfor hvert element tilnærmes løsningen med enkle funksjoner (typisk polynomer). Til slutt settes elementene sammen for å få en tilnærmet løsning over hele området.

En numerisk metode for å løse differensiallikninger ved å dele opp domenet i mindre elementer. (Finite Element Method (FEM))

3.2 Fremgangsmåte

1. Finn $u : \Omega \rightarrow \mathbb{R}$ slik at $\mathcal{L}u = f$ i Ω og $u|_{\partial\Omega} = g$ på randen $\partial\Omega$.
2. Finn $u \in V := \mathcal{H}_0^1(\Omega)$ slik at $\langle \mathcal{L}u, \mathcal{L}\phi \rangle_\omega = \langle f, \phi \rangle_\omega \quad \forall \phi \in V$.
3. Finn $u_h \in V_h \subset V := \mathcal{H}_0^1(\Omega)$ slik at $\langle \mathcal{L}u_h, \mathcal{L}\phi_i \rangle_\omega = \langle f, \phi_i \rangle_\omega \quad \forall 1 \leq i \leq n$.
4. Løs $A\mathbf{u} = \mathbf{f}$ with $A_{ij} := \langle \mathcal{L}\phi_i, \mathcal{L}\phi_j \rangle$.

1. **Diskretisering 1:** Del **Domenet** Ω inn i enkle geometriske elementer Ω_e og tilnærme løsningen med en lineærkombinasjon av **Basisfunksjoner**:

$$u(x) \approx \sum_{i=1}^N c_i \phi_i(x)$$

2. **Svak formulering:** Multipliser differensiallikningen med en **Testfunksjon** $v(x)$ og integrer over **Domenet** Ω :

$$\int_{\Omega} v(x) \mathcal{L}(u) dx = \int_{\Omega} v(x) f(x) dx$$

3. **Galerkin-prosedyre:** Tilnærme løsningen og **Testfunksjonen** med **Basisfunksjoner**:

$$u(x) \approx \sum_{i=1}^N c_i \phi_i(x) \quad \text{og} \quad v(x) \approx \sum_{j=1}^N d_j \phi_j(x)$$

4. **Diskretisering 2:** Sett inn tilnærmingene i den svake formen og diskretiser:

$$\sum_{j=1}^N d_j \int_{\Omega} \phi_j(x) \mathcal{L} \left(\sum_{i=1}^N c_i \phi_i(x) \right) dx = \sum_{j=1}^N d_j \int_{\Omega} \phi_j(x) f(x) dx$$

5. **Matriseform:** Skriv den diskretiserte formen som et ligningssett:

$$\mathbf{K}\mathbf{c} = \mathbf{F}$$

Og løs ligningssystemet for å finne koeffisientene c_i :

$$\mathbf{c} = \mathbf{K}^{-1}\mathbf{F}$$

1. **Definer problemet:** Finn [Partiell differensiallikning](#), over hvilket domene og rand.

$$u : \Omega \rightarrow \mathbb{R}, \quad \mathcal{L}(u) = f(x) \quad \text{for } x \in \Omega \quad \text{og} \quad u|_{\partial\Omega} = g$$

2. **Diskretiser domenet:** Del opp Ω i enkle geometriske elementer.

3. **Lokal approksimering for hvert element:** Innenfor hvert element (ukjente området/løsning) tilnærmer vi løsningen med en lineærkombinasjon av basisfunksjoner.

$$u(x) \approx \sum_{i=1}^N c_i \phi_i(x)$$

4. **Fra Sterk til Svak formulering:** Formuler [Partiell differensiallikning](#) på [Svak formulering](#).

$$\int_{\Omega} v(x) \mathcal{L}(u(x)) dx = \int_{\Omega} v(x) f(x) dx$$

$$u \in \mathcal{C}^2 \quad \text{s.t.} \quad \mathcal{L}(u) = f \text{ for } x \in \Omega, \quad u|_{\partial\Omega} = g \implies u \in V \quad \text{s.t.} \quad \langle \mathcal{L}(u), \mathcal{L}(\phi) \rangle = \langle f, \phi \rangle \quad \forall \phi \in V$$

5. **Formuler elementeneligningene:** Sett sammen elementene til et system av ligninger, ved å bruke [Svak formulering](#) av [Partiell differensiallikning](#).

$$\mathbf{K}\mathbf{u} = \mathbf{F}$$

6. **Randbetingelser:** Sett opp ligningssystemet med randbetingelser.

7. **Løs ligningssystemet:** Løs ligningssystemet for å finne koeffisientene u_i .

$$\mathbf{u} = \mathbf{K}^{-1}\mathbf{F}$$

3.3 Elementrom og basisfunksjoner

3.3.1 Definisjon av elementrommet $\mathcal{X}_h^p \subset H^1(\Omega)$

Finite Element Space \mathcal{X}_h^p er et rom av stykkvis polynomer av grad p begrenset på elementet $\mathcal{K}_i \in \mathcal{T}_h \subset \Omega$. Hvor \mathcal{T}_h er en partisjon av domenet Ω i elementer K_i .

Example 7. Elementer

Eksempler på elementer er trekanten, firkanten, tetraedre osv.

Definition 20. Finite Element Space

$$\mathcal{X}_h^p = \{v \in \mathcal{C}^0(\Omega) \mid v|_K \in \mathbb{P}_p \text{ for alle } K_i \in \mathcal{T}_h\} \quad (3.1)$$

- $u \in \mathcal{C}^0(\Omega)$ betyr at u er kontinuerlig \mathcal{C}^0 i domenet Ω .
- $v|_K \in \mathbb{P}_p$ betyr at v er et polynom av grad p i elementet K .
- \mathcal{T}_h er en partisjon av domenet Ω i elementer K .

3.3.2 Valg av basisfunksjoner

En sentral del av FEM-metoden er valget av basisfunksjoner for å representere løsningen. Disse funksjonene danner fundamentet for approksimeringen av løsningen i det diskrete rommet.

Egenskaper til basisfunksjoner

Basisfunksjoner er funksjoner som brukes til å representere den approksimerte løsningen u_h i det diskrete rommet \mathcal{X}_h^p . Gode basisfunksjoner for FEM har følgende egenskaper:

- **Lokal støtte:** $\phi_i(x) = 0$ utenfor elementene der node i inngår
- **Interpolasjon:** $\phi_i(x_j) = \delta_{ij}$ (Kronecker delta)
- **Partisjon av enheten:** $\sum_{i=1}^n \phi_i(x) = 1$ for alle $x \in \Omega$
- **Kontinuitet:** Basisfunksjonene er kontinuerlige på tvers av elementgrensene
- **Representasjon:** Gir en god approksimering $u_h(x) = \sum_{i=1}^n u_i \phi_i(x)$

Med disse egenskapene kan vi representere enhver funksjon $v \in X_h^p$ som:

$$v(x) = \sum_{i=0}^M v_i \phi_i(x) \in \mathbb{P}_p$$

Lagrange-basisfunksjoner

Lagrange-basisfunksjoner er den mest brukte typen polynomielle basisfunksjoner i FEM. De er spesielt praktiske fordi koeffisientene u_i direkte representerer funksjonsverdiene ved nodene.

