

M4

Del I
Analyse

1 Laplacetransform

Laplacetransform er en teknikk vi skal bruke til løse ordinære differensiallikninger. For det første er det en mye mer elegant teknikk enn den du lærte i M3, og for det andre takler den en bredere klasse av likninger.

Definisjon og grunnleggende egenskaper

La f være en funksjon definert for $t \geq 0$. Laplace-transformen til f er:

$$\mathcal{L}(f) = \int_0^\infty f(t)e^{-st} dt = F(s).$$

Først kan man spørre seg hvilke funksjoner det er naturlig å finne laplacetransformen til. Vi noterer oss at $\mathcal{L}(f)$ er definert ved et uegentlig integral, og at dette bør konvergere.

Teorem 1.1. *La f en stykkvis kontinuertlig funksjon, og la a og $M > 0$ være reelle tall slik at*

$$|f(t)| \leq Me^{at} \quad \text{for } t \geq 0.$$

Integralet

$$\int_0^\infty f(t)e^{-st} dt.$$

konvergerer absolutt dersom $s > a$.

Bevis. Vi beregner:

$$\left| \int_0^\infty f(t)e^{-st} dt \right| \leq \int_0^\infty |f(t)|e^{-st} dt \leq M \int_0^\infty e^{at}e^{-st} dt \leq M \int_0^\infty e^{(a-s)t} dt = \frac{M}{s-a}. \quad \square$$

Eksempel 1.2. Vi kan ikke beregne

$$\mathcal{L}(e^{t^2}) = \int_0^\infty e^{t^2-st} dt,$$

for e^{t^2} vokser for fort. Siden

$$\lim_{t \rightarrow \infty} e^{t^2-st} = \infty$$

kan integralet ikke konvergere. \triangle

Eksempel 1.3. La f være funksjonen

$$f(x) = \begin{cases} 1 & \text{rasjonale } x \\ 0 & \text{irrasjonale } x \end{cases}$$

Vi kan ikke laplacetransformere denne funksjonen, for den er ikke stykkvis kontinuertlig. Den er faktisk diskontinuertlig overalt, og ikke en gang riemannintegrerbar. \triangle

Eksempel 1.4. Så lenge $s > a$, kan vi fint beregne

$$\mathcal{L}(e^{at}) = \int_0^\infty e^{(a-s)t} dt = \frac{1}{s-a}. \quad \triangle$$

Eksempel 1.5. Vi kan også beregne

$$\mathcal{L}(0) = \int_0^\infty 0e^{-st} dt = 0. \quad \triangle$$

Heretter skal vi alltid anta at s er valgt slik at integralet konvergerer. Vi skal også anta f er stykkvis kontinuertlig og tilfredsstillende $|f| \leq Me^{at}$.

Et viktig spørsmål er hvorvidt man kan invertere laplacetransformen: Dersom to funksjoner har samme laplacetransform, må de være like som funksjoner? Slik du har lært funksjoner til nå, er svaret definitivt nei, som følgende eksempel viser.

Eksempel 1.6. La

$$g(t) = \begin{cases} e^{at} & t \neq 1 \\ 0 & t = 1 \end{cases}$$

Som i eksemplet over, kan vi for $s > a$ beregne

$$\begin{aligned} \mathcal{L}(g) &= \int_0^1 e^{(a-s)t} dt + \int_1^\infty e^{(a-s)t} dt \\ &= \frac{1}{s-a}, \end{aligned}$$

altså den samme som i eksempel 1.4. \triangle

Eksemplet over kan fremstå som litt patologisk, men det er et viktig matematisk poeng at et integral ikke endres av at man bytter ut enkelte funksjonsverdier. I og med at laplacetransformen er et integral, kan vi med andre ord ikke forvente at det skal finnes en entydig invers, med mindre man er litt streng på hvilke funksjoner man laplacetransformerer.

En alternativ løsning er å omdefinere hva man mener med at to funksjoner er like; i avansert matematisk analyse bryr man seg ikke om at funksjoner har forskjellige funksjonsverdier i et og annet punkt, og da blir det lettere å snakke om invers laplacetransform. Vi har ikke mulighet til å gå inn på dette, men det er egentlig ikke så viktig for en ingeniør. Ingeniører bruker laplacetransform til å løse differensiallikninger, og da vil det gjerne være åpenbart fra anvendelsen hva som er den korrekte inverstransformen, selv om det ikke finnes noe matematisk entydig valg. Følgende teorem tar vi nå med uansett.

Teorem 1.7. *La f og g være stykkvis kontinuerlige funksjoner. Dersom*

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

må

$$f = g$$

overalt der f og g er kontinuertlige.

Dette teoremet forteller at dersom vi ser bort fra funksjonsverdier akkurat i sprang, vil den inverse laplacetransformen være entydig. Beviset er dessverre altfor hardt for dette kurset. Det finnes formler for invers laplacetransform. Men de er også for harde for oss, så vi skal bruke tabell i stedet. Dette er den vanligste måten å invertere laplacetransformer på.

Regneregler

Man løser ordinære differensiallikninger ved å bruke laplacetransform til å skrive dem om til algebraiske likninger. Det er duket for å introdusere triksene vi skal bruke.

Teorem 1.8. *Laplacetransformen er lineær. Dersom a og b er reelle tall, er*

$$\mathcal{L}(af + bg) = a\mathcal{L}(f) + b\mathcal{L}(g).$$

Beviset for dette teoremet er så enkelt at vi dropper det.

Eksempel 1.9.

$$\begin{aligned} \mathcal{L}(\cos t) &= \int_0^\infty e^{-st} \cos t \, dt \\ &= \frac{1}{2} \int_0^\infty (e^{it} + e^{-it}) e^{-st} \, dt \\ &= \frac{1}{2} \int_0^\infty e^{(i-s)t} + e^{-(i+s)t} \, dt \\ &= \frac{1}{2} \left(\frac{1}{s-i} + \frac{1}{s+i} \right) = \frac{s}{s^2+1}. \end{aligned}$$

Her har vi brukt en omskrivning av Eulers formel:

$$\cos t = \frac{e^{it} + e^{-it}}{2}. \quad \triangle$$

Eksempel 1.10.

$$\begin{aligned} \mathcal{L}(\cosh t) &= \mathcal{L}\left(\frac{e^t + e^{-t}}{2}\right) \\ &= \frac{1}{2} (\mathcal{L}(e^t) + \mathcal{L}(e^{-t})) \\ &= \frac{1}{2} \left(\frac{1}{s-1} + \frac{1}{s+1} \right) = \frac{s}{s^2-1}. \quad \triangle \end{aligned}$$

Teorem 1.11. *Dersom $s > 0$ og $|f'(t)| \leq Me^{at}$ er stykkvis kontinuert, er*

$$\mathcal{L}(f') = s\mathcal{L}(f) - f(0).$$

Bevis. La $s > 0$. Vi beregner

$$\begin{aligned} \mathcal{L}(f') &= \int_0^\infty f'(t) e^{-st} \, dt \\ &= f(t) e^{-st} \Big|_0^\infty + s \int_0^\infty f(t) e^{-st} \, dt \\ &= s\mathcal{L}(f) - f(0). \quad \square \end{aligned}$$

Eksempel 1.12.

$$\begin{aligned} \mathcal{L}(\sinh t) &= s\mathcal{L}(\cosh t) + \cosh(0) \\ &= -s \frac{s}{s^2-1} + 1 = \frac{1}{s^2+1}. \quad \triangle \end{aligned}$$

Under liknende betingelser på de høyere ordens deriverte, vil gjentatt bruk av teorem 1.14 gi formler av typen

$$\mathcal{L}(f'') = s^2\mathcal{L}(f) - sf(0) - f'(0)$$

og

$$\mathcal{L}(f''') = s^3\mathcal{L}(f) - s^2f(0) - sf'(0) - f''(0)$$

og så videre.

Eksempel 1.13. Vi beregner først

$$\mathcal{L}(1) = \int_0^\infty e^{-st} \, dt = \frac{1}{s}.$$

La nå $f(t) = t^n$. Den n -te deriverte av f er:

$$f^n(t) = n!$$

og siden $f^k(0) = 0$ uansett k , får vi

$$n!\mathcal{L}(1) = s^n\mathcal{L}(t^n),$$

slik at

$$\mathcal{L}(t^n) = \frac{n!}{s^{n+1}}. \quad \triangle$$

Vi tar to teoremer til.

Teorem 1.14. *La $g(t) = \int_0^t f(u) \, du$. Dersom g tilfredsstiller vekstkravet, har vi:*

$$\mathcal{L}(g) = \frac{1}{s}\mathcal{L}(f)$$

Bevis. La $s > 0$. Vi beregner

$$\begin{aligned} \mathcal{L}(g) &= \int_0^\infty g(t) e^{-st} \, dt \\ &= -\frac{1}{s} g(t) e^{-st} \Big|_0^\infty + \frac{1}{s} \int_0^\infty f(t) e^{-st} \, dt \\ &= \frac{1}{s} \mathcal{L}(f). \quad \square \end{aligned}$$

Eksempel 1.15.

$$\begin{aligned} \mathcal{L}(\sin t) &= \frac{1}{s} \mathcal{L}(\cos t) \\ &= \frac{1}{s} \frac{s}{s^2+1} = \frac{1}{s^2+1}. \quad \triangle \end{aligned}$$

Teorem 1.16. *La $g(t) = tf(t)$ og $\mathcal{L}(f) = F$. Anta at F er deriverbar. Da har vi:*

$$\mathcal{L}(g) = -F'$$

Bevis. Dersom f er stykkvis kontinuert og tilfredsstiller vekstkravet, gjelder dette også for g . Vi beregner

$$\begin{aligned} \mathcal{L}(g) &= \int_0^\infty tf(t) e^{-st} \, dt = \int_0^\infty -\frac{d}{ds} f(t) e^{-st} \, dt \\ &= -\frac{d}{ds} \int_0^\infty f(t) e^{-st} \, dt = -F'. \quad \square \end{aligned}$$

Eksempel 1.17.

$$\mathcal{L}(t \sin t) = -\frac{d}{ds} \frac{1}{s^2+1} = \frac{2s}{(s^2+1)^2}. \quad \triangle$$

To spesielle funksjoner

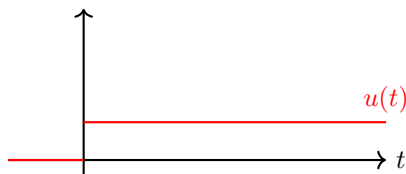
I det følgende skal vi beskrive noen helt nye typer funksjoner som dukker opp i anvendelser.

Heavisidefunksjonen

Heavisidefunksjonen er gitt ved

$$u(t) = \begin{cases} 0 & \text{for } t < 0 \\ 1 & \text{for } t \geq 0. \end{cases}$$

Man kan tenke på denne som en funksjon som slår noe på ved $t = 0$.

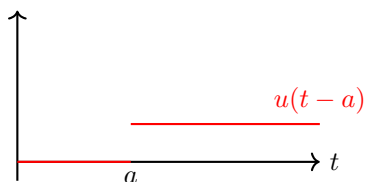


Eksempel 1.18.

$$u(t)e^t = \begin{cases} 0 & \text{for } t < 0 \\ e^t & \text{for } t \geq 0 \end{cases} \quad \triangle$$

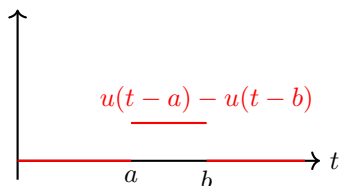
Vi kan slå på ved tiden $t = a$ istedet:

$$u(t-a) = \begin{cases} 0 & \text{for } t < a \\ 1 & \text{for } t \geq a, \end{cases}$$



Vi kan også slå på ved $t = a$ og av igjen ved $t = b$:

$$u(t-a) - u(t-b) = \begin{cases} 0 & \text{for } t < a \\ 1 & \text{for } a \leq t < b \\ 0 & \text{for } t \geq b. \end{cases}$$



Eksempel 1.19.

$$(u(t-a) - u(t-b))e^t = \begin{cases} 0 & \text{for } t < a \\ e^t & \text{for } a \leq t < b \\ 0 & \text{for } t \geq b \end{cases} \quad \triangle$$

Eksempel 1.20.

La $f(t) = u(t-a)$.

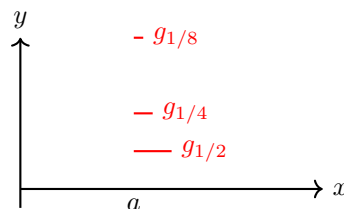
$$\mathcal{L}(f) = \int_a^\infty e^{-st} dt = \frac{e^{-sa}}{s} \quad \triangle$$

Deltafunksjonen

La

$$g_k(t-a) = \begin{cases} 1/k & \text{for } a < t < a+k \\ 0 & \text{ellers} \end{cases}$$

Slik ser de ut:



Vi definerer

$$\delta(t) = \lim_{k \rightarrow 0} g_k(t) = \begin{cases} \infty & \text{for } t = a \\ 0 & \text{ellers} \end{cases}$$

Deltafunksjonen brukes til å modellere impuls, altså energitilførsler der den påtrykte kraften er ekstremt høy og ekstremt kortvarig, for eksempel hammerslag. Man kan også tenke på den som noe som plukker ut funksjonsverdier:

$$\int_0^\infty f(t)\delta(t-a) dt = \lim_{k \rightarrow 0} \frac{1}{k} \int_a^{a+k} f(t) dt = f(a).$$

Eksempel 1.21.

$$\mathcal{L}(\delta(t-a)) = \int_0^\infty \delta(t-a)e^{-st} dt = e^{-as} \quad \triangle$$

Deltafunksjonen er strengt tatt ikke noe funksjon i ordets rette forstand, og heavisidefunksjonen er ikke deriverbar. Det kan allikevel være fruktbart å tenke på deltafunksjonen som et forsøk på å sette opp den deriverte til heavisidefunksjonen.

To skifteteoremer

Det neste teoremet kalles gjerne 's-skift'.

Teorem 1.22. La $g(t) = e^{at}f(t)$, $\mathcal{L}(f) = F$ og $\mathcal{L}(g) = G$. Da har vi

$$G(s) = F(s-a).$$

Bevis. Dersom $f(t)$ er stykkvis kontinuert og tilfredsstiller vekstkravet, gjelder dette også for $e^{at}f(t)$.

$$\begin{aligned} \mathcal{L}(g) &= \int_0^\infty e^{at}f(t)e^{-st} dt \\ &= \int_0^\infty f(t)e^{(a-s)t} dt \\ &= \int_0^\infty f(t)e^{-(s-a)t} dt = F(s-a). \quad \square \end{aligned}$$

Eksempel 1.23.

$$\mathcal{L}(e^{at} \sin(t)) = \frac{1}{(s-a)^2 + 1}$$

$$\mathcal{L}(e^{at} \cos(t)) = \frac{s-a}{(s-a)^2 + 1} \quad \triangle$$

Følgende teorem kalles *t*-skift.

Teorem 1.24. La $g(t) = f(t - a)u(t - a)$, $\mathcal{L}(f) = F$ og $\mathcal{L}(g) = G$. Vi har

$$G(s) = e^{-as}F(s).$$

Bevis.

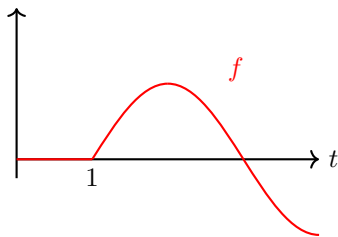
$$\begin{aligned} G(s) &= \int_0^\infty u(t-a)f(t-a)e^{-st} dt \\ &= \int_a^\infty f(t-a)e^{-st} dt \\ &= \int_0^\infty f(v)e^{-s(v+a)} dv \\ &= e^{-sa} \int_0^\infty f(v)e^{-sv} dv = e^{-as}F(s). \quad \square \end{aligned}$$

Eksempel 1.25. La

$$f(t) = u(t-1)\sin(t-1) = \begin{cases} 0 & \text{for } t < 1 \\ \sin(t-1) & \text{for } t \geq 1 \end{cases}$$

Vi beregner

$$\mathcal{L}(f) = \frac{e^{-s}}{s^2 + 1}. \quad \triangle$$



Konvolusjon

En konvolusjon mellom f og g , er et integral på formen

$$f * g = \int f(p)g(t-p) dp.$$

Det finnes mange typer konvolusjoner, og det er forskjellige integrasjonsgrenser som skiller dem. Hvilken type integrasjonsgrenser som er mest relevant, kommer litt an på anvendelsen.

I dette kapitlet skal vi bruke konvolusjonen

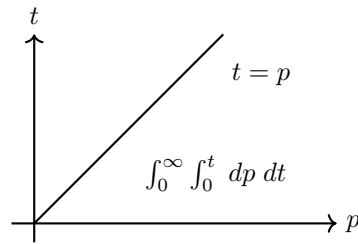
$$f * g = \int_0^t f(p)g(t-p) dp$$

Når vi løser differensiallikninger med laplacetransform, kan vi støte på produkter av laplacetransformere. Følgende teorem, som kalles konvolusjonsteoremet, forteller oss hvordan vi skal nøste opp i et slikt produkt.

Teorem 1.26. Anta at både $\mathcal{L}(f * g)$, $\mathcal{L}(f)$ og $\mathcal{L}(g)$ eksisterer. Da er

$$\mathcal{L}(f * g) = \mathcal{L}(f)\mathcal{L}(g).$$

Bevis. Vi laplacetransformerer $f * g$ og bytter integrasjonsvariable, slik som i M2. Hjelpfiguren under kan være grei å ha i bakhodet.



$$\begin{aligned} \mathcal{L}(f * g) &= \int_0^\infty \int_0^t f(p)g(t-p) dp e^{-st} dt \\ &= \int_0^\infty \int_p^\infty f(p)g(t-p)e^{-st} dt dp \\ &= \int_0^\infty \int_0^\infty f(p)g(u)e^{-s(u+p)} du dp \\ &= \int_0^\infty f(p)e^{-sp} dp \int_0^\infty g(u)e^{-su} du \\ &= \mathcal{L}(f)\mathcal{L}(g) \quad \square \end{aligned}$$

Å forklare hva konvolusjon er sånn rent fysisk sett, er ikke så enkelt. Men mange artige ting er basert på konvolusjon. Noen eksempler er reverb-knappen på gitarforsterkeren din, bakgrunnsuskarpheiten i det vakre konfirmasjonsbildet ditt, eller autofokusskanningen på speilreflekskameraet du fikk til ovennevnte konfirmasjon.

Differensiallikninger

Oppskriften for å løse differensiallikninger med laplacetransform, er alltid å laplacetransformere hele differensiallikningen, bruke regnereglene, løse for laplacetransformen til den ukjente, og inverstransformere. Vi begynner med to elementære eksempler.

Eksempel 1.27.

$$y' + y = 0 \quad y(0) = 1.$$

Vi laplacetransformerer likningen. Venstresiden blir

$$\begin{aligned} \mathcal{L}(y' + y) &= \mathcal{L}(y') + \mathcal{L}(y) \\ &= s\mathcal{L}(y) - 1 + \mathcal{L}(y) \\ &= (s+1)\mathcal{L}(y) - 1. \end{aligned}$$

Merk bruken av initialbetingelsen $y(0) = 1$. Høyresiden blir

$$\mathcal{L}(0) = 0.$$

Etter laplacetransformering står vi igjen med den algebraiske likningen

$$(s+1)\mathcal{L}(y) - 1 = 0,$$

slik at

$$\mathcal{L}(y) = \frac{1}{s+1}.$$

Dette er en laplacetransform vi har beregnet tidligere:

$$y(t) = e^{-t}. \quad \triangle$$

Eksempel 1.28. Vi løser likningen

$$y'' + y = 0 \quad y(0) = 1 \quad y'(0) = 0.$$

Vi laplacetransformerer likningen

$$\mathcal{L}(y'') + \mathcal{L}(y) = s^2 \mathcal{L}(y) - s + \mathcal{L}(y) = 0,$$

slik at

$$\mathcal{L}(y) = \frac{s}{s^2 + 1},$$

og

$$y(t) = \cos t. \quad \triangle$$

Følgende eksempel illustrerer derivasjonsregelen 1.16.

Eksempel 1.29. Vi løser initialverdiproblemet

$$ty' - 2y = 0 \quad y(0) = 0$$

Vi laplacetransformerer likningen

$$\mathcal{L}(ty') - 2\mathcal{L}(y) = 0,$$

og lar $Y = \mathcal{L}(y)$. Først kan vi skrive

$$\mathcal{L}(ty') = -\frac{d}{ds}(\mathcal{L}(y')) = -\frac{d}{ds}(sY) = -Y - sY'.$$

Setter vi denne inn i likningen, får vi

$$-Y - sY' - Y = -sY' - 3Y = 0.$$

eller

$$sY' + 3Y = 0.$$

Dette er en differensiallikning for Y . Integrerende faktor er s^2 , slik at

$$(s^3 Y)' = 0.$$

og

$$Y = \frac{C}{s^3}$$

Vi inverstransformerer og endrer litt på konstanten:

$$y = Ct^2$$

og bruker initialverdibetingelsen for andre gang:

$$y = t^2. \quad \triangle$$

De tre foregående eksemplene løses enklere med metodene du kan fra før. Men disse teknikkene kommer til kort i eksemplene vi skal ta nå. Her er et par gamle eksamensoppgaver.

Eksempel 1.30. Vi løser

$$y'' + y = \delta(t - a) - \delta(t - b)$$

med initialbetingelser

$$y(0) = y'(0) = 0.$$

Vi får

$$(s^2 + 1)\mathcal{L}(y) = e^{-as} - e^{-bs},$$

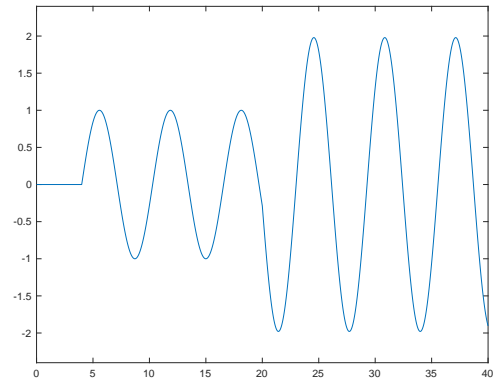
slik at

$$\mathcal{L}(y) = \frac{e^{-as}}{(s^2 + 1)} - \frac{e^{-bs}}{(s^2 + 1)}.$$

Vi bruker t -skift, og får

$$y(t) = \sin(t - a)u(t - a) - \sin(t - b)u(t - b).$$

Ligningens venstreside beskriver en kloss og en fjær på friksjonsfritt underlag. Ved tiden $t = 0$ er systemet i ro. Ved tiden $t = a$ blir klossen dengt av en hammer fra venstre, som gir opphav til svingningen $\sin(t - a)$. Ved tiden $t = b$ blir klossen dengt av en hammer fra høyre. Dette slaget gir en ny svingning $\sin(t - b)$ som legges oppå den gamle svingningen. \triangle



Noen ganger kan det være komplisert å invers-transformere. I det neste eksemplet må man bruke både t -skift og s -skift.

Eksempel 1.31.

$$y'' + 2y' + y = \delta(t - 1) \quad y(0) = y'(0) = 0.$$

Vi transformerer likningen til

$$s^2 \mathcal{L}(y) + 2s \mathcal{L}(y) + \mathcal{L}(y) = \mathcal{L}(\delta(t - 1))$$

slik at

$$\mathcal{L}(y) = \frac{e^{-s}}{(s + 1)^2}.$$

Her lukter det s -skift på grunn av $(s + 1)$, og t -skift på grunn av e^{-s} . Formelen for s -skift gir

$$\mathcal{L}(te^{-t}) = \frac{1}{(s + 1)^2}.$$

Vi bruker i tillegg t -skift, slik at

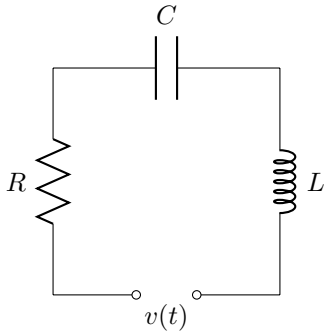
$$y = (t - 1)e^{-(t-1)}u(t - 1). \quad \triangle$$

Elektrofolk er forresten veldig glade i laplacetransform. Det er nok fordi de trenger spenning som kan slås av og på. Heavisidefunksjonen gir dem en mulighet for dette, og laplacetransform takler heavisidefunksjonen med letthet.

Eksempel 1.32. Strømmen $i(t)$ i kretsen under tilfredsstillende integro-differensiallikningen

$$Li'(t) + Ri(t) + \frac{1}{C} \int_0^t i(\tau) d\tau = \delta(t - 1),$$

der $R = 2$, $L = 1$, $C = 0.5$ og δ er Diracs deltafunksjon. Vi setter $i(0) = 0$ og finner strømmen $i(t)$.



Vi laplacetransformerer likningen, og får

$$s\mathcal{L}(i) + 2\mathcal{L}(i) + \frac{2}{s}\mathcal{L}(i) = e^{-s},$$

og løser vi for $\mathcal{L}(i)$, får vi

$$\mathcal{L}(i) = \frac{se^{-s}}{s^2 + 2s + 2}.$$

Her er det enklest å fullføre kvadratet $s^2 + 2s + 2 = (s + 1)^2 + 1$, og så skrive

$$\mathcal{L}(i) = \frac{(s + 1)e^{-s}}{(s + 1)^2 + 1} - \frac{e^{-s}}{(s + 1)^2 + 1}.$$

Formelen for s -skift gir

$$\mathcal{L}(e^{-t} \cos t) = \frac{s + 1}{(s + 1)^2 + 1}$$

og

$$\mathcal{L}(e^{-t} \sin t) = \frac{1}{(s + 1)^2 + 1}.$$

Inverstransformerer vi

$$\mathcal{L}(i) = \frac{(s + 1)e^{-s}}{(s + 1)^2 + 1} - \frac{e^{-s}}{(s + 1)^2 + 1}.$$

med t -skift, får vi altså

$$i(t) = e^{-(t-1)}u(t-1) (\cos(t-1) - \sin(t-1)).$$

Dette eksemplet modellerer en fugl som ved tiden $t = 1$ setter seg på en høyspentledning. Merk at strømmen er null frem til $t = 1$. I $t = 1$ opplever fuglen en voldsom påtrykt spenning i det den setter seg på høyspenten. Den faller umiddelbart av, slik at spenningsøkningen bare varer et kort øyeblikk. For $t > 1$ beskriver i strømmen i fuglens kropp mens den daler ned mot bakken. Se figuren under.