Definition 21. Lagrange-basisfunksjoner

Lagrange-basisfunksjoner av grad p er definert ved interpolasjon over nodepunkter $\{x_i\}_{i=0}^p$. For hvert nodepunkt i , er den tilsvarende Lagrange-basisfunksjonen $\ell_i(x)$ gitt ved:

$$\ell_i(x) = \prod_{j=0, j \neq i}^p \frac{x - x_j}{x_i - x_j} \quad (3.2)$$

med egenskapene:

1. $\ell_i(x_j) = \delta_{ij}$ (Kronecker-delta)
2. $\ell_i(x)$ er et polynom av grad p
3. $\sum_{i=0}^p \ell_i(x) = 1$ for alle x (partisjon av enheten)

Lagrange-basisfunksjoner har flere fordeler i FEM:

- De gir en direkte tolkning av koeffisientene som funksjonsverdier ved nodene
- De forenkler håndteringen av randbetingelser
- De har kompakt støtte, noe som gir sparse matriser i implementeringen

Lagrange-basisfunksjoner i 1D I én dimensjon er Lagrange-basisfunksjonen for node i definert som:

$$\phi_i(x) = \prod_{j=0, j \neq i}^M \frac{x - x_j}{x_i - x_j}$$

Example 8. Lineære basisfunksjoner i 1D ($p = 1$)

For rommet \mathcal{X}_h^1 av stykkvis lineære funksjoner er basisfunksjonene gitt ved:

$$\phi_i(x) = \begin{cases} \frac{x - x_{i-1}}{x_i - x_{i-1}}, & x_{i-1} < x \leq x_i, \\ \frac{x_{i+1} - x}{x_{i+1} - x_i}, & x_i < x < x_{i+1}, \\ 0, & \text{ellers.} \end{cases}$$

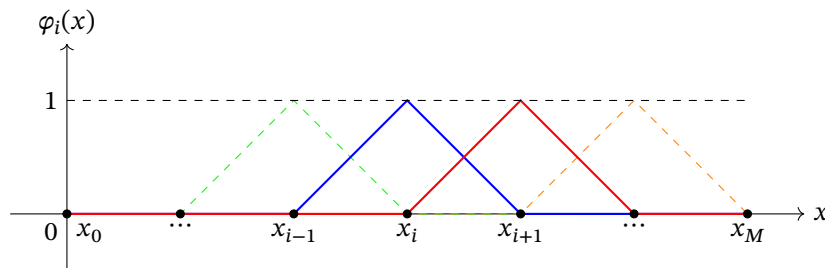


Figure 3.1: Basisfunksjoner $\phi_i(x)$ for $i = 0, 1, \dots, M$ i rommet \mathcal{X}_h^1

På et enkelt element $K_i = [x_i, x_{i+1}]$ er bare to basisfunksjoner ulik null:

$$\phi_i(x) = \frac{x_{i+1} - x}{x_{i+1} - x_i}, \quad \phi_{i+1}(x) = \frac{x - x_i}{x_{i+1} - x_i}, \quad \text{for } x \in [x_i, x_{i+1}].$$

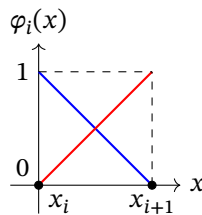


Figure 3.2: Basisfunksjoner $\phi_i(x)$ (blå) og $\phi_{i+1}(x)$ (rød) på elementet $K_i = [x_i, x_{i+1}]$

Nøkkelobservasjoner:

- **Lokal støtte:** Hver $\phi_i(x)$ er ulik null kun på de to elementene K_{i-1} og K_i
- **Partisjon av enheten:** $\sum_{i=0}^M \phi_i(x) = 1$ for alle $x \in [x_0, x_M]$
- **Interpolerende:** For en funksjon $v(x)$, gir $v_h(x) = \sum_{i=0}^M v(x_i)\phi_i(x)$ den stykkevis lineære interpolasjonen av v på nodepunktene

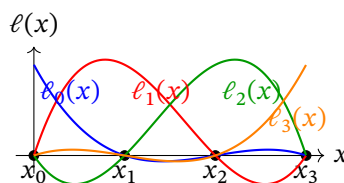


Figure 3.3: Kubiske Lagrange-basisfunksjoner ($p = 3$) på intervallet $[0, 3]$

Dette kan generaliseres til polynomer av høyere grad ved å bruke flere nodepunkter innenfor hvert element, noe som gir bedre nøyaktighet i løsningen.

Lagrange-basisfunksjoner i 2D I to dimensjoner kan Lagrange-basisfunksjoner konstrueres ved bruk av tensorprodukt:

Definition 22. Lagrange-basisfunksjoner i 2D

$$\phi_i(x, y) = \prod_{j=0, j \neq i}^M \frac{x - x_j}{x_i - x_j} \cdot \prod_{k=0, k \neq i}^M \frac{y - y_k}{y_i - y_k}$$

hvor (x_i, y_i) er nodepunktet i elementet, med egenskapene:

- $\phi_i(x_j, y_j) = \delta_{ij}$ (Kronecker delta)
- $\phi_i(x, y)$ er lik null utenfor elementene som inneholder node i
- $\phi_i(x, y)$ er et polynom av grad p i x og y innenfor hvert element

For triangulære elementer med lineære basisfunksjoner ($p = 1$) brukes ofte barysentriske koordinater, som gir spesielt elegante uttrykk for basisfunksjonene.

Remark 2. FEM funksjonsrom

Vi må også kreve noe av rommet vi jobber med.

$$V := \{v : v \in \mathcal{C}[0, 1], v' \text{ er stykkvis kont. og bundet på } [0, 1] \text{ hvor } v(0) = v(1) = 0\}$$

3.4 FEM-betingelser

Betingelser For å bruke [Elementmetoden](#) til å løse et [Partiell differensiallikning](#)-problem, må problemet oppfylle følgende betingelser:

- **Linearitet:** Problemet må være lineært, dvs. ligningene kan uttrykkes som $\mathcal{L}(u) = f$, hvor \mathcal{L} er en lineær operator.
- **Kontinuerlig Differensierbar:** Løsningen $u(x)$ må være kontinuerlig differensierbar i Ω .
- **Geometrisk Enkelhet:** [Domenet](#) Ω bør kunne deles opp i enkle geometriske elementer (f.eks. trekanten, firkanter i 2D, tetraedre i 3D):

$$\Omega = \bigcup_{e=1}^E \Omega_e$$

- **Kvantiserbarhet:** Problemet må være kvantiserbart, dvs. løsningen kan tilnærmes godt ved hjelp av en endelig [Basisfunksjon](#):

$$u_h(x) = \sum_{i=1}^N c_i \phi_i(x)$$

- **Randbetingelser:** Randbetingelsene må være compatible med valg av funksjonsrom:

$$u|_{\partial\Omega} = g \quad \text{eller} \quad \frac{\partial u}{\partial n} \Big|_{\partial\Omega} = h$$

For å anvende FEM på en PDE, må følgende betingelser være oppfylt:

3.4.1 Linearitet

Problemet må kunne uttrykkes i formen $\mathcal{L}(u) = f$, der \mathcal{L} er en lineær differensialoperator.

3.4.2 Regularitet

Løsningen $u(x)$ må ha tilstrekkelig regularitet (kontinuitet og differensierbarhet) innenfor Ω .

$$u \in C^2(\Omega) \quad \text{og} \quad u \in C^1(\partial\Omega)$$

3.4.3 Diskretisering av domenet

Domenet Ω må kunne dekomponeres i ikke-overlappende elementer:

$$\Omega = \bigcup_{e=1}^E \Omega_e, \quad \Omega_i \cap \Omega_j = \emptyset \text{ for } i \neq j$$

3.4.4 Approksimering av funksjonsrommet

Løsningen må kunne representeres tilfredsstillende ved hjelp av basisfunksjoner:

$$u(x) \approx u_h(x) = \sum_{i=1}^N c_i \phi_i(x)$$

3.4.5 Veldefinerte randbetingelser

Problemet må ha klart definerte randbetingelser:

- Dirichlet-betingelser: $u = g_D$ på $\partial\Omega_D$
- Neumann-betingelser: $\nabla u \cdot \mathbf{n} = g_N$ på $\partial\Omega_N$
- Robin-betingelser: $\alpha u + \beta \nabla u \cdot \mathbf{n} = g_R$ på $\partial\Omega_R$

3.5 Sterk formulering

Gitt $f(x)$, finn $u(x)$ slik at

$$\begin{aligned} u''(x) &= f(x) \quad \text{for alle } 0 \leq x \leq 1, \\ u(0) &= 0, \quad u'(1) = 0. \end{aligned}$$

Introduserer testfunksjonene $v(x)$.