I denne oppgaven kunne vi også brukt delbrøksoppspaltning:

$$\frac{s}{s^2 + 2s + 2} = \frac{s}{(s - 1 - i)(s - 1 + i)}$$

men regningen blir noe svineri. \triangle

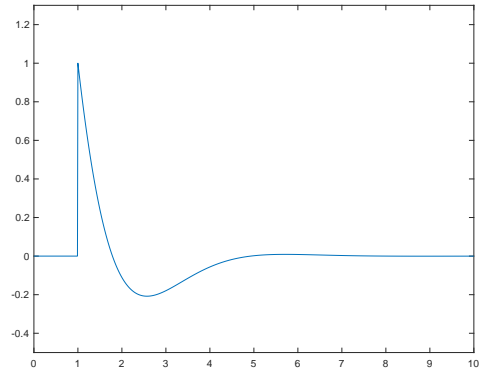
Nå tar vi to eksempler der vi bruker konvolusjon.

Eksempel 1.33. Vi løser initialverdiproblemet

$$y + t * y' = t \quad y(0) = 0$$

Vi laplacetransformerer

$$\mathcal{L}(y) + \mathcal{L}(t * y') = \frac{1}{s^2}$$



og bruker konvolusjonsteoremet

$$\mathcal{L}(y) + \mathcal{L}(t)\mathcal{L}(y') = \mathcal{L}(y) + \frac{1}{s}\mathcal{L}(y) = \frac{1}{s^2}$$

eller

$$s\mathcal{L}(y) + \mathcal{L}(y) = \frac{1}{s}$$

Herfra er regningen standard, og vi får

$$\mathcal{L}(y) = \frac{1}{s(s + 1)} = \frac{1}{s} - \frac{1}{s + 1},$$

slik at

$$y = 1 - e^{-t}. \quad \triangle$$

Eksempel 1.34. La oss løse initialverdiproblemet

$$y'' + y = \sin t \quad y(0) = y'(0) = 0$$

Vi laplacetransformerer alt, og får

$$(s^2 + 1)\mathcal{L}(y) = \frac{1}{s^2 + 1}$$

eller

$$\mathcal{L}(y) = \frac{1}{(s^2 + 1)^2} = (\mathcal{L}(\sin t))^2$$

som gir at

$$y(t) = \int_0^t \sin(t-u) \sin u \, du$$

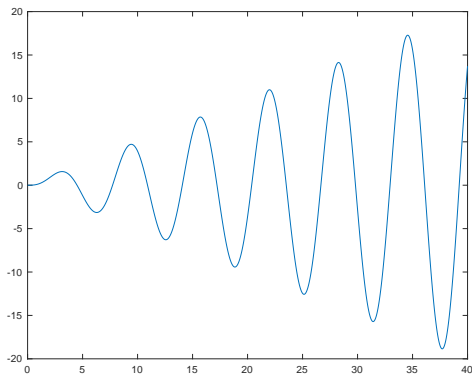
Dette siste integralet kan vi løse ved å bruke relasjonen

$$\sin u = \frac{e^{iu} - e^{-iu}}{2i}$$

og integrere i vei

$$\begin{aligned} y(t) &= \int_0^t \sin(t-u) \sin u \, du \\ &= -\frac{1}{4} \int_0^t (e^{i(t-u)} - e^{-i(t-u)})(e^{iu} - e^{-iu}) \, du \\ &= -\frac{1}{4} \int_0^t e^{it} - e^{i(t-2u)} - e^{-i(t-2u)} + e^{-it} \, du \\ &= -\frac{1}{2} \int_0^t \cos t - \cos(t-2u) \, du \\ &= \frac{1}{2} (\sin t - t \cos t) \end{aligned}$$

Dette eksemplet beskriver for eksempel en kloss og en fjær, der den påtrykte kraften $\sin t$ har samme frekvens som systemets naturlige svingefrekvens. Da oppstår det resonans. Resonansen kommer til uttrykk gjennom leddet $t \sin t$, som vokser mot uendelig når $t \rightarrow \infty$. Når dette skjer i et PA-anlegg, kalles det feedback, og alle må holde seg for ørene. \triangle



Eksemplet over viser at selv om det er litt regning med laplacetransform, er det en penere løsningsmetode enn å dele opp i homogen og partikulær løsning, og så gjette i vei. Eksemplet illustrerer også en teknikk vi får bruk for når vi skal løse varmelikningen på hele x -aksen i kapittel 5.

2 Fourierrekker

I M1 lærte du at mange glatte funksjoner kan skrives som en potensrekke. En mye større klasse av funksjoner kan skrives som rekker av sinus- og cosinusfunksjoner.

Bevisene i dette kapitlet er ganske tekniske. Ikke les dem før du har regnet en haug med gamle eksamensoppgaver. Begynn med å forstå hvordan du setter opp en funksjons fourierrekke.

Komplekse funksjoner av en reell variabel

I M1 og M2 var alle funksjoner reelle. I M3 begynte vi så smått med komplekse funksjoner av en reell variabel når vi løste systemer av differensiallikninger. I dette kapitlet skal vi fortsette med det.

En *kompleks funksjon av en reell variabel* er en funksjon

$$f: \mathbb{R} \rightarrow \mathbb{C}$$

Man kan tenke på f som en parametrisering av en kurve i det komplekse planet.

Eksempel 2.1. Funksjonen

$$f(x) = (1 + i)x \quad t \geq 0$$

beskriver en rett linje i første kvadrant i det komplekse planet. \triangle

Eksempel 2.2. Fra M3 husker du kanskje Eulers formel

$$e^{ix} = \cos x + i \sin x.$$

Denne funksjonen kalles en *kompleks eksponensialfunksjon*. Merk at funksjonsverdiene ligger på enhets-sirkelen i det komplekse planet, og funksjonen kan betraktes som en parametrisering av denne. \triangle

I mange tilfeller er det gunstig å skrive en reell funksjon som en kompleks funksjon.

Eksempel 2.3. Dersom vi erstatter x med $-x$ i Eulers formel, får vi

$$e^{-ix} = \cos x - i \sin x.$$

Disse to likningene kan kombineres til

$$\cos x = \frac{e^{ix} + e^{-ix}}{2}$$

og

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i}.$$

Med andre ord: om vi har en funksjon som involverer sinus- og cosinusfunksjoner, kan den like gjerne skrives med komplekse eksponensialfunksjoner. \triangle

Sinus- og cosinusfunksjoner har du jobbet med siden gymnaset, og derfor er de intuitive å jobbe med, og dessuten er de lette å visualisere. Komplekse eksponensialfunksjoner er vanskeligere å se for seg, men ofte mer praktiske i bruk. Noen beviser blir veldig hårete på reell form, og fouriertransform, som kommer i neste kapittel, er praktisk talt umulig å forstå uten komplekse eksponensialfunksjoner.

Periodiske funksjoner

Når man skal forstå fourierrekker er det ikke mulig å komme utenom periodiske funksjoner. En funksjon sies å ha *periode* $p > 0$ dersom

$$f(x + p) = f(x) \quad (2.1)$$

for alle x i definisjonsmengden til f . Den minste p slik at (2.1) holder, kalles *fundamentalperioden* til f .

Eksempel 2.4. La

$$f(x) = \sin x$$

Denne funksjonen har perioder $2n\pi$ for alle $n \in \mathbb{N}$. Fundamentalperioden er 2π . \triangle

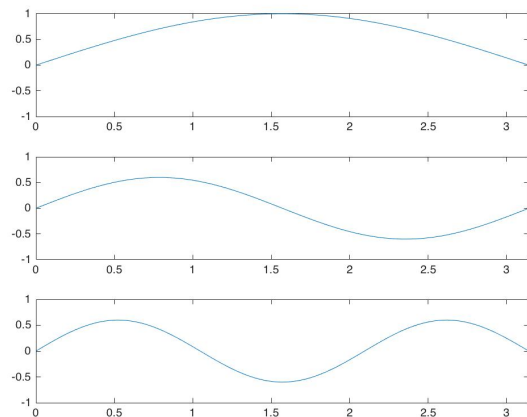
Dersom f har periode p , og $g(x) = f(kx)$, har g periode p/k , for

$$\begin{aligned} g(x) &= f(kx) = f(kx + p) \\ &= f(k(x + p/k)) = g(x + p/k). \end{aligned}$$

Eksempel 2.5. La

$$f(x) = \sin(3x)$$

Denne funksjonen har perioder $2n\pi/3$ for alle $n \in \mathbb{N}$. Fundamentalperioden er $2\pi/3$. Under er plot av funksjonene $\sin x$, $\sin 2x$ og $\sin 3x$. \triangle



Eksempel 2.6. Den komplekse eksponensialfunksjonen $e^{ix} = \cos x + i \sin x$ også åpenbart fundamentalperiode 2π . \triangle

Ortogonale systemer

La f og g være komplekse funksjoner av en reell variabel på intervallet $[a, b]$. Integralet

$$\int_a^b f \bar{g} dx$$

er en generalisering av indreproduktet mellom vektorer, og har alle de samme egenskapene. Vi sier at f og g er *ortogonale* på $[a, b]$ dersom

$$\int_a^b f \bar{g} dx = 0.$$

Et *ortogonalt system* $\{\phi_n\}$ er familie av funksjoner ϕ_n som er innbyrdes ortogonale. Merk at siden $f\bar{f} = |f|^2 \geq 0$, er

$$\int_a^b f\bar{f} dx \geq 0,$$

og den eneste funksjonen som tilfredsstiller

$$\int_a^b f\bar{f} dx = 0$$

er $f = 0$. På samme vis som $\int_a^b f\bar{g} dx$ er en generalisering av indreproduktet mellom vektorer, er

$$\|f\| = \sqrt{\int_a^b f\bar{f} dx}$$

en generalisering av lengden til en vektor.

Dersom en familie $\{\phi_n\}$ av funksjoner er innbyrdes ortogonale, sier vi at $\{\phi_n\}$ er et *ortogonalt system*, og dersom de tilfredsstiller

$$\int_a^b \phi_n \overline{\phi_m} dx = \begin{cases} 1 & \text{for } n = m \\ 0 & \text{for } n \neq m \end{cases}$$

sier vi at familien er *ortonormal*.

Eksempel 2.7. De komplekse eksponensialfunksjonene

$$e^{inx} \quad n \in \mathbb{Z}$$

utgjør et ortogonalt system på intervallet $[-\pi, \pi]$:

$$\int_{-\pi}^{\pi} e^{inx} \overline{e^{imx}} dx = \int_{-\pi}^{\pi} e^{inx} e^{-imx} dx = \int_{-\pi}^{\pi} e^{i(n-m)x} dx = \begin{cases} 2\pi & \text{for } n = m \\ 0 & \text{for } n \neq m \end{cases} \quad \Delta$$

Eksempel 2.8. Funksjonene

$$\frac{1}{2\pi} e^{inx} \quad n \in \mathbb{Z}$$

er et ortonormalt system på intervallet $[-\pi, \pi]$. Δ

Dersom f og g er reelle, blir $f = \bar{f}$ og $g = \bar{g}$, slik at indreproduktet kan skrives

$$\int_a^b f g dx.$$

Reelle funksjoner kan også danne ortogonale systemer.

Eksempel 2.9. La $m, n \geq 1$. Siden

$$\int_{-\pi}^{\pi} \cos nx \cos mx dx = \begin{cases} \pi & \text{for } n = m \\ 0 & \text{for } n \neq m \end{cases}$$

$$\int_{-\pi}^{\pi} \sin nx \sin mx dx = \begin{cases} \pi & \text{for } n = m \\ 0 & \text{for } n \neq m \end{cases}$$

og

$$\int_{-\pi}^{\pi} \cos nx \sin mx dx = 0$$

utgjør $\{\sin nx, \cos mx\}$ også et ortogonalt system på $[-\pi, \pi]$. Det er ikke så vanskelig å se at disse er ortogonale. Vi beviser det for cosinusfunksjonene. Først skriver vi om litt:

$$\int_{-\pi}^{\pi} \cos nx \cos mx dx = \frac{1}{2} \int_{-\pi}^{\pi} \cos(n+m)x + \cos(n-m)x dx$$

Det siste integralet er lett å beregne. Det forsvinner for alle verdier av m og n , unntatt når $m = n$, for da er

$$\frac{1}{2} \int_{-\pi}^{\pi} \cos(n-m)x dx = \int_{-\pi}^{\pi} dx = \pi,$$

slik at

$$\int_{-\pi}^{\pi} \cos nx \cos mx dx = \begin{cases} \pi & \text{for } n = m \\ 0 & \text{for } n \neq m \end{cases}$$

De to andre formlene bevises på samme måte. Δ

Eksempel 2.10. Legendrepolynomene er en familie av polynomer gitt ved rekursjonen

$$P_0 = 1$$

$$P_1 = x$$

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x).$$

og er et ortogonalt system på intervallet $[-1, 1]$. De første fem polynomene er:

$$P_0 = 1$$

$$P_1 = x$$

$$P_2 = \frac{1}{2}(3x^2 - 1)$$

$$P_3 = \frac{1}{2}(5x^3 - 3x)$$

$$P_4 = \frac{1}{8}(35x^4 - 30x^2 + 3)$$

Vi skal ikke ha bruk for legendrepolynomene i dette kapitlet, men vi skal møte dem igjen i kapitlene om interpolasjon og numerisk integrasjon. Legendrepolynomene dukker opp i mange sammenhenger, både i matematikken og i andre fagfelt. Det kanskje mest kjente eksemplet utenfor matematikken, er Schrödingers likning for hydrogenatomet, der disse dukker opp i familien av løsninger. Δ

Dersom en funksjon kan skrives som en lineærkombinasjon av funksjoner i et ortogonalt system, er det lett å utlede formler for koeffisientene i lineærkombinasjonen.

Teorem 2.11. Dersom $\{\phi_n\}$ er et ortonormalt system på $[a, b]$, og

$$f(x) = \sum_{n=1}^N c_n \phi_n(x),$$

er

$$c_n = \int_a^b f(x) \overline{\phi_n(x)} dx$$

Bevis. Vi ganger f med $\overline{\phi_m}$:

$$f(x)\overline{\phi_m} = \sum_{n=1}^N c_n \phi_n \overline{\phi_m},$$

integrerer

$$\int_a^b f(x)\overline{\phi_m} dx = \sum_{n=1}^N c_n \int_a^b \phi_n \overline{\phi_m} dx,$$

og bruker ortogonaliteten:

$$\int_a^b f(x)e^{-imx} dx = c_m. \quad \square$$

Koeffisientene c_n kalles *fourierkoeffisientene*. De har følgende artige egenskap dersom vi begrenser oss til funksjoner som tilfredsstillers $\int_a^b |f(x)|^2 dx < \infty$.

Teorem 2.12. Dersom

$$f(x) = \sum_{n=1}^N c_n \phi_n(x),$$

er

$$\sum_{n=1}^N |c_n|^2 = \int_a^b |f(x)|^2 dx$$

Bevis. Vi ganger nå f med \overline{f} :

$$f\overline{f} = \sum_{n=1}^N \sum_{m=1}^N c_n \overline{c_m} \phi_n \overline{\phi_m},$$

integrerer

$$\int_a^b f\overline{f} dx = \sum_{n=1}^N \sum_{m=1}^N c_n \overline{c_m} \int_a^b \phi_n \overline{\phi_m} dx,$$

og bruker igjen ortogonaliteten:

$$\int_a^b f(x)\overline{f(x)} dx = \sum_{n=1}^N c_n \overline{c_n}. \quad \square$$

Dersom f ikke kan skrives som en lineærkombinasjon av funksjonene ϕ_n , kan vi projisere f ned i rommet spent ut disse funksjonene. Lineærkombinasjonen med fourierkoeffisientene som vektor er den beste approksimasjonen til f i rommet utspent av ϕ_n , så lenge vi måler kvaliteten med funksjonsnormen definert ovenfor.

Teorem 2.13. Dersom

$$h(x) = \sum_{n=1}^N c_n \phi_n(x),$$

med

$$c_n = \int_a^b f(x)\overline{\phi_n(x)} dx$$

og

$$g(x) = \sum_{n=1}^N d_n \phi_n(x),$$

er

$$\int_a^b |f(x) - h(x)|^2 dx \leq \int_a^b |f(x) - g(x)|^2 dx$$

Bevis. Denne beregningen er litt hårete, men du finner nok ut av det med penn og papir:

$$\begin{aligned} \int_a^b |f(x) - g(x)|^2 dx &= \int_a^b |f(x)|^2 dx - \int_a^b f(x)\overline{g(x)} dx \\ &\quad - \int_a^b \overline{f(x)}g(x) dx + \int_a^b |g(x)|^2 dx = \\ \int_a^b |f(x)|^2 dx - \sum_{n=1}^N c_n \overline{d_n} - \sum_{n=1}^N \overline{c_n} d_n + \sum_{n=1}^N d_n \overline{d_n} &= \\ \int_a^b |f(x)|^2 dx - \sum_{n=1}^N c_n \overline{c_n} + \sum_{n=1}^N |c_n - d_n|^2. \end{aligned}$$

Det siste uttrykket er helt klart minimert dersom man velger $d_n = c_n$. \square

Det neste teoremet kalles gjerne *Bessels ulikhet*, ihvertfall dersom $n \rightarrow \infty$.

Teorem 2.14. Dersom

$$h(x) = \sum_{n=1}^N c_n \phi_n(x),$$

med

$$c_n = \int_a^b f(x)\overline{\phi_n(x)} dx$$

er

$$\sum_{n=1}^N |c_n|^2 \leq \int_a^b |f(x)|^2 dx$$

Bevis. Fra forrige bevis vet vi at

$$\int_a^b |f(x) - h(x)|^2 dx = \int_a^b |f(x)|^2 dx - \sum_{n=1}^N c_n \overline{c_n}.$$

Siden

$$\int_a^b |f(x) - h(x)|^2 dx \geq 0,$$

må

$$\int_a^b |f(x)|^2 dx \geq \sum_{n=1}^N c_n \overline{c_n}. \quad \square$$

Merk til slutt at dersom vi har uendelig mange funksjoner ϕ_n , impliserer Bessels ulikhet at

$$\sum_{n=1}^{\infty} |c_n|^2$$

er en konvergent rekke, og at $\lim_{n \rightarrow \infty} |c_n| = 0$.

Fourierrekker

En *generell fourierrekke* er et uttrykk på formen

$$\sum_{n=1}^{\infty} c_n \phi_n(x).$$

En *trigonometrisk rekke* er en fourierrekke av trigonometriske funksjoner

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}.$$

Merk at dette er en geometrisk rekke med multiplikasjonsfaktor e^{ix} . Fra nå av skal alt handle om trigonometriske rekker.

Definisjon. La f være en funksjon. Vi definerer f sin *fourierrekke* som

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}$$

der

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx,$$

og skriver

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}.$$

Vi skal nå stille oss spørsmålet for hvilke f det er riktig å skrive

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}.$$

Ved første øyekast kan man bli forledet til å tro at kun harmoniske svingebevegelser kan skrives på denne måten, og det trodde de fleste matematikere helt frem til 1800-tallet en gang. Det viser seg imidlertid at nesten alle funksjoner mellom himmel og jord kan skrives som trigonometriske rekker, bare man løser litt på konvergensbegrepet. Det har ikke vi mulighet til å gjøre ordentlig, men vi skal bevise et klassisk konvergensteorem. Først må vi bare ha litt mer maskineri.

Uttrykket

$$S_N = \sum_{n=-N}^N c_n e^{inx}$$

kalles *den N-te partialsummen* til fourierrekken

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}.$$

Uttrykket

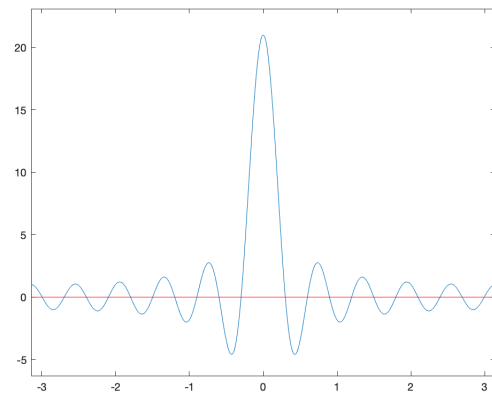
$$D_N(x) = \sum_{n=-N}^N e^{inx}$$

kalles *dirichletkjernen*. Den er en reell funksjon, for vi kan skrive

$$D_N = \sum_{n=-N}^N e^{inx} = 1 + 2 \sum_{n=1}^N \cos nx.$$

Under er et plot av dirichletkjernen for $N = 10$.

Vi kan også lage en enda penere formel. Først utnytter vi at en trigonometrisk rekke er en geometrisk



rekke med faktor e^{ix} , og skriver

$$e^{iNx} \sum_{n=-N}^N e^{inx} = \sum_{n=0}^{2N} e^{inx} = \frac{1 - e^{i(2N+1)x}}{1 - e^{ix}}.$$

Siden

$$\frac{e^{-i(N+1/2)x}}{e^{-ix/2}} e^{iNx} = 1$$

kan vi beregne

$$\begin{aligned} \sum_{n=-N}^N e^{inx} &= \frac{e^{-i(N+1/2)x}}{e^{-ix/2}} e^{iNx} \sum_{n=-N}^N e^{inx} \\ &= \frac{e^{-i(N+1/2)x} + e^{i(N+1/2)x}}{e^{-ix/2} - e^{ix/2}} \\ &= \frac{\sin(N + \frac{1}{2})x}{\sin \frac{1}{2}x}. \end{aligned}$$

for $x \neq 0$. Siden $D_N(0) = 2N + 1$, og

$$\lim_{x \rightarrow 0} \frac{\sin(N + \frac{1}{2})x}{\sin \frac{1}{2}x} = \frac{N + \frac{1}{2}}{\frac{1}{2}} = 2N + 1,$$

kan vi skrive

$$D_N(x) = \begin{cases} \frac{\sin(N + \frac{1}{2})x}{\sin \frac{1}{2}x} & x \neq 0 \\ 2N + 1 & x = 0 \end{cases}$$

Vi kan bruke dirichletkjernen til å skrive partialsummen til en fourierrekke som en konvolusjon:

$$\begin{aligned} 2\pi \sum_{n=-N}^N c_n e^{inx} &= \sum_{n=-N}^N \int_{-\pi}^{\pi} f(y) e^{-iny} dy e^{inx} \\ &= \int_{-\pi}^{\pi} f(y) \sum_{n=-N}^N e^{in(x-y)} dy \\ &= \int_{-\pi}^{\pi} f(y) D_N(x-y) dy = f * D_N. \end{aligned}$$

Det er ikke så vanskelig demonstrere punktvis konvergens for funksjoner som er litt mer enn bare kontinuerlige.

Teorem 2.15. La f være en 2π -periodisk funksjon, og la x være et punkt slik at

$$|f(x-t) - f(x)| \leq M|t|$$

for alle $t \in (-\pi, \pi)$. Da er

$$\lim_{N \rightarrow \infty} \sum_{n=-N}^N c_n e^{inx} = f(x).$$

Bevis. Husk dirichletkjernen

$$D_N(x) = \sum_{n=-N}^N e^{inx}.$$

For det første er det lett å se at

$$\int_{-\pi}^{\pi} D_N(t) dt = \sum_{n=-N}^N \int_{-\pi}^{\pi} e^{inx} dt = 2\pi,$$

og siden D_N er en 2π -periodisk funksjon, kan vi like gjerne skrive

$$\int_{-\pi}^{\pi} D_N(t) dt = \int_{-\pi}^{\pi} D_N(t-x) dt = 2\pi.$$

På grunn av denne likningen, kan vi, siden både f og D_N er 2π -periodiske, skrive

$$f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) D_N(t) dt.$$

Siden både f og D_N er 2π -periodiske, kan vi skrive

$$\begin{aligned} \sum_{n=-N}^N c_n e^{inx} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) D_N(x-t) dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t-x) D_N(t) dt. \end{aligned}$$

Alt dette kan vi slå sammen til

$$f(x) - \sum_{n=-N}^N c_n e^{inx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} (f(t-x) - f(x)) D_N(t) dt.$$

Vi tar en titt på funksjonen

$$g(t) = \frac{f(t-x) - f(x)}{\sin t/2}$$

Siden $|f(t-x) - f(x)| \leq M|t|$, er denne funksjonen er begrenset på $[-\pi, \pi]$. Den er ikke definert for $t = 0$, men det gjør ikke noe, for vi skal bare integrere den. Nå skriver vi

$$\begin{aligned} f(x) - \sum_{n=-N}^N c_n e^{inx} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} g(t) \sin(N+1/2)t dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} g(t) \cos t/2 \sin Nt dt \\ &\quad + \frac{1}{2\pi} \int_{-\pi}^{\pi} g(t) \sin t/2 \cos Nt dt. \end{aligned}$$

Bruker vi Eulers formel på $\sin Nt$ og $\cos Nt$ ser vi at de to siste integralene er satt sammen av fourierkoeffisientene til funksjonene $g(t) \sin t/2$ og $g(t) \cos t/2$ og de komplekskonjugerte av disse. Fra Bessels ulikhet vet vi at alle disse fire må gå mot null når $N \rightarrow \infty$, og med andre ord må både

$$\lim_{n \rightarrow \infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} g(t) \cos t/2 \sin Nt dt = 0$$

og

$$\lim_{n \rightarrow \infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} g(t) \sin t/2 \cos Nt dt = 0$$

slik at

$$\lim_{N \rightarrow \infty} \sum_{n=-N}^N c_n e^{inx} = f(x). \quad \square$$

Kommentar. Det finnes noen berømte eksempler på kontinuerlige funksjoner der fourierrekken ikke konvergerer i alle punkter. Så at f er kontinuerlig, er ikke nok. Betingelsen

$$|f(x-t) - f(x)| \leq M|t|$$

for alle $t \in (-\pi, \pi)$, kalles gjerne *Lipschitz-kontinuitet*, og er litt strengere enn vanlig kontinuitet.