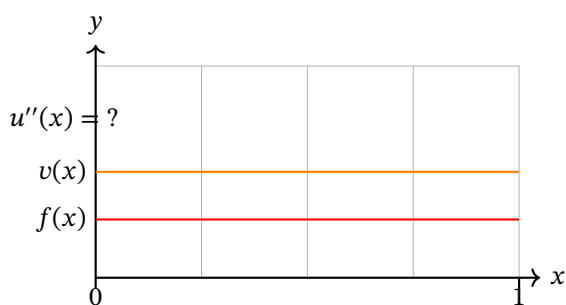
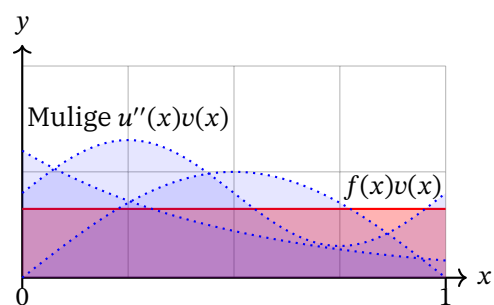
$$u''(x)v(x) = f(x)v(x)$$

Så integrerer vi over intervallet $[0, 1]$.

$$\int_0^1 -u''(x)v(x) dx = \int_0^1 f(x)v(x) dx$$

Alt vi har gjort til nå er å multiplisere med $v(x)$ og integrere over intervallet $[0, 1]$, som er generelt lov så lenge $v(x)$ er kontinuerlig og $u(x)$ er to ganger kontinuerlig deriverbar. Noe den er fra antagelsen om at $u(x)$ er to ganger kontinuerlig deriverbar.

Hvis vi sammenligner integralene på hver side, sier vi egentlig at arealet for $u''(x)v(x)$ er lik arealet for $f(x)v(x)$.

Figure 3.4: Vilkrårlig testfunksjon $v(x)$, og kjent $f(x)$ Figure 3.5: Flere mulige løsninger for $u''(x)v(x)$, som har samme areal som $f(x)v(x)$ Figure 3.6: Flere mulige løsninger for $u''(x)v(x)$, som har samme areal som $f(x)v(x)$

I figur ?? er det flere mulige løsninger for $u''(x)v(x)$, som har samme areal som $f(x)v(x)$.

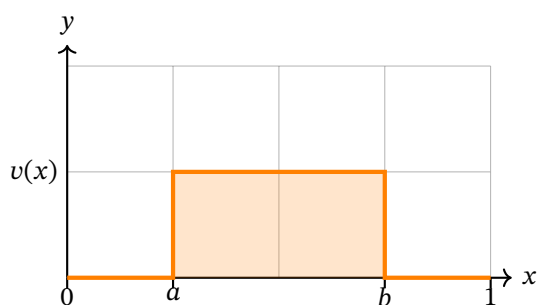
I dette tilfellet er testfunksjonen vår $v(x) = 1$ forteller ikke dette oss noe mer om hvordan $u''(x)$ ser ut. Altså det gir oss ikke noe mer informasjon om $u''(x)$.

Målet vårt nå er å finne/lage en testfunksjon $v(x)$ som faktisk vil kunne gi oss mer informasjon om $u''(x)$.

La oss se på testfunksjonen

$$v(x) = \begin{cases} 0 & \text{for } x < a, \\ 1 & \text{for } a \leq x \leq b, \\ 0 & \text{for } b < x. \end{cases}$$

Denne testfunksjonen er lik 1 bare i intervallet $[a, b]$ og 0 ellers, som vist i figuren under.



Vi vet at $u''(x) = f(x)$ fra den sterke formuleringen. Med vår valgte testfunksjon $v(x)$ får vi:

$$\begin{aligned} \int_0^1 u''(x) \cdot v(x) dx &= \int_0^1 f(x) \cdot v(x) dx \\ \int_0^1 u''(x) \cdot v(x) dx &= \int_a^b f(x) \cdot 1 dx \\ \int_0^1 u''(x) \cdot v(x) dx &= \int_a^b f(x) dx \end{aligned}$$

Merk at integralet på høyre side er begrenset til intervallet $[a, b]$ siden $v(x) = 0$ utenfor dette intervallet. På samme måte, på venstre side, bidrar $u''(x)$ kun til integralet når $x \in [a, b]$.

En nyttig egenskap ved denne tilnærmingen er at hvis vi lar intervallet $[a, b]$ bli veldig lite, slik at $b \rightarrow a$, kan vi intuitivt tenke at:

$$\int_a^b u''(x) dx \approx u''(a) \cdot (b - a)$$

Tilsvarende for høyresiden:

$$\int_a^b f(x) dx \approx f(a) \cdot (b - a)$$

Dette gir oss $u''(a) \approx f(a)$ når $b \rightarrow a$, som stemmer med vår opprinnelige differensialligning.

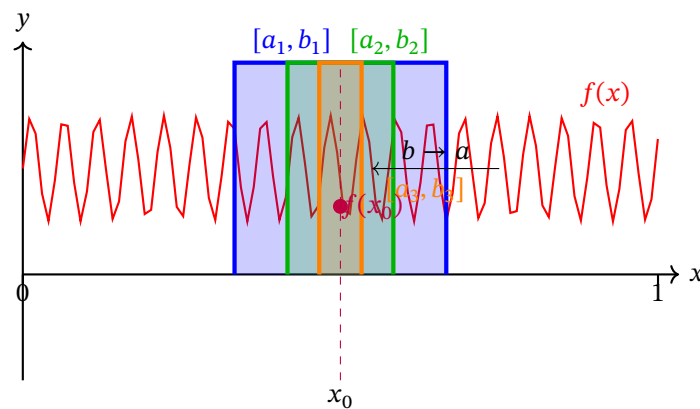


Figure 3.7: Når vi gjør intervallet $[a, b]$ stadig mindre, får vi til slutt punktverdien $u''(x_0) = f(x_0)$

Men den virkelige styrken med svak formulering kommer når vi bruker delvis integrasjon for å redusere derivasjonsordenene.

3.5.1 Delvis integrasjon og svakere krav til regularitet

La oss nå anvende delvis integrasjon på den svake formuleringen. Vi har generelt at:

$$\int_0^1 u''(x)v(x) dx = u'(x)v(x)|_0^1 - \int_0^1 u'(x)v'(x) dx$$

Med randbetingelsene våre $u(0) = 0$ og $u'(1) = 0$, og hvis vi krever at $v(x)$ oppfyller $v(0) = 0$ (siden $u(0) = 0$), får vi:

$$\begin{aligned}
 \int_0^1 u''(x)v(x) dx &= u'(x)v(x)|_0^1 - \int_0^1 u'(x)v'(x) dx \\
 &= u'(1)v(1) - u'(0)v(0) - \int_0^1 u'(x)v'(x) dx \\
 &= 0 \cdot v(1) - u'(0) \cdot 0 - \int_0^1 u'(x)v'(x) dx \\
 &= - \int_0^1 u'(x)v'(x) dx
 \end{aligned}$$

Dermed blir den svake formuleringen:

$$\int_0^1 u'(x)v'(x) dx = \int_0^1 f(x)v(x) dx$$

Dette er en viktig omformulering fordi:

1. Vi trenger nå bare at $u(x)$ er én gang deriverbar, ikke to ganger som i den sterke formuleringen.
2. Randbetingelsen $u'(1) = 0$ er naturlig inkorporert i formuleringen.
3. Vi kan nå bruke stykkevis lineære funksjoner som har veldefinert første deriverte (nesten overalt), men ikke andre deriverte.