Det går an å vise et litt mer generelt teorem, men det skal ikke vi gjøre, vi bare skriver det opp.

Teorem 2.16. La f være en stykkvis kontinuerlig deriverbar funksjon på $[-\pi, \pi]$, der den venstre- og høyrederiverte eksisterer i eventuelle bruddpunkter, og la

$$f \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}.$$

Hvis f er kontinuerlig i punktet x , konvergerer fourierrekken til $f(x)$ i x . Dersom f ikke er kontinuerlig i x , konvergerer fourierrekken til

$$\lim_{h \rightarrow 0^+} \frac{f(x-h) + f(x+h)}{2}.$$

I endepunktene konvergerer fourierrekken til

$$\lim_{h \rightarrow 0^+} \frac{f(\pi-h) + f(-\pi+h)}{2}.$$

Dersom f er kontinuerlig deriverbar, er

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

på $(-\pi, \pi)$, og dersom f er kontinuerlig deriverbar med periode 2π , er

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

på $[-\pi, \pi]$.

En haug med eksempler

I dette avsnittet skal jeg beregne en helt syk mengde eksempler for dere.

Eksempel 2.17. Vi finner fourierrekken til

$$f(x) = x \quad \text{der} \quad x \in (-\pi, \pi).$$

Vi beregner

$$c_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} x \, dx = 0,$$

og

$$\begin{aligned} c_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} x e^{-inx} \, dx = -\frac{1}{2in} (e^{-in\pi} + e^{in\pi}) \\ &= -\frac{1}{in} \cos n\pi = \frac{(-1)^{n+1}}{in}. \end{aligned}$$

Siden x er glatt på intervallet $(-\pi, \pi)$, blir fourierrekken

$$\begin{aligned} x &= \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{(-1)^{n+1}}{in} e^{inx} \\ &= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{in} e^{inx} - \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{in} e^{-inx} \\ &= 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin nx. \quad \triangle \end{aligned}$$

Fourierrekker på reell form

Dersom f er reell, blir fourierrekken også reell, som illustrert i forrige eksempel. Vi kan fint sette opp fourierrekken direkte med sinus- og cosinusfunksjoner:

$$f \sim a_0 + \sum_{n=1}^{\infty} a_n \cos nx + b_n \sin nx,$$

for disse funksjonene utgjør også et ortogonalt system. koeffisientene blir

$$\begin{aligned} a_0 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \, dx & a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx. \end{aligned}$$

Eksempel 2.18. Vi finner fourierrekken til

$$f(x) = x \quad \text{der} \quad x \in (-\pi, \pi).$$

en gang til. Vi beregner

$$\begin{aligned} a_0 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} x \, dx = 0 & a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} x \cos nx \, dx = 0 \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} x \sin nx \, dx = \frac{2}{n} (-1)^{n+1} \end{aligned}$$

Merk at symmetri gir a_0 og a_n , mens b_n må beregnes. Fourierrekken blir som kjent

$$x = 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin nx. \quad \triangle$$

Eksempel 2.19. Vi finner heavisidefunksjonen

$$u(x) = \begin{cases} 1 & \text{for } x \geq 0 \\ 0 & \text{for } x < 0. \end{cases}$$

sin reelle fourierrekke på $(-\pi, \pi)$. Vi beregner

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} u(x) \, dx = \frac{1}{2\pi} \int_0^{\pi} dx = \frac{1}{2}$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} u(x) \cos nx \, dx = \frac{1}{\pi} \int_0^{\pi} \cos nx \, dx = 0$$

$$\begin{aligned} b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} u(x) \sin nx \, dx = \frac{1}{\pi} \int_0^{\pi} \sin nx \, dx \\ &= \begin{cases} \frac{2}{n\pi} & \text{for } n \text{ oddetall} \\ 0 & \text{for } n \text{ partall.} \end{cases} \end{aligned}$$

Altså kan vi skrive

$$u(x) \sim \frac{1}{2} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1}{2n-1} \sin(2n-1)x$$

Fourierrekken konvergerer til u på intervallene $(-\pi, 0)$ og $(0, \pi)$, og til $1/2$ i $x = 0$ og $x = \pm\pi$. Partialsummer er funksjoner på formen

$$S_N = \frac{1}{2} + \frac{2}{\pi} \sum_{n=1}^N \frac{1}{2n-1} \sin(2n-1)x$$

Under er plot av partialsummer for $n = 2$, $n = 5$ og $n = 100$. Jeg har plottet på intervallet $[-2\pi, 2\pi]$ for å illustrere hvordan fourierrekken oppfører seg utenfor intervallet $[-\pi, \pi]$. \triangle

Formler for overgang mellom kompleks og reell fourierrekke

Vi har sett i eksempler at dersom f er en reell funksjon, blir fourierrekken også reell, til tross for at de trigonometriske eksponensialene er komplekse funksjoner. Dette følger av formelen for koeffisientene c_n . Dersom f er reell, kan vi skrive

$$\begin{aligned} c_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} \, dx \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \overline{e^{inx}} \, dx \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{inx} \, dx = \overline{c_{-n}} \end{aligned}$$

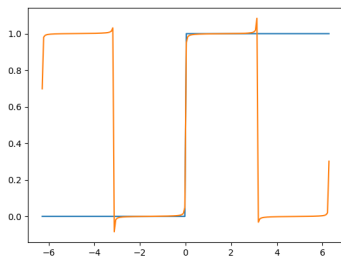
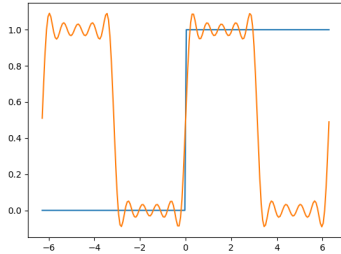
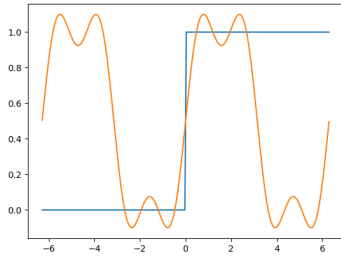
Dette betyr at

$$c_n e^{inx} + c_{-n} e^{-inx} = c_n e^{inx} + \overline{c_n e^{inx}}$$

blir en reell funksjon, og følgelig er hele fourierrekken reell.

Vi kan utlede formel for overgangen mellom fourierrekker på reell og kompleks form:

$$\begin{aligned} c_n + c_{-n} &= \frac{1}{2\pi} \left(\int_{-\pi}^{\pi} f(x) e^{-inx} \, dx + \int_{-\pi}^{\pi} f(x) e^{inx} \, dx \right) \\ &= \frac{1}{2\pi} \left(\int_{-\pi}^{\pi} f(x) (e^{-inx} + e^{inx}) \, dx \right) \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx = a_n \end{aligned}$$



$$\begin{aligned}
 c_n - c_{-n} &= \frac{1}{2\pi} \left(\int_{-\pi}^{\pi} f(x) e^{-inx} dx - \int_{-\pi}^{\pi} f(x) e^{inx} dx \right) \\
 &= \frac{1}{2\pi} \left(\int_{-\pi}^{\pi} f(x) (e^{-inx} - e^{inx}) dx \right) \\
 &= \frac{-i}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx = -ib_n
 \end{aligned}$$

Vi kan selvfølgelig også snu om på disse formlene, og skrive

$$c_n = \frac{1}{2}(a_n - ib_n)$$

$$c_{-n} = \frac{1}{2}(a_n + ib_n).$$

Fourierrekker på andre intervaller

Å utlede formler for fourierrekker på andre intervaller enn $[-\pi, \pi]$ er ikke vanskelig. For intervallet $[-L, L]$ skriver man

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{i \frac{n\pi x}{L}}. \quad (2.2)$$

eller

$$f \sim a_0 + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{L} + b_n \sin \frac{n\pi x}{L}.$$

Koeffisientene er gitt ved

$$c_n = \frac{1}{2L} \int_{-L}^L f(x) e^{-\frac{n\pi x}{L}} dx,$$

og

$$a_0 = \frac{1}{2L} \int_{-L}^L f(x) dx$$

$$a_n = \frac{1}{L} \int_{-L}^L f(x) \cos \frac{n\pi x}{L} dx$$

$$b_n = \frac{1}{L} \int_{-L}^L f(x) \sin \frac{n\pi x}{L} dx$$

henholdsvis. Utledning er identisk med utledning på intervallet $[-\pi, \pi]$. På samme måte kan man sette opp fourierrekker på intervallet $[a, b]$, men det dropper vi.

Eksempel 2.20. Vi beregner den komplekse fourierrekken til heavisidefunksjonen på $[-1, 1]$. Koeffisientene blir

$$\begin{aligned}
 c_n &= \frac{1}{2} \int_{-1}^1 f(x) e^{-i \frac{n\pi x}{L}} dx \\
 &= \frac{1}{2} \int_0^1 e^{-i \frac{n\pi x}{L}} dx = \begin{cases} \frac{1}{2} & \text{for } n = 0 \\ \frac{1}{n\pi i} & \text{for odde } n \\ 0 & \text{for jevne } n, \end{cases}
 \end{aligned}$$

slik at

$$u(x) \sim \frac{1}{2} + \frac{1}{\pi i} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{1}{(2n-1)} e^{(2n-1)\pi i x}.$$

Merk nok en gang at

$$c_n e^{n\pi i x} + c_{-n} e^{-n\pi i x} = \frac{2}{n\pi} \sin n\pi x,$$

slik at

$$u(x) \sim \frac{1}{2} + \frac{1}{\pi i} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{1}{(2n-1)} e^{(2n-1)\pi i x}$$

$$= \frac{1}{2} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1}{2n-1} \sin(2n-1)\pi x. \quad \triangle$$

Odde og jevne funksjoner

Vi sier at en funksjon er odde dersom

$$f(-x) = -f(x)$$

og jevn dersom

$$f(-x) = f(x).$$

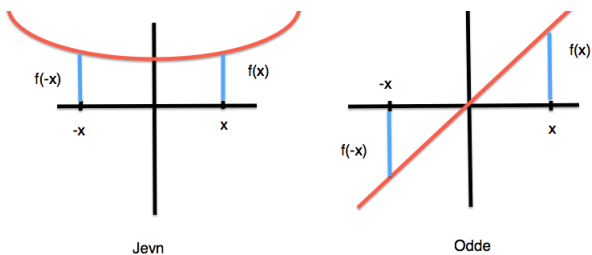
for alle x i definisjonsmengden til f . Grafen til en odde funksjon blir identisk dersom du dreier den π radianer om origo, mens grafen til en jevn funksjon blir identisk dersom du speiler den om y -aksen. En rask kikk på figur overbeviser oss om at

$$\int_{-L}^L f(x) dx = 0$$

for odde funksjoner, og

$$\int_{-L}^L f(x) dx = 2 \int_0^L f(x) dx$$

for jevne funksjoner.



Eksempel 2.21. Man kan vise at produktet av to jevne eller to odde funksjoner blir en jevn funksjon, og at produktet av en jevn og en odde funksjon blir en odde funksjon. La f være odde og g være jevn. Siden $f(-x) = -f(x)$, og $g(-x) = g(x)$, får vi

$$f(-x)g(-x) = -f(x)g(x),$$

altså er fg en odde funksjon. De andre tilfellene bevises på samme måte. \triangle

Eksempel 2.22. Merk at $u - \frac{1}{2}$ er en odde funksjon. Denne strukturen kommer til syne i fourierrekken til u . \triangle

Merk at dersom f er odde, er $a_n = 0$ for alle n , mens dersom f er jevn, er $b_n = 0$ for alle n . Fourierrekken til en odde funksjon inneholder følgelig kun sinusfunksjoner, mens fourierrekken til jevne funksjoner inneholder kun cosinusfunksjoner.

Odde og jevne utvidelser

Dersom en funksjon f har definisjonsmengde $(0, L)$, kan vi definere den odde utvidelsen

$$f_o = \begin{cases} f(x) & \text{for } x = (0, L) \\ -f(-x) & \text{for } x = (-L, 0) \end{cases}$$

og den jevne utvidelsen

$$f_j = \begin{cases} f(x) & \text{for } x = (0, L) \\ f(-x) & \text{for } x = (-L, 0). \end{cases}$$

Siden både f_o og f_j er identiske med f på $(0, L)$, vil fourierrekken deres konvergere til $f(x)$ på $(0, L)$. Man kan således velge mellom sinus eller cosinus når man skal fourierutvikle f på $(0, L)$. Disse kalles henholdsvis sinus- og cosinusrekken til f på $(0, L)$.