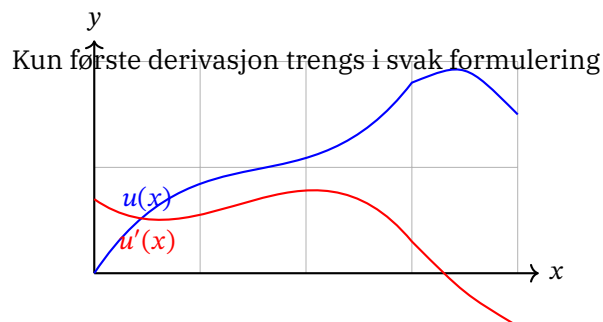


Figure 3.8: I svak formulering trenger vi bare første deriverte av $u(x)$

Sterk formulering

$$u''(x) = f(x) \\ u(0) = 0, u'(1) = 0$$

Krav:

- $u \in C^2[0, 1]$
- $\partial\Omega$ må oppfylles eksakt

Svak formulering

$$\int_0^1 u'(x)v'(x) dx = \int_0^1 f(x)v(x) dx \\ \forall v \in V = \{v \in H^1[0, 1] : v(0) = 0\}$$

Fordeler:

- $u \in H^1[0, 1]$
- Noen $\partial\Omega$ automatisk oppfylt

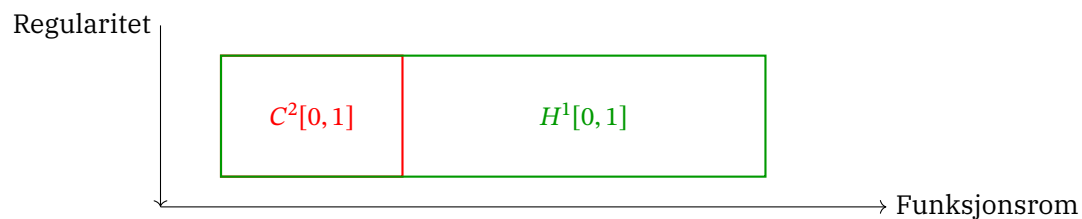


Figure 3.9: Sammenligning av sterk og svak formulering av differensialligninger

3.5.2 Sammenheng med FEM

Den svake formuleringen er grunnlaget for endelig element-metoden (FEM). I FEM antar vi at:

$$u(x) \approx \sum_{i=1}^n c_i \varphi_i(x) \\ v(x) = \varphi_j(x) \quad \text{for } j = 1, 2, \dots, n$$

hvor $\varphi_i(x)$ er basisfunksjoner (typisk stykkevis lineære funksjoner). Ved å sette disse uttrykkene inn i den svake formuleringen, får vi et lineært ligningssystem for koeffisientene c_i :

$$\sum_{i=1}^n c_i \int_0^1 \varphi_i'(x) \varphi_j'(x) dx = \int_0^1 f(x) \varphi_j(x) dx \quad \text{for } j = 1, 2, \dots, n$$

Dette kan skrives på matriseform som:

$$\mathbf{Kc} = \mathbf{F}$$

hvor:

$$K_{ji} = \int_0^1 \varphi_i'(x) \varphi_j'(x) dx \\ F_j = \int_0^1 f(x) \varphi_j(x) dx$$

Denne formuleringen er grunnlaget for numerisk løsning av differensialligninger ved hjelp av endelig element-metoden.

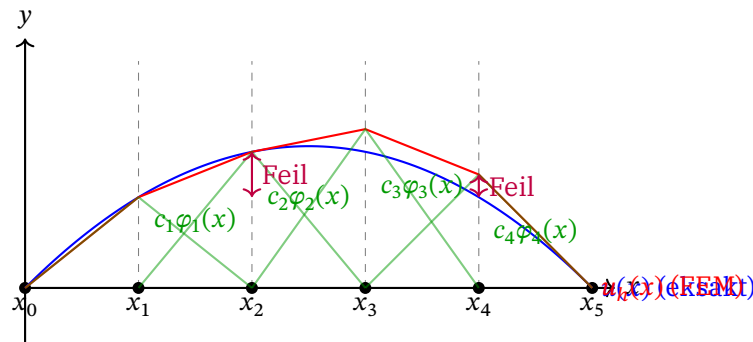


Figure 3.10: FEM-løsning ($u_h(x)$) som en sum av vektete basisfunksjoner, sammenlignet med den eksakte løsningen ($u(x)$)

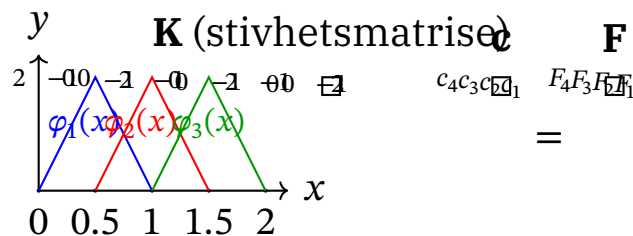


Figure 3.11: FEM diskretisering: basisfunksjoner $\varphi_i(x)$ og tilhørende ligningssystem $\mathbf{K}\mathbf{c} = \mathbf{F}$

Her skriver vi PDE direkte i differensialform. Det betyr at vi forutsetter at løsningen $u(x)$ er glatt nok til at alle deriverte eksisterer. Da har vi:

$$\mathcal{L}(u) = f,$$

samt presise krav på grenseverdier, for eksempel:

$$u|_{\partial\Omega} = g \quad \text{eller} \quad \frac{\partial u}{\partial n}|_{\partial\Omega} = h.$$

Hvis løsningen u ikke er glatt nok, bruker vi en [Svak formulering](#) av problemet.

3.6 Svak form

Her tester vi u med en [Testfunksjon](#) $v(x)$ over hele domenet:

$$\int_{\Omega} v^{(k)}(x) \mathcal{L}(u(x)) dx = \int_{\Omega} v(x) f(x) dx.$$

Denne metoden gjør det mulig å finne løsninger i et bredere funksjonsrom.

[Testfunksjon](#) er en vilkårlig funksjon som tilfredsstiller randbetingelsene, og velges ofte til å være det samme som [Basisfunksjon](#):

$$v(x) = \sum_{j=1}^n v_j \phi_j(x) = \mathbf{v}^T \boldsymbol{\phi}(x)$$

3.6.1 Fra sterk form

Her skriver vi [Partiell differensiallikning](#)-en direkte i differensialform. Det betyr at vi forutsetter at løsningen u er glatt nok til at alle deriverte eksisterer. Da har vi:

$$\mathcal{L}(u) = f,$$

samt presise krav på grenseverdier, for eksempel:

$$u|_{\partial\Omega} = g \quad \text{eller} \quad \left. \frac{\partial u}{\partial n} \right|_{\partial\Omega} = h.$$

3.6.2 Til svak form

For løsninger som ikke er tilstrekkelig glatte, bruker vi en [Svak formulering](#):

1. Multipliser PDE-en med en [Testfunksjon](#) $v(x)$
2. Integrer over domenet Ω :

$$\int_{\Omega} v^{(k)}(x) \mathcal{L}(u(x)) dx = \int_{\Omega} v(x) f(x) dx$$

3.6.3 Fordeler med svak form

- Tillater løsninger i bredere funksjonsrom
- Reduserer glatthetskrav
- Gir mer fleksible randbetingelser

En [Testfunksjon](#) $v(x)$ tilfredsstiller randbetingelsene og uttrykkes i Galerkin-metoden ved:

$$v(x) = \sum_{j=1}^n v_j \phi_j(x) = \mathbf{v}^T \boldsymbol{\phi}(x)$$

3.7 Basisfunksjoner (Formfunksjoner)

Basisfunksjoner er byggsteinene i [Elementmetoden](#)-metoden. De er enkle funksjoner som gjør at vi kan representere en komplisert funksjon ved hjelp av enkle byggesteiner.

En [Basisfunksjon](#) $\phi_i(x)$ er en lokal funksjon som:

- Er null overalt unntatt i nærheten av node i (lokalt definert).
- Har verdien 1 i node i og 0 i alle andre noder
- Til sammen kan bygge opp løsningen vår $u(x)$ som en sum:

$$u(x) = \sum_{i=1}^n u_i \phi_i(x) = \mathbf{u}^T \boldsymbol{\phi}(x)$$

- u_i er koeffisientene som bestemmer hvor mye av hver [Basisfunksjon](#) som skal brukes.

3.7.1 Stykkvis lineære basisfunksjoner

La $u(x)$ være en tilnærming til løsningen av et [Partiell differensiallikning](#), og la $u(x)$ være gitt ved en lineærkombinasjon av [Basisfunksjoner](#):

$$u(x) = \sum_{i=1}^6 u_i N_i(x)$$

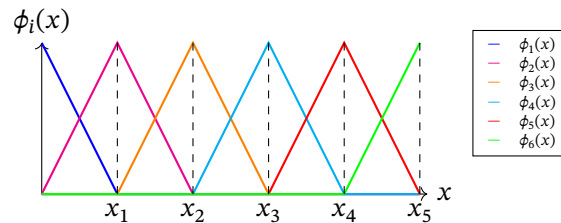


Figure 3.12: Piecewise Linear Basis Functions $\phi_i(x)$

FEM for Poisson-ligningen

$$\begin{cases} -\frac{d^2 u(x)}{dx^2} = f(x), & x \in (0, 1) \\ u(0) = \alpha, \quad u(1) = \beta \end{cases} \quad (\text{randbetingelser}) \quad (3.3)$$

$$\mathbf{F} = - \int_0^1 f(x) \mathbf{N}(x) dx \quad (3.4)$$

La $f(x) = \bar{f} = C$ være konstant.

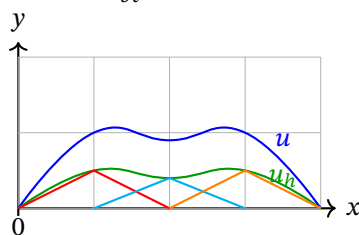
Vi antar at $u(x_m)$ er ukjent for $m = 1, \dots, M$ punkter (noder) i det diskrete [Domenetet](#) Ω_h .

I mellom nodene definerer vi *elementene* \implies *formfunksjoner* $N_i(x)$.

Example 9

- **Eksakt løsning:** u
- **Numerisk løsning:** u_h
- **Feil:** $e = \|u - u_h\|$

La $\|u - v\|_{\mathcal{H}^1}$ for en tilfeldig $v \in \mathcal{X}_h^1$.



3.8 Interpolasjonsoperator

Vi definerer en interpolasjonsoperator $\Pi_h^1 : \mathbb{H}(\Omega)^1 \rightarrow \mathbb{X}_h^1$.

$$\begin{aligned}\Pi_h^1 v(x_i) &= v(x_i) \quad i = 0, 1, \dots, M \\ \Pi_h^1 v(x) &= \sum_{i=0}^M v(x_i) \varphi_i(x) \in \mathbb{X}_h^1 = V_h\end{aligned}$$

Med interpolasjonsfeilen:

$$e(x) = v(x) - \Pi_h^1 v(x)$$

som vi ønsker å finne en øvre grense for $\|e(x)\|_{\mathbb{H}^1}$.

La nå $H^2(\Omega) = \{v \in H^1(\Omega) : v_{xx} \in L^2(\Omega)\}$ være et Hilbert-rom med norm $\|v\|_{H^2}^2 = \int_{\Omega} v_{xx}^2 dx$.

$$\begin{aligned}\|e(x)\|_{\mathbb{H}^1} &= \int_0^1 e^2 dx + \int_0^1 e_x^2 dx \\ &= \sum_{k \in \mathcal{T}_h} \int_k e^2 dx + \sum_{k \in \mathcal{T}_h} \int_k e_x^2 dx\end{aligned}$$

Vi vet at $e(x_k) = e(x_{k+1}) = 0$.

For $x > \xi_k$:

$$e_x(x) = \int_{\xi_k}^x e_{xx} ds = \int_{\xi_k}^x v_{xx} ds$$

Fordi $(\Pi_h^1 v)_{xx} = 0$.

For $x < \xi_k$:

$$\begin{aligned}
 e_x(x) &= - \int_x^{\xi_k} v_{xx} \, ds \\
 |e_x|_K &\leq \int_{K_k} |v_{xx}| \, dx = \int_{K_k} 1 \cdot |v_{xx}| \, dx \\
 &= \langle 1, |v_{xx}| \rangle_{L^2(K_k)} \leq \|1\|_{L^2(K_k)} \|v_{xx}\|_{L^2(K_k)} \quad (\text{Cauchy-Schwarz}) \\
 |e_x|_{K_k} &\leq \underbrace{\sqrt{\int_{K_k} 1^2 \, ds}}_{h_k} \times \sqrt{\int_{K_k} v_{xx}^2 \, ds} \quad \text{hvor } h_k = x_{k+1} - x_k \\
 |e_x|_{K_k}^2 &\leq h_k \int_{K_k} v_{xx}^2 \, ds \\
 h &= \max_{k \in \mathcal{J}_h} h_k \\
 |e|_{H^1(K)}^2 &= \int_{K_k} e_x^2 \, dx \leq h_k^2 \int_{K_k} v_{xx}^2 \, dx \\
 |e|_{H^1(\Omega)}^2 &= \sum_{k \in \mathcal{J}_h} |e_x|_{H^1(K)}^2 \leq h^2 \sum_{k \in \mathcal{J}_h} \int_{K_k} v_{xx}^2 \, dx = h^2 \int_{\Omega} v_{xx}^2 \, dx = h^2 |v|_{H^2(\Omega)}^2 \\
 \|e\|_{L^2(\Omega)} &\leq h^4 |v|_{H^2(\Omega)}^2 \\
 \|e\|_{H^1(\Omega)} &= \|e\|_{L^2(\Omega)} + |e|_{H^1(\Omega)} \leq (h^4 + h^2) |v|_{H^2(\Omega)}^2 \leq K^2 h^2 \|v\|_{H^2(\Omega)}^2
 \end{aligned}$$

Hvis vi bruker Ceas lemma med $v = \Pi_h^1(u)$, får vi:

$$\|u - u_h\|_{H^1(\Omega)} \leq \frac{M}{\alpha} |u|_{H^2(\Omega)} h \leq C |u|_{H^2(\Omega)} h$$

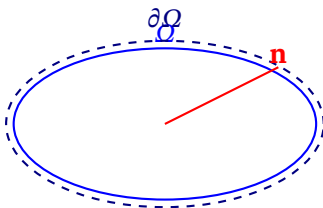
Er en første ordens metode. Fungerer kun hvis $u \in H^2(\Omega)$.