Vi ser på fourierutviklingen til f_o . For den har vi at $a_n = 0$ for alle n , og at

$$\begin{aligned} b_n &= \frac{1}{L} \int_{-L}^L f_o(x) \sin \frac{n\pi x}{L} dx \\ &= \frac{2}{L} \int_0^L f(x) \sin \frac{n\pi x}{L} dx. \end{aligned}$$

For fourierutviklingen til f_j , har vi tilsvarende at

$$\begin{aligned} a_n &= \frac{1}{L} \int_{-L}^L f_j(x) \cos \frac{n\pi x}{L} dx \\ &= \frac{2}{L} \int_0^L f(x) \cos \frac{n\pi x}{L} dx \end{aligned}$$

samt

$$a_0 = \frac{1}{2L} \int_{-L}^L f_j(x) dx = \frac{1}{L} \int_0^L f(x) dx.$$

Eksempel 2.23. Vi beregner cosinusrekken til $f(x) = x$ på $(0, \pi)$. Koeffisientene blir

$$a_0 = \frac{1}{\pi} \int_0^\pi x dx = \pi/2$$

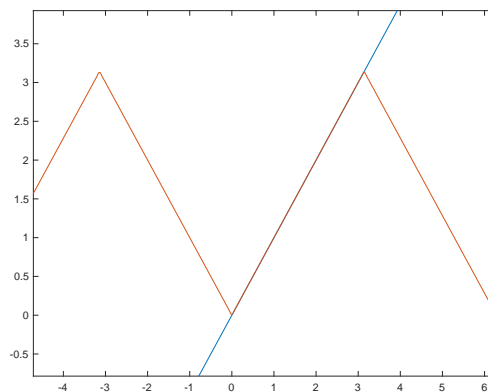
og

$$a_n = \frac{2}{\pi} \int_0^\pi x \cos nx dx = \begin{cases} \frac{-4}{\pi n^2} & \text{for odde } n \\ 0 & \text{for for jevne } n. \end{cases}$$

slik at

$$x = \frac{\pi}{2} - \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} \cos(2n-1)x$$

på intervallet $(0, \pi)$. Under er et plot af $f(x) = x$ og cosinusrekken på et litt større intervall enn $[0, \pi]$. \triangle



Parsevals identitet

Teorem 2.24. Anta integralet

$$\int_{-L}^L f^2(x) dx$$

eksisterer, og at

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{i \frac{n\pi x}{L}}.$$

Da gjelder Parsevals identitet:

$$\begin{aligned} \frac{1}{2L} \int_{-L}^L f^2(x) dx &= \sum_{n=-\infty}^{\infty} |c_n|^2 \\ &= a_0^2 + \frac{1}{2} \sum_{n=1}^{\infty} a_n^2 + b_n^2 \end{aligned}$$

Bevis. Denne er lettest å bevise om man starter på kompleks form

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{i \frac{n\pi x}{L}}.$$

Vi ganger denne likheten med seg selv

$$\begin{aligned} f^2(x) &= \sum_{n=-\infty}^{\infty} c_n e^{i \frac{n\pi x}{L}} \sum_{m=-\infty}^{\infty} c_m e^{i \frac{m\pi x}{L}} \\ &= \sum_{n=-\infty}^{\infty} c_n e^{i \frac{n\pi x}{L}} \sum_{m=-\infty}^{\infty} c_{-m} e^{i \frac{-m\pi x}{L}} \\ &= \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} c_n c_{-m} e^{i \frac{(n-m)\pi x}{L}} \end{aligned}$$

og integrerer

$$\begin{aligned} \int_{-L}^L f^2(x) dx &= \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} c_n c_{-m} \int_{-L}^L e^{i \frac{(n-m)\pi x}{L}} dx \\ &= 2L \sum_{n=-\infty}^{\infty} c_n c_{-n} \\ &= 2L \sum_{n=-\infty}^{\infty} c_n \bar{c}_n = 2L \sum_{n=-\infty}^{\infty} |c_n|^2. \end{aligned}$$

Likningen

$$\frac{1}{2L} \int_{-L}^L f^2(x) dx = a_0^2 + \frac{1}{2} \sum_{n=1}^{\infty} a_n^2 + b_n^2$$

vises på samme måte. \square

Noen artige anvendelser

Her er en samling eksempler som ofte dukker opp på eksamen i M4.

Eksempel 2.25. Vi kan bruke fourierrekken til heavisidefunksjonen til å finne summen til rekken

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots = \sum_{n=1}^{\infty} \frac{(-1)^n}{2n-1}$$

Siden

$$u(x) \sim \frac{1}{2} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1}{2n-1} \sin(2n-1)x$$

og u er glatt i $x = \pi/2$, ser vi at

$$\begin{aligned} 1 = u(\pi/2) &= \frac{1}{2} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1}{2n-1} \sin(2n-1) \frac{\pi}{2} \\ &= \frac{1}{2} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n}{2n-1}. \end{aligned}$$

Dette betyr at

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{2n-1} = \left(1 - \frac{1}{2}\right) \frac{\pi}{2} = \frac{\pi}{4}. \quad \triangle$$

Eksempel 2.26. Vi kan bruke Parsevals identitet til å finne summen til den kjente og kjære rekken

$$1 + \frac{1}{4} + \frac{1}{9} + \dots = \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

Siden

$$x \sim 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin nx,$$

gir Parsevals identitet at

$$\frac{2\pi^2}{3} = \frac{1}{\pi} \int_{-\pi}^{\pi} x^2 dx = \sum_{n=1}^{\infty} b_n^2 = 4 \sum_{n=1}^{\infty} \frac{1}{n^2},$$

eller

$$\frac{\pi^2}{6} = \sum_{n=1}^{\infty} \frac{1}{n^2}. \quad \triangle$$

Epilog

Du husker kanskje projeksjon fra M3. Vi gjengir et avsnitt. Denne epilogen er ikke essensiell lesning dersom du bare ønsker å komme deg gjennom eksamen, men viktig hvis du ønsker å sette fourieranalysen i sammenheng med ting du har lært tidligere.

Hvis vi har en ortogonal basis for et rom, er det veldig lett å finne en vektors komponenter i rommet. La oss si at vi ønsker å finne vektoren \mathbf{v} sine komponenter i basisen $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$. Komponentene til \mathbf{v} er gitt ved likningen

$$\mathbf{v} = x_1 \mathbf{u}_1 + x_2 \mathbf{u}_2 + \dots + x_n \mathbf{u}_n = U \mathbf{x}$$

Hvis vi ganger begge sider av denne likningen med U , får vi

$$U^* \mathbf{v} = U^* U \mathbf{x},$$

og siden kolonnene til U er ortogonale, blir den kvadratiske matrisen $U^* U$ diagonal:

$$U^* U = \begin{bmatrix} \mathbf{u}_1^* \mathbf{u}_1 & 0 & 0 & \dots & 0 \\ 0 & \mathbf{u}_2^* \mathbf{u}_2 & 0 & \dots & 0 \\ 0 & 0 & \mathbf{u}_3^* \mathbf{u}_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \mathbf{u}_n^* \mathbf{u}_n \end{bmatrix}.$$

Følgelig er løsningen av systemet $U^* \mathbf{v} = U^* U \mathbf{x}$ enkel å skrive opp.

Teorem 2.27. La $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ være en ortogonal basis for V , og la $\mathbf{v} \in V$. I så fall kan \mathbf{v} skrives

$$\mathbf{v} = \frac{\mathbf{u}_1^* \mathbf{v}}{\mathbf{u}_1^* \mathbf{u}_1} \mathbf{u}_1 + \frac{\mathbf{u}_2^* \mathbf{v}}{\mathbf{u}_2^* \mathbf{u}_2} \mathbf{u}_2 + \dots + \frac{\mathbf{u}_n^* \mathbf{v}}{\mathbf{u}_n^* \mathbf{u}_n} \mathbf{u}_n$$

Vi kan også projisere en vektor ned i et rom der den ikke hører hjemme. Projeksjonen minimerer avstanden fra rommet til vektoren.

Teorem 2.28. La $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ være en ortogonal basis for V , og la $\mathbf{v} \notin V$. Punktet

$$\mathbf{v}' = \frac{\mathbf{u}_1^* \mathbf{v}}{\mathbf{u}_1^* \mathbf{u}_1} \mathbf{u}_1 + \frac{\mathbf{u}_2^* \mathbf{v}}{\mathbf{u}_2^* \mathbf{u}_2} \mathbf{u}_2 + \dots + \frac{\mathbf{u}_n^* \mathbf{v}}{\mathbf{u}_n^* \mathbf{u}_n} \mathbf{u}_n$$

er det punktet i V som har kortest avstand til \mathbf{v} :

$$\|\mathbf{v} - \mathbf{v}'\| = \min_{\mathbf{w} \in V} \|\mathbf{v} - \mathbf{w}\|$$

Disse teoremene beviste vi for ortogonale systemer i dette kapitlet. Fourieranalyse spinner videre

på dette, og setter opp de samme teoremene, bare i uendelig mange dimensjoner. Vi definerer indreproduktet

$$\int_{-\pi}^{\pi} f \bar{g} dx,$$

lager oss en ortogonal mengde,

$$e^{inx} \quad n \in \mathbb{Z}$$

og prøver å skrive funksjoner som

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

der

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx.$$

Uttrykket

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}$$

kan betraktes som projeksjonen av f ned i det uendeligdimensjonale rommet utspent av funksjonene e^{inx} , og koeffisienten

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx$$

er f sin komponent langs vektoren e^{inx} . Beregningen

$$\int_{-\pi}^{\pi} e^{inx} e^{imx} dx = \begin{cases} 2\pi & \text{for } n = m \\ 0 & \text{for } n \neq m \end{cases}$$

ville vi i M3 skrevet

$$\mathbf{u}_n^* \mathbf{u}_n = 2\pi,$$

og

$$\mathbf{u}_n^* \mathbf{v} = \int_{-\pi}^{\pi} f(x) e^{-inx} dx,$$

og så satt sammen alt ved

$$\begin{aligned} \mathbf{v} &= \frac{\mathbf{u}_1^* \mathbf{v}}{\mathbf{u}_1^* \mathbf{u}_1} \mathbf{u}_1 + \frac{\mathbf{u}_2^* \mathbf{v}}{\mathbf{u}_2^* \mathbf{u}_2} \mathbf{u}_2 + \dots \\ &= \sum_{n=-\infty}^{\infty} c_n e^{inx} \end{aligned}$$

Det kan også nevnes at Parsevals teorem er en uendeligdimensjonal variant av Pytagoras' teorem, som du lærte allerede på ungdomsskolen.

Med dette i bakhuert, er det naturlig å spørre seg hva slags vektorrom funksjonene e^{inx} er basis for. Dette er et komplisert spørsmål, og ble ikke ordentlig besvart før tidlig på 1900-tallet. Rommet heter L^2 , og vi skal snakke litt løst om ingrediensene i konstruksjonen.

For å få et meningsfylt svar, må man for det første gå over til å betrakte ekvivalensklasser av funksjoner, heller enn funksjoner. Grunnen er at funksjoner som kun er forskjellige i enkelte punkter her og der, vil ha den samme fourierrekken, og dette kan vi ikke ha noe av, for et element i rommet bør helst kunne representeres på en entydig måte i basisen $\{e^{inx}\}$.

For det andre må man justere litt på konvergensbegrepet. Hva vil det si at rekken

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}$$

konvergerer til f ? Punktvis konvergens, altså at fourierrekken konvergerer til f for hver x , er greit hvis man skal bevise noe i et grunnleggende kurs i fourieranalyse, men om man løsner litt på konvergensbegrepet, vil man se at det finnes en mye større klasse av funksjoner kan representeres ved sin fourierrekke. Vi beviser heller at

$$\int_{-\pi}^{\pi} \left(f(x) - \sum_{n=-N}^N c_n e^{inx} \right)^2 dx \rightarrow 0 \quad \text{når } N \rightarrow \infty.$$

som kalles L^2 -konvergens. Denne konvergens sier i prinsippet at arealet mellom f og partialsummen skal gå mot null når partialsummen går mot fourierrekken til f .

Når alt dette er sagt og gjort, ekvivalensklasser av funksjoner er definert, og L^2 -konvergens er tatt i bruk, kan man vise at L^2 består, litt forenklet, av alle funksjoner som tilfredsstiller

$$\int_{-\pi}^{\pi} f \bar{f} dx = \int_{-\pi}^{\pi} |f|^2 dx < \infty.$$

Legendrepolynomene er en ortogonal basis for $L^2[-1, 1]$.