3.8.1 Modell problem

$$-\Delta u(x) = f(x) \quad x \in \Omega$$

$$u(x) = 0 \quad x \in \partial\Omega$$

- Hva er den svake formuleringen av dette problemet?
- Er LM-teoremet oppfylt?
- Hva er X_h^1 i dette tilfellet?



- Ω er en åpen, begrenset og sammenhengende delmengde.
- $\partial\Omega$ er den lukkede mengden av alle punkter i Ω .
- $\overline{\Omega} = \Omega \cup \partial\Omega$ er den lukkede mengden av alle punkter i Ω .
- $v : \Omega \rightarrow \mathbb{R}$
- $\mathbf{F} : \Omega \rightarrow \mathbb{R}^2$

Vektor notasjon

Gradient

$$\nabla u(x) = \begin{pmatrix} \frac{\partial u}{\partial x_1} \\ \frac{\partial u}{\partial x_2} \end{pmatrix} = \begin{pmatrix} u_{x_1} \\ u_{x_2} \end{pmatrix} = (u_{x_1}, u_{x_2})^T$$

Divergens

$$\nabla \cdot \mathbf{F} = \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} = F_{x_1} + F_{x_2}$$

Laplace operator

$$\Delta u(x) = \nabla^2 u(x) = \nabla \cdot \nabla u(x) = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = u_{x_1 x_1} + u_{x_2 x_2}$$

Theorem 10. Divergensteoremet

$$\int_{\Omega} \nabla \cdot \mathbf{F} \, dx = \int_{\partial\Omega} \mathbf{F} \cdot \mathbf{n} \, ds$$

- \mathbf{F} er en vektorfelt.
- \mathbf{n} er en normalvektor til grensen $\partial\Omega$.
- ds er et infinitesimalt areal på grensen $\partial\Omega$.
- dx er et infinitesimalt areal i domenet Ω .
- $\int_{\partial\Omega} \mathbf{F} \cdot \mathbf{n} \, ds$ er fluks gjennom grensen $\partial\Omega$.
- $\int_{\Omega} \nabla \cdot \mathbf{F} \, dx$ er divergensen av vektorfeltet \mathbf{F} i domenet Ω .

Laplace operator

$$\Delta u(x) = \nabla^2 u(x) = \nabla \cdot \nabla u(x) = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = u_{x_1 x_1} + u_{x_2 x_2}$$

Vektor Kalkulus

$$\nabla \cdot (v\mathbf{F}) = \nabla v \cdot \mathbf{F} + v \nabla \cdot \mathbf{F}$$

Green's teorem La $u : \Omega \rightarrow \mathbb{R}$ og $\mathbf{F} = \nabla u$ være en vektorfelt. Da har vi at:

$$\int_{\Omega} \nabla v \cdot \nabla u \, d\Omega + \int_{\Omega} v \Delta u \, d\Omega = \oint_{\partial\Omega} v \nabla u \cdot \mathbf{n} \, d\gamma$$

$-\Delta u = f$ på Ω og $u = 0$ på $\partial\Omega$ gir oss:

$$\int_{\Omega} \nabla v \cdot \nabla u \, d\Omega = \oint_{\partial\Omega} v \nabla u \cdot \mathbf{n} \, d\gamma$$

Krever at $v = 0$ på $\partial\Omega$.

Definerer:

$$H^1(\Omega) = \{v : \Omega \rightarrow \mathbb{R}, v, v_{x_1}, v_{x_2} \in L^2(\Omega)\}$$

$$L^2(\Omega) = \{v : \int_{\Omega} v^2 \, dx < \infty\}$$

$$\langle u, v \rangle_{L^2(\Omega)} = \int_{\Omega} uv \, d\Omega, \quad \|v\|_{L^2(\Omega)}^2 = \int_{\Omega} v^2 \, d\Omega$$

$$H_0^1(\Omega) = \{v \in H^1(\Omega), v = 0 \text{ på } \partial\Omega\}$$

Variasjonsformulering (Variational form) Finn $u \in H_0^1(\Omega)$ slik at $\int_{\Omega} \nabla u \cdot \nabla v \, d\Omega = \int_{\Omega} f v \, d\Omega$ for alle $v \in H_0^1(\Omega)$.

$$V = H_0^1(\Omega)$$

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, d\Omega, \quad F(v) = \int_{\Omega} f v \, d\Omega$$

Example 10. Poisson

Sterk form Finn $u : \Omega \rightarrow \mathbb{R}$ slik at

$$\begin{aligned} -\Delta u &= f \quad \text{i } \Omega, \\ u &= 0 \quad \text{på } \partial\Omega, \end{aligned}$$

Svak form Finn $u \in V$ slik at

$$\langle \nabla u, \nabla \phi \rangle = \langle f, \phi \rangle \quad \forall \phi \in V := H_0^1(\Omega) = \{u \in H^1(\Omega) : u = 0 \text{ på } \partial\Omega\}$$

Diskret svak form Finn $u_h \in V_h \subset V$ slik at

$$\langle \nabla u_h, \nabla \phi_i \rangle = \langle f, \phi_i \rangle \quad \forall i = 1, \dots, n,$$

Løsning Løs det lineære systemet $A\vec{u} = \vec{f}$, der

$$\begin{aligned} \langle \nabla u_h, \nabla \phi_i \rangle &= \langle f, \phi_i \rangle, \quad \Leftrightarrow \quad A\vec{u} = \vec{f} \\ A_{ij} &= \langle \nabla \phi_j, \nabla \phi_i \rangle, \quad \vec{u}_i, \quad \vec{f}_i = \langle f, \phi_i \rangle \end{aligned}$$

3.8.2 Ekvivalente former

Sterk form

Finn $u \in \mathcal{C}^2(\Omega)$ slik at

$$\begin{aligned} -u'' &= f & \text{i } \Omega, \\ u &= 0 & \text{på } \partial\Omega, \end{aligned}$$

Svak form

Finn $u \in V = H_0^1(\Omega)$ slik at

$$\langle u', \phi' \rangle = \langle f, \phi \rangle \quad \forall \phi \in V := H_0^1(\Omega)$$

Energiminimeringsform

La $F : V \rightarrow \mathbb{R}$ være en energifunksjon definert som

$$F(u) = \frac{1}{2} \langle u', u' \rangle - \langle f, u \rangle$$

Der $F(u)$ representerer energien i systemet. Den første termen representerer den kinetiske energien, mens den andre termen representerer den potensielle energien.

Finn $u \in V$ slik at

$$F(u) \leq F(\phi) \quad \forall \phi \in V$$

Dette er en variabel minimeringsoppgave der vi ønsker å finne den funksjonen u som minimerer energien i systemet.¹

3.9 Eksempler

3.9.1 Poisson-ligningen

Example 11. Poisson-ligningen

Poisson-ligningen er en vanlig [Partiell differensiallikning](#) som beskriver mange fysiske fenomener, inkludert varmeledning og elektrisk potensial. Den kan skrives som:

$$\begin{cases} -\Delta u(x) = f(x), & x \in (0, 1) \\ u(0) = \alpha, \quad u(1) = \beta & \text{(randbetingelser)} \end{cases} \quad (3.5)$$

$$\mathbf{F} = - \int_0^1 f(x) \mathbf{N}(x) dx \quad (3.6)$$

¹Det er ikke alltid at denne formen eksisterer for problemet vårt.

$$\begin{cases} -\frac{d^2 u}{dx^2} = f(x), & x \in (0, 1) \\ u(0) = \alpha, \quad u(1) = \beta \quad (\text{randbetingelser}) \end{cases} \quad (3.7)$$

La $f(x) = \bar{f}$ være konstant. Vi antar at $u(x_m)$ er ukjent for $m = 1, \dots, M$ punkter (noder) i det diskrete domenet Ω_h .

I mellom nodene definerer vi *elementene* \Rightarrow *formfunksjoner* $N_i(x)$.

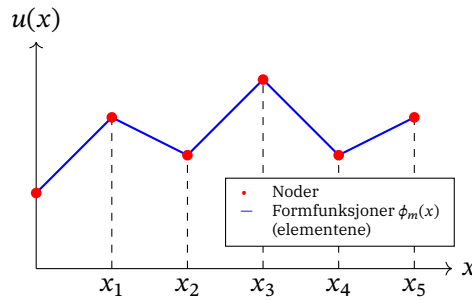


Figure 3.13: Tilfeldig valgt formfunksjoner $N_i(x)$ mellom nodene x_i .

Definerer formfunksjonene som:

$$\phi_i(x) = \begin{cases} 1 - 2|x - x_i|, & |x - x_i| < 0.5 \\ 0, & \text{ellers} \end{cases}$$

Og testfunksjonene som:

$$v(x) = \sum_{j=1}^n v_j \phi_j(x) = \mathbf{v}^T \boldsymbol{\phi}(x)$$

Den svake formuleringen blir:

$$\begin{aligned} \int_0^1 \left(\sum_{j=1}^n v_j \phi_j'(x) \right) \left(\sum_{i=1}^n u_i \phi_i'(x) \right) dx &= \int_0^1 \left(\sum_{j=1}^n v_j \phi_j(x) \right) f(x) dx \\ \mathbf{v}^T \int_0^1 \boldsymbol{\phi}' \boldsymbol{\phi}'^T dx \mathbf{u} &= \mathbf{v}^T \int_0^1 -f(x) \boldsymbol{\phi} dx \\ \mathbf{v}^T \mathbf{K} \mathbf{u} &= \mathbf{v}^T \mathbf{F} \\ \mathbf{K} \mathbf{u} &= \mathbf{F} \end{aligned}$$

Hvor:

$$\mathbf{K} = \int_0^1 \boldsymbol{\phi}' \boldsymbol{\phi}'^T dx = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$

$$\mathbf{F} = -\bar{f} \int_0^1 \mathbf{N}(x) dx = -\bar{f} \begin{bmatrix} 0.1 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.1 \end{bmatrix}$$

Løser ligningssystemet for å finne koeffisientene u_i :

$$\mathbf{u} = \mathbf{K}^{-1} \mathbf{F}$$

3.9.2 Svak formulering av Poisson-ligningen

Vi kan skrive den svake formuleringen som:

$$\int_0^1 v'(x) u'(x) dx = \int_0^1 v(x) f(x) dx,$$

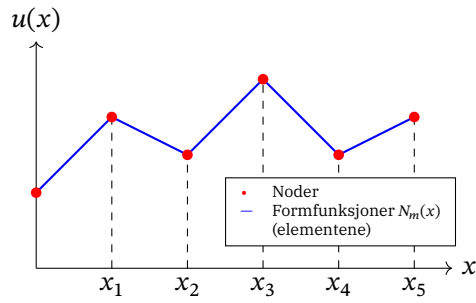
Velger [Basisfunksjoner](#) til å være:

$$\phi_i(x) = \begin{cases} 1 - 2|x - x_i|, & |x - x_i| < 0.5 \\ 0, & \text{ellers} \end{cases}$$

Med [Testfunksjonene](#) $v(x) = \sum_{j=1}^n v_j \phi_j(x)$.

Da får vi:

$$\begin{aligned} \int_0^1 \left(\sum_{j=1}^n v_j \phi_j'(x) \right) \left(\sum_{i=1}^n u_i \phi_i'(x) \right) dx &= \int_0^1 \left(\sum_{j=1}^n v_j \phi_j(x) \right) f(x) dx \\ \mathbf{v}^T \int_0^1 \boldsymbol{\phi}' \boldsymbol{\phi}'^T dx \mathbf{u} &= \mathbf{v}^T \int_0^1 -f(x) \boldsymbol{\phi} dx \\ \mathbf{v}^T \mathbf{K} \mathbf{u} &= \mathbf{v}^T \mathbf{F} \\ \mathbf{K} \mathbf{u} &= \mathbf{F} \end{aligned}$$

Figure 3.14: Tilfeldig valgt formfunksjoner $N_i(x)$ mellom nodene x_i .

$$\mathbf{F} = -\bar{f} \int_0^1 \mathbf{N}(x) dx = -\bar{f} \begin{bmatrix} 0.1 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.1 \end{bmatrix}$$

Example 12. Eksamen Vår 2024

Vi har ett en-dimesjonalt elliptisk randverdi-problem.

$$\begin{aligned} -\mathcal{L}u &= -\frac{d}{dx} \left((1+x) \frac{du}{dx} \right) + 2u = 2x \quad x \in (0, 1) \\ u(0) &= \sqrt{2}, \quad u(1) = \sqrt{3} \end{aligned}$$

La nå $M \in \mathbb{N}$ og $T_h = \bigcup_{i=1}^M K_i$ være en triangulering av $(0, 1)$ med $K_i = (x_{i-1}, x_i)$, $x_i = ih$ og $h = \frac{1}{M}$.

$$T_h = \{K_1, K_2, \dots, K_M\} = \{(x_0, x_1), (x_1, x_2), \dots, (x_{M-1}, x_M)\}$$

Finn **elementmatrisen** A^{K_i} og **lastvektoren** F^{K_i} for Lagrange FEM i \mathbb{P}_1 på trianguleringen T_h for problemet i (??).

På \mathbb{P}_1 er standard Lagrange-FEM-basisfunksjonene gitt ved:

$$\begin{aligned} \phi_1(x) &= \frac{x_i - x}{h} & \phi_1(x_{i-1}) &= 1, \phi_1(x_i) = 0 \\ \phi_2(x) &= \frac{x - x_{i-1}}{h} & \phi_2(x_{i-1}) &= 0, \phi_2(x_i) = 1 \end{aligned}$$

Først finner vi testfunksjonene $v \in \mathcal{H}_0^1(0, 1)$ og basisfunksjonene $\phi_i(x)$ for $i = 1, \dots, M$.

$$v(x) = \sum_{i=1}^M v_i \phi_i(x)$$

Svak formulering

Den svake formuleringen av problemet i (??) finner man ved å gange PDE med testfunksjonen v på begge sidene og integrere over $(0, 1)$:

$$\begin{aligned} - \int_0^1 \mathcal{L}u(x)v(x) dx &= 0 \\ \int_0^1 \left[-\frac{d}{dx} \left((1+x) \frac{du}{dx} \right) + 2u \right] v(x) dx &= 2 \int_0^1 xv(x) dx \\ - \int_0^1 \frac{d}{dx} ((1+x)u'(x)) v(x) dx + 2 \int_0^1 u(x)v(x) dx &= 2 \int_0^1 xv(x) dx \end{aligned}$$

For det første leddet i LHS bruker vi delvis integrasjon:

$$\begin{aligned} - \int_0^1 \overbrace{\frac{d}{dx} ((1+x)u'(x))}^{w'(x)} v(x) dx &= - [w(x)v(x)]_0^1 + \int_0^1 w(x)v'(x) dx \\ &= - [(1+x)u'(x)v(x)]_0^1 + \int_0^1 (1+x)u'(x)v'(x) dx \\ &= - [(1+1)u'(1)v(1) - (1+0)u'(0)v(0)] + \int_0^1 (1+x)u'(x)v'(x) dx \\ &= - [2u'(1)v(1) - u'(0)v(0)] + \int_0^1 (1+x)u'(x)v'(x) dx \\ &= 0 + \int_0^1 (1+x)u'(x)v'(x) dx \end{aligned}$$

Siden $v \in \mathcal{H}_0^1(0, 1)$, så er $v(0) = v(1) = 0$ får vi at $[w(x)v(x)]_0^1 = 0$.