Joseph Fourier oppfant fourierrekker da han prøvde å løse varmelikningen tidlig på 1800-tallet. Ideene hans ble i noen grad avskrevet av samtidige matematikere, for bevisene hans holdt ikke alltid vann. Men eksemplene han oppdrev på at en diskontinuerlig funksjon kan skrives som en uendelig rekke av glatte funksjoner, var banebrytende, og stimulerer andre matematikere til å utvikle nytt maskineri for å rydde opp i disse konvergensspørsmålene, som Fourier bare gjettet tildels riktig på.

Andre artige opplysninger om Joseph Fourier, er at at han ble satt i fengsel under den franske revolusjon, dro til Egypt med Napoleon Bonaparte, og regnes som første som oppdaget drivhuseffekten.

3 Fouriertransform

Fouriertransform er egentlig et spesialtilfelle av laplacetransform, men vi kommer ikke til å komme så langt at vi ser det. Fourierrekker og fouriertransform er også to spesialtilfeller av en mer generell konstruksjon, men vi kommer ikke til å se det heller.

Vi skal senere bruke fouriertransform til å løse differensiallikninger på akkurat samme måte som vi gjorde med laplacetransform, men fouriertransform er teknisk vanskeligere å håndtere, og vi skal løse likninger som er mer kompliserte. I dette kapitlet skal vi gå gjennom fouriertransformens grunnleggende egenskaper.

Definisjon og grunnleggende egenskaper

Definisjon. Fouriertransformen til f er

$$\mathcal{F}(f) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-iwx} dx = \hat{f}(w),$$

der w er en reell variabel.

Akkurat som med laplacetransform, må vi ta stilling til hvilke funksjoner det er naturlig å ta fouriertransform til. Siden

$$\hat{f}(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) dx,$$

ser vi at det gir ingen mening å stappe inn en funksjon som ikke lar seg integrere på hele x -aksen. Det er vanlig å kreve at f er absolutt integrerbar, altså at

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

Det er mulig å slakke noe på dette kravet, med det skal ikke vi gjøre. Merk at e^{iwx} ikke går mot 0, slik som e^{-st} i laplacetransform. Funksjonen $e^{-iwx} = \cos wx - i \sin wx$ er en parametrisering av enhetsirkelen, med $|e^{-iwx}| = 1$ for alle w og x .

Teorem 3.1. Dersom f er absolutt integrerbar, konvergerer integralet

$$\int_{-\infty}^{\infty} f(x)e^{-iwx} dx.$$

absolutt.

Bevis. Dersom w og x er reelle, er $|e^{-iwx}| = 1$, slik at

$$\begin{aligned} \int_{-\infty}^{\infty} |f(x)e^{-iwx}| dx &= \int_{-\infty}^{\infty} |f(x)| \cdot |e^{-iwx}| dx \\ &= \int_{-\infty}^{\infty} |f(x)| dx < \infty. \quad \square \end{aligned}$$

Eksempel 3.2. La $a > 0$. Vi beregner

$$\begin{aligned} \mathcal{F}(e^{-a|x|}) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-a|x|} e^{-iwx} dx \\ &= \frac{1}{\sqrt{2\pi}} \left(\int_0^{\infty} e^{-ax} e^{-iwx} dx + \int_{-\infty}^0 e^{ax} e^{-iwx} dx \right) \\ &= \frac{1}{\sqrt{2\pi}} \left(\int_0^{\infty} e^{-x(a+iw)} dx + \int_{-\infty}^0 e^{x(a-iw)} dx \right) \\ &= \frac{1}{\sqrt{2\pi}} \left(\frac{1}{a+iw} + \frac{1}{a-iw} \right) \\ &= \frac{1}{\sqrt{2\pi}} \frac{2a}{a^2+w^2} = \sqrt{\frac{2}{\pi}} \frac{a}{a^2+w^2} \quad \triangle \end{aligned}$$

På samme måte som for laplacetransform, er det naturlig å spørre seg om fouriertransformen er entydig - kan to forskjellige funksjoner ha samme fouriertransform? Svaret er i prinsippet det samme som for laplacetransform - to funksjoner kan godt ha samme fouriertransform dersom de er forskjellige i isolerte punkter, men to kontinuerlige funksjoner kan bare ha samme fouriertransform dersom de er like.

Teorem 3.3. La f og g være stykkvis kontinuerlige funksjoner. Dersom

$$\mathcal{F}(f) = \mathcal{F}(g)$$

er

$$f = g$$

overalt der f og g er kontinuerlige.

Motivasjon - relasjon til fourierrekker

Man kan tenke at fouriertransform er en slags fourierrekke der $[-L, L]$ strekkes til å bli hele x -aksen. La f være en kontinuerlig og absolutt integrerbar funksjon. Vi setter opp fourierrekken til f på intervallet $(-L, L)$, og gjør en omskrivning:

$$\begin{aligned} f(x) &= \sum_{n=-\infty}^{\infty} c_n e^{i\frac{n\pi x}{L}} \\ &= \sum_{n=-\infty}^{\infty} \left(\frac{1}{2L} \int_{-L}^L f(y) e^{-i\frac{n\pi y}{L}} dy \right) e^{i\frac{n\pi x}{L}} \\ &= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \left(\frac{\pi}{L} \int_{-L}^L f(y) e^{-i\frac{n\pi y}{L}} dy \right) e^{i\frac{n\pi x}{L}}. \end{aligned}$$

Det siste uttrykket kan tolkes som en riemannsum på aksene der n telles fra $-\infty$ til ∞ . Gitteravstanden er $\frac{\pi}{L}$, og punktene er gitt ved $\frac{n\pi}{L}$. Hvis vi lar $L \rightarrow \infty$, vifter vi det vi kan med armer og bein og får

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(y) e^{-iwy} dy \right) e^{iwx} dw.$$

Vi kjenner igjen det innerste integralet som fouriertransformen til f :

$$\hat{f}(w) = \mathcal{F}(f) = \int_{-\infty}^{\infty} f(x) e^{-iwx} dx.$$

Det ytterste integralet kalles den inverse fouriertransformen, og vi skriver

$$\mathcal{F}^{-1}(\hat{f}) = \int_{-\infty}^{\infty} \hat{f}(x)e^{iwx} dw.$$

Teorem 3.4. Dersom \hat{f} er absolutt integrerbar, har vi

$$f(x) = \mathcal{F}^{-1}(\hat{f}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(w)e^{iwx} dw.$$

I motsetning til laplacetransform har vi tilgang på en formel for den inverse transform. Men det er allikevel vanlig å finne invers fouriertransform ved å slå opp i en tabell.

Regneregler

Fouriertransformasjonen er en lineær operasjon.

Teorem 3.5. Dersom a og b er reelle tall, og f og g er funksjoner, er

$$\mathcal{F}(af + bg) = a\mathcal{F}(f) + b\mathcal{F}(g).$$

Beviset er så enkelt at vi dropper det.

Teorem 3.6. La f' være absolutt integrerbar på x -aksen, og anta at $f(x) \rightarrow 0$ når $x \rightarrow \pm\infty$. Da er

$$\mathcal{F}(f') = iw\mathcal{F}(f).$$

Bevis. Vi tar en delvis integrasjon, og regner ut

$$\begin{aligned} \mathcal{F}(f') &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f'(x)e^{-iwx} dx \\ &= \frac{1}{\sqrt{2\pi}} f(x)e^{-iwx} \Big|_{-\infty}^{\infty} + \frac{iw}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-iwx} dx \\ &= iw\mathcal{F}(f). \end{aligned} \quad \square$$

Eksempel 3.7. Vi beregner

$$\mathcal{F}(e^{-x^2}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2} e^{-iwx} dx$$

Dette er litt jobb. Derivasjonsregelen over gir

$$\mathcal{F}(-2xe^{-x^2}) = iw\mathcal{F}(e^{-x^2}).$$

Men vi kan også observere at

$$\begin{aligned} \mathcal{F}(-2xe^{-x^2}) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} -2xe^{-x^2} e^{-iwx} dx \\ &= \frac{-2i}{\sqrt{2\pi}} \int_{-\infty}^{\infty} -ixe^{-x^2} e^{-iwx} dx \\ &= \frac{-2i}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2} \frac{d}{dw} e^{-iwx} dx \\ &= -2i \frac{d}{dw} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2} e^{-iwx} dx \\ &= -2i \frac{d}{dw} \mathcal{F}(e^{-x^2}). \end{aligned}$$

Hvis vi setter disse uttrykkene lik hverandre, får vi differensiallikningen

$$iw\mathcal{F}(e^{-x^2}) = -2i \frac{d}{dw} \mathcal{F}(e^{-x^2})$$

eller

$$\frac{d}{dw} \mathcal{F}(e^{-x^2}) + \frac{w}{2} \mathcal{F}(e^{-x^2}) = 0$$

for $\mathcal{F}(e^{-x^2})$. Integrerende faktor er

$$e^{w^2/4},$$

slik at

$$\frac{d}{dw} \left(e^{w^2/4} \mathcal{F}(e^{-x^2}) \right) = 0$$

eller

$$\mathcal{F}(e^{-x^2}) = C e^{-w^2/4}.$$

Husk fra nøtten i øving 2 at

$$\hat{f}(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2} dx = \frac{1}{\sqrt{2\pi}} \sqrt{\pi} = \frac{1}{\sqrt{2}},$$

slik at

$$\mathcal{F}(e^{-x^2}) = \frac{1}{\sqrt{2}} e^{-w^2/4}.$$

Dersom $a > 0$, kan vi gjøre den samme beregningen og få

$$\mathcal{F}(e^{-ax^2}) = \frac{1}{\sqrt{2a}} e^{-w^2/4a}. \quad \triangle$$

Konvolusjon

Konvolusjon kommer i mange former. I dette kapitlet definerer vi kovolusjon mellom to funksjoner f og g som

$$f * g = \int_{-\infty}^{\infty} f(v)g(x-v) dv.$$

Teorem 3.8.

$$\mathcal{F}(f * g) = \sqrt{2\pi} \mathcal{F}(f) \mathcal{F}(g).$$

Bevis. Vi bruker varablelskiftet $u = x - v$, $v = v$, og beregner

$$\begin{aligned} \mathcal{F}(f * g) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(v)g(x-v) dv e^{-iwx} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(v)g(x-v) e^{-iwx} dv dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(v)g(u) e^{-iw(u+v)} dv du \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(u) e^{-iwu} \int_{-\infty}^{\infty} f(v) e^{-iuv} dv du \\ &= \int_{-\infty}^{\infty} g(u) e^{-iwu} \mathcal{F}(f) du \\ &= \mathcal{F}(f) \int_{-\infty}^{\infty} g(u) e^{-iwu} du \\ &= \sqrt{2\pi} \mathcal{F}(f) \mathcal{F}(g) \end{aligned} \quad \square$$

Konvolusjon kan fremstå som noe umotivert. Men grunnen til at vi trenger det, er veldig enkel. Vi skal løse differensiallikninger med fouriertransform, og da vil vi få bruk for å inverstransformere produkter av fouriertransformer. Konvolusjonsteoremet forteller oss nøyaktig hvordan vi inverstransformerer et slikt produkt.

4 Partielle differensiallikninger

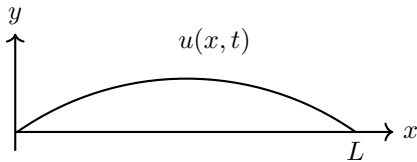
Fysiske situasjoner der det er behov for mer enn en uavhengig variabel, beskrives gjerne av partielle differensiallikninger. Vi skal ta for oss tre av de mest grunnleggende likningene, nemlig bølgelikningen, varmelikningen, og Laplaces likning.

Bølgelikningen

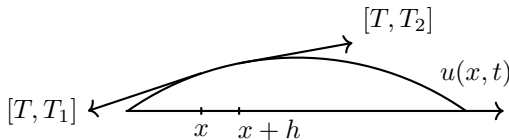
Bølgelikningen er en matematisk beskrivelse av en vibrerende streng, eller en stående luftbølge i en orgelpipe. Vi skal ta for oss selve likningen, hvor den kommer fra, og to forskjellige løsningsteknikker - en for et intervall på x -aksen, og en for hele x -aksen.

Utleddning

Vi tenker at vi har en vibrerende streng som er spent opp i $x = 0$ og $x = L$. La $u(x, t)$ være en funksjon som for hvert tidspunkt t og hvert punkt x beskriver utslaget fra likevektslinjen, som ligger langs x -aksen. Strengen har konstant massetetthet ρ [kg/m].



Vi tar en nærmere titt på strekkraftene på et lite stykke av strengen. Vi antar at tyngdekraften er neglisjerbar, og at strengen er helt elastisk, slik at strengestrekking, som virker parallelt med strengen, er eneste kraft. Vi antar at hvert punkt på strengen kun beveger seg loddrett, og at den horisontale komponenten av strengestrekking er konstant lik T .



Vi setter opp Newtons andre lov for den lille biten fra x til $x + h$. Massen til en bit med lengde h er $h\rho$, og akselerasjonen til strengen i punktet x er $u_{tt}(x, t)$. Netto kraft på biten er gitt ved $T_2 + T_1$, slik at

$$h\rho u_{tt}(x, t) = T_2 + T_1,$$

eller

$$\begin{aligned} \frac{\rho}{T} u_{tt}(x, t) &= \frac{T_2/T + T_1/T}{h} \\ &= \frac{u_x(x+h, t) - u_x(x, t)}{h}, \end{aligned}$$

siden stigningstallet til tangenten til strengen er gitt ved u_x . Lar vi nå $h \rightarrow 0$, får vi bølgelikningen

$$u_{tt}(x, t) = c^2 u_{xx}(x, t),$$

der $c^2 = \frac{T}{\rho}$.

Oppstilling av problem

Det er ikke nok med en differensialligning som beskriver strengens bevegelse. Vi må også ha informasjon om hvordan bevegelsen blir satt igang, og hvor strengen er spent opp. Et fullstendig oppstilt problem er:

$$u_{tt}(x, t) = c^2 u_{xx}(x, t), \quad (4.1)$$

med randkrav

$$u(0, t) = u(L, t) = 0, \quad (4.2)$$

og initialkrav

$$u(x, 0) = f(x) \quad u_t(x, 0) = g(x). \quad (4.3)$$

Selve differensialligningen (4.1) forteller oss hva slags fysiske lover som skal tilfredsstilles (i dette tilfelle Newtons andre lov), eller hva slags oppførsel vi kan forvente av løsningen, for eksempel at det er en vibrerende streng det er snakk om. Randkravene (4.2) forteller oss at strengen er spent opp i $x = 0$ og $x = L$, slik at løsningen står helt i ro der. Initialkravene (4.3) forteller oss noen om hvordan bevegelsen settes i gang; f angir strengens posisjon ved $t = 0$, mens g angir strengens fart ved $t = 0$. Når man spiller en tone på en gitar ved å dra i strengen og slippe den, slik man vanligvis gjør, er $g = 0$. Randkravene

$$u(0, t) = u(L, t) = 0, \quad (4.4)$$

kalles *dirichletrandkrav*.

Løsning på intervall - separasjon av variable

Et stort geni har engang tenkt at løsningen på bølgelikningen kan skrives

$$u(x, t) = F(x)G(t).$$

Han hadde rett. Innsetting i (4.1) gir

$$F(x)G''(t) = c^2 F''(x)G(t).$$

Vi deler på $c^2 F(x)G(t)$ og får

$$\frac{F''(x)}{F(x)} = \frac{G''(t)}{c^2 G(t)}.$$

Siden x og t skal kunne varieres uavhengig av hverandre, må vi ha at

$$\frac{F''(x)}{F(x)} = \frac{G''(t)}{c^2 G(t)} = k$$

der k er en foreløpig ubestemt konstant. Vi ganger opp med $c^2 F(x)G(t)$ og bytter fortegn på k , slik at

$$F''(x) + kF(x) = 0$$

og

$$G''(t) + kc^2 G(t) = 0.$$

Vi skal først prøve å finne ut hva k kan være. Vi kan bruke F og randkravene

$$u(0, t) = u(L, t) = 0,$$

til dette. La oss kikke på

$$F''(x) + kF(x) = 0.$$

Dersom $k = 0$, får vi

$$F''(x) = 0$$

som gir

$$F(x) = Ax + B.$$

Er dette en interessant løsning? Vel, nei. Dersom

$$F(0)G(t) = u(0, t) = 0,$$

må enten $F(0) = 0$ eller $G(t) = 0$. At $G(t) = 0$, slik at $u(x, t) = 0$, er en gyldig løsning av bølge­likningen, som også tilfredsstiller randkravene. Men dette er åpenbart ikke en spesielt interessant løsning, så jeg tror vi går for $F(0) = 0$, som impliserer $B = 0$. Det andre randkravet

$$u(L, t) = 0$$

gir likeledes at $F(L) = 0$, altså at $AL = 0$, som impliserer at $A = 0$. Dette impliserer igjen at $u(x, t) = 0$, vi konkluderer at $k = 0$ og $F(x) = 0$ ikke er en interessant løsning av problemet.

La oss prøve $k < 0$. I så fall løses

$$F''(x) + kF(x) = 0$$

av

$$F(x) = Ae^{\sqrt{-k}x} + Be^{-\sqrt{-k}x}$$

Dersom vi nå bruker randkravene, får vi de to likningene

$$\begin{aligned} A + B &= 0 \\ Ae^{\sqrt{-k}L} + Be^{-\sqrt{-k}L} &= 0 \end{aligned}$$

Dersom $k \neq 0$ (som vi jo allerede vet), er determinanten til dette systemet gitt ved

$$e^{\sqrt{-k}L} + e^{-\sqrt{-k}L} \neq 0$$

og vi konkluderer med at $A = B = 0$, slik at $F(x) = 0$. Altså er heller ikke $k < 0$ en interessant løsning.

Dersom $k > 0$, går alt så meget bedre, og vi får

$$F(x) = A \cos \sqrt{k}x + B \sin \sqrt{k}x.$$

Bruker vi randkravet $u(x, t) = 0$, får vi

$$A = 0,$$

og krever vi

$$F(L) = B \sin(\sqrt{k}L)G(t) = 0,$$

kan dette oppnås ved å sette $B = 0$, som er uinteressant siden da blir $u(x, t) = 0$, eller ved å kreve

$$\sqrt{k}L = n\pi.$$

slik at

$$k = \left(\frac{n\pi}{L}\right)^2.$$

Vi ser også at $n > 0$, for dersom $n < 0$ byttes bare fortegnet på B , som ennå er ubestemt. Ved å ta en titt på det endelige løsningen av problemet nedenfor, ser man at B kommer til å bli overflødig, så vi velger $B = 1$.

Ligningene

$$G''(t) + c^2 \left(\frac{n\pi}{L}\right)^2 G(t) = 0.$$

løses av

$$G_n(t) = A_n \cos c \frac{n\pi}{L} t + B_n \sin c \frac{n\pi}{L} t,$$

så de generelle løsningene til bølge­likningen med randkrav (4.2) blir

$$\begin{aligned} u_n(x, t) &= F(x)G_n(t) \\ &= \left(A_n \cos c \frac{n\pi}{L} t + B_n \sin c \frac{n\pi}{L} t \right) \sin \frac{n\pi}{L} x. \end{aligned}$$

Vi har ennå ikke tatt stilling til initialkravene (4.3). Det kan vi klare ved å skrive

$$\begin{aligned} u(x, t) &= \sum_{n=1}^{\infty} u_n(x, t) \\ &= \sum_{n=1}^{\infty} \left(A_n \cos c \frac{n\pi}{L} t + B_n \sin c \frac{n\pi}{L} t \right) \sin \frac{n\pi}{L} x. \end{aligned}$$

Dersom vi nå krever

$$f(x) = u(x, 0) = \sum_{n=1}^{\infty} u_n(x, 0) = \sum_{n=1}^{\infty} A_n \sin \frac{n\pi}{L} x,$$

ser vi at summen til høyre bør være fourierrekken til den odde utvidelsen til f , og følgelig bør

$$A_n = \frac{1}{L} \int_{-L}^L f(x) \sin \frac{n\pi}{L} x dx = \frac{2}{L} \int_0^L f(x) \sin \frac{n\pi}{L} x dx.$$

På samme vis, dersom

$$g(x) = u_t(x, 0) = \sum_{n=1}^{\infty} B_n c \frac{n\pi}{L} \sin \frac{n\pi}{L} x,$$

bør summen til høyre være fourierrekken til den odde utvidelsen til g , og følgelig må

$$B_n c \frac{n\pi}{L} = \frac{2}{L} \int_0^L g(x) \sin \frac{n\pi}{L} x dx.$$

slik at

$$B_n = \frac{2}{cn\pi} \int_0^L g(x) \sin \frac{n\pi}{L} x dx.$$

Når vi skal oppsummere, må vi være litt forsiktige, for vi har ikke sagt noe om deriverbarhet. En løsning av bølge­likningen bør helst være to ganger kontinu­erlig deriverbar, men dersom du drar i en gitarstreng med fingeren og slipper strengen, vil f har form som en sagtann, og ikke være deriverbar i punktet der fingeren din holder strengen. (Ihvertfall hvis strengen er elastisk, og fingeren din er uendelig tynn.) Av disse grunnene bør man kanskje ikke kalle oppsum­meringen for et teorem.

Bølgeligningen

$$u_{tt}(x, t) = c^2 u_{xx}(x, t),$$

med randkrav

$$u(0, t) = u(L, t) = 0,$$

og initialkrav

$$u(x, 0) = f(x) \quad u_t(x, 0) = g(x),$$

løses av

$$u(x, t) =$$

$$\sum_{n=1}^{\infty} \left(A_n \cos c \frac{n\pi}{L} t + B_n \sin c \frac{n\pi}{L} t \right) \sin \frac{n\pi}{L} x,$$

der

$$A_n = \frac{2}{L} \int_0^L f(x) \sin \frac{n\pi}{L} x \, dx$$

og

$$B_n = \frac{2}{cn\pi} \int_0^L g(x) \sin \frac{n\pi}{L} x \, dx.$$

Løsning for fløyte

En stående trykkbølge inne i fløyte, beskrives av problemet

$$u_{tt}(x, t) = c^2 u_{xx}(x, t),$$

med randkrav

$$u_x(0, t) = u_x(L, t) = 0,$$

og initialkrav

$$u(x, 0) = f(x) \quad u_t(x, 0) = g(x).$$

Konstanten c avhenger nå av trykket og lydshastigheten i mediet der lydbølgene produseres. (Dersom du spiller på fløyten inni en gassballong full av helium, blir c høyere enn i luft.) Randkravene kalles *von-Neumann-randkrav*, og L er lengden på fløyten.

Vi løser problemet på samme måte som for den vibrerende strengen. Den eneste forskjellen blir i forbindelse med løsning av

$$F''(x) + kF(x) = 0.$$

På samme vis må $k > 0$, slik at

$$F(x) = A \cos \sqrt{k}x + B \sin \sqrt{k}x.$$

Siden

$$F'(x) = -A\sqrt{k} \sin \sqrt{k}x + B\sqrt{k} \cos \sqrt{k}x$$

impliserer

$$u_x(0, t) = F'(0) = 0,$$

at $B = 0$, mens

$$u_x(L, t) = F'(L) = -A\sqrt{k} \sin \sqrt{k}L = 0,$$

gir som før at

$$\sqrt{k}L = n\pi.$$

Merk at $n = 0$ er en gyldig løsning her siden en konstant funksjon løser både bølgeligningen og von-Neumann-randkravene. Resten blir som før, men vi må bruke cosinusrekken til f og g istedet for sinusrekken. Vi tar ikke resten av regningen.

Bølgeligningen

$$u_{tt}(x, t) = c^2 u_{xx}(x, t),$$

med randkrav

$$u_x(0, t) = u_x(L, t) = 0,$$

og initialkrav

$$u(x, 0) = f(x) \quad u_t(x, 0) = g(x),$$

løses av

$$u(x, t) =$$

$$A + \sum_{n=1}^{\infty} \left(A_n \cos c \frac{n\pi}{L} t + B_n \sin c \frac{n\pi}{L} t \right) \cos \frac{n\pi}{L} x,$$

der A er en vilkårlig konstant,

$$A_n = \frac{2}{L} \int_0^L f(x) \cos \frac{n\pi}{L} x \, dx$$

og

$$B_n = \frac{2}{cn\pi} \int_0^L g(x) \cos \frac{n\pi}{L} x \, dx.$$

Løsning på hele x -aksen - D'Alembert

Hvis man ikke ønsker å bruke bølgeligningen til å beskrive en oppspennet streng, men heller bølgene fra et steinkast på et endimensjonalt og uendelig langt hav, må man studere bølgeligningen

$$u_{tt}(x, t) = c^2 u_{xx}(x, t),$$

med initialkrav

$$u(x, 0) = f(x) \quad u_t(x, 0) = g(x),$$

og ingen randkrav. Her skal løsningen gjelde for alle x , ikke bare på intervallet $[0, L]$

I noen tilfeller kan dette gjøre alt vanskeligere, men akkurat i tilfellet bølgeligningen, blir alt fryktelig enkelt. Det er enkelt å sjekke at funksjonen

$$u(x, t) = \phi(x + ct) + \psi(x - ct)$$

passer i likningen

$$u_{tt}(x, t) = c^2 u_{xx}(x, t).$$

uansett hva ϕ og ψ er, så lenge de er to ganger kontinuerlig deriverbare. Spørsmålet blir bare da hvordan vi skal klare å innpasse initialkravene.

Dersom man bruker

$$u(x, 0) = f(x)$$

får man

$$\phi(x) + \psi(x) = f(x)$$

og dersom man bruker

$$u_t(x, 0) = g(x),$$

får man

$$c\phi'(x) - c\psi'(x) = g(x),$$

eller

$$\phi(x) - \psi(x) = \frac{1}{c} \int^x g(t) dt + C$$

Vi har nå et lineært 2×2 -likningssystem for ϕ og ψ .

Legger vi likningene sammen, får vi

$$2\phi(x) = f(x) + \frac{1}{c} \