Dermed kan vi skrive om den svake formuleringen til:

$$\begin{aligned} \int_0^1 (1+x)u'(x)v'(x) dx + 2 \int_0^1 u(x)v(x) dx &= 2 \int_0^1 xv(x) dx \\ \int_0^1 (1+x)u'(x)v'(x) + 2u(x)v(x) dx &= 2 \int_0^1 xv(x) dx \end{aligned}$$

Vi approksimerer u med Lagrange-FEM-basisfunksjoner $\phi_i(x)$ i \mathbb{P}_1 :

$$\begin{aligned} u(x) &\approx \sum_{i=1}^M u_i \phi_i(x) = \mathbf{u}^T \boldsymbol{\phi}(x) \\ v(x) &\approx \sum_{i=1}^M v_i \phi_i(x) = \mathbf{v}^T \boldsymbol{\phi}(x) \end{aligned}$$

Setter inn i den svake formuleringen:

$$\begin{aligned}
\int_0^1 (1+x) \left(\sum_{i=1}^M u_i \phi'_i(x) \right) \left(\sum_{j=1}^M v_j \phi'_j(x) \right) dx + 2 \left(\sum_{i=1}^M u_i \phi_i(x) \right) \left(\sum_{j=1}^M v_j \phi_j(x) \right) dx &= 2 \int_0^1 x \left(\sum_{i=1}^M u_i \phi_i(x) \right) \left(\sum_{j=1}^M v_j \phi_j(x) \right) dx \\
\int_0^1 (1+x) (\mathbf{u}^\top \boldsymbol{\phi}'(x)) (\mathbf{v}^\top \boldsymbol{\phi}'(x)) dx + \int_0^1 2 (\mathbf{u}^\top \boldsymbol{\phi}(x)) (\mathbf{v}^\top \boldsymbol{\phi}(x)) dx &= 2 \int_0^1 x (\mathbf{u}^\top \boldsymbol{\phi}(x)) (\mathbf{v}^\top \boldsymbol{\phi}(x)) dx \\
\int_0^1 (1+x) (\mathbf{u}^\top \boldsymbol{\phi}'(x)) (\boldsymbol{\phi}'(x)^\top \mathbf{v}) dx + \int_0^1 2 (\mathbf{u}^\top \boldsymbol{\phi}(x)) (\boldsymbol{\phi}(x)^\top \mathbf{v}) dx &= 2 \int_0^1 x (\mathbf{u}^\top \boldsymbol{\phi}(x)) (\boldsymbol{\phi}(x)^\top \mathbf{v}) dx \\
\mathbf{u}^\top \int_0^1 (1+x) \boldsymbol{\phi}'(x) \boldsymbol{\phi}'(x)^\top dx \mathbf{v} + 2 \mathbf{u}^\top \int_0^1 \boldsymbol{\phi}(x) \boldsymbol{\phi}(x)^\top dx \mathbf{v} &= 2 \mathbf{u}^\top \int_0^1 x \boldsymbol{\phi}(x) \boldsymbol{\phi}(x)^\top dx \mathbf{v} \\
\mathbf{u}^\top \left[\int_0^1 (1+x) \boldsymbol{\phi}'(x) \boldsymbol{\phi}'(x)^\top + 2 \boldsymbol{\phi}(x) \boldsymbol{\phi}(x)^\top dx \right] \mathbf{v} &= 2 \mathbf{u}^\top \left[\int_0^1 x \boldsymbol{\phi}(x) \boldsymbol{\phi}(x)^\top dx \right] \mathbf{v}
\end{aligned}$$

Vi definerer nå elementmatrisen A^{Ki} og lastvektoren \mathbf{F}^{Ki} som:

$$\begin{aligned}
A^{Ki} &= \int_0^1 (1+x) \boldsymbol{\phi}'(x) \boldsymbol{\phi}'(x)^\top + 2 \boldsymbol{\phi}(x) \boldsymbol{\phi}(x)^\top dx \\
\mathbf{F}^{Ki} &= 2 \int_0^1 x \boldsymbol{\phi}(x) \boldsymbol{\phi}(x)^\top dx
\end{aligned}$$

$$\mathbf{u}^\top A^{Ki} \mathbf{v} = \mathbf{u}^\top \mathbf{F}^{Ki}$$

$$\begin{aligned}
A^{Ki} &= \int_0^1 (1+x) \boldsymbol{\phi}'(x) \boldsymbol{\phi}'(x)^\top + 2 \boldsymbol{\phi}(x) \boldsymbol{\phi}(x)^\top dx \\
&= \int_0^1 (1+x) \begin{bmatrix} \phi'_1(x) & \phi'_2(x) \end{bmatrix} \begin{bmatrix} \phi'_1(x) \\ \phi'_2(x) \end{bmatrix} + 2 \begin{bmatrix} \phi_1(x) & \phi_2(x) \end{bmatrix} \begin{bmatrix} \phi_1(x) \\ \phi_2(x) \end{bmatrix} dx \\
&= \int_0^1 (1+x) \begin{bmatrix} \phi'_1(x) \phi'_1(x) & \phi'_1(x) \phi'_2(x) \\ \phi'_2(x) \phi'_1(x) & \phi'_2(x) \phi'_2(x) \end{bmatrix} + 2 \begin{bmatrix} \phi_1(x) \phi_1(x) & \phi_1(x) \phi_2(x) \\ \phi_2(x) \phi_1(x) & \phi_2(x) \phi_2(x) \end{bmatrix} dx \\
&= \int_0^1 (1+x) \begin{bmatrix} \phi'_1(x)^2 & \phi'_1(x) \phi'_2(x) \\ \phi'_2(x) \phi'_1(x) & \phi'_2(x)^2 \end{bmatrix} dx + 2 \int_0^1 \begin{bmatrix} \phi_1(x)^2 dx & \phi_1(x) \phi_2(x) dx \\ \phi_2(x) \phi_1(x) dx & \phi_2(x)^2 dx \end{bmatrix} dx
\end{aligned}$$

Hvor de deriverte av Lagrange-basisfunksjonene er:

$$\begin{aligned}
\phi'_1(x) &= -\frac{1}{h}, & \phi'_1(x)^2 &= \frac{1}{h^2}, \\
\phi'_2(x) &= \frac{1}{h}, & \phi'_2(x)^2 &= \frac{1}{h^2}, \\
\phi'_1(x) \phi'_2(x) &= \phi'_2(x) \phi'_1(x) = \frac{-1}{h^2} = -\frac{1}{h^2}
\end{aligned}$$

Glossary

This document is incomplete. The external file associated with the glossary ‘main’ (which should be called `main.gls`) hasn’t been created.

Check the contents of the file `main.gls`. If it’s empty, that means you haven’t indexed any of your entries in this glossary (using commands like `\gls` or `\glsadd`) so this list can’t be generated. If the file isn’t empty, the document build process hasn’t been completed.

If you don’t want this glossary, add `nomain` to your package option list when you load `glossaries-extra.sty`. For example:

```
\usepackage[nomain]{glossaries-extra}
```

Try one of the following:

- Add `automake` to your package option list when you load `glossaries-extra.sty`. For example:

```
\usepackage[automake]{glossaries-extra}
```

- Run the external (Lua) application:
`makeglossaries-lite.lua "main"`
- Run the external (Perl) application:
`makeglossaries "main"`

Then rerun \LaTeX on this document.

This message will be removed once the problem has been fixed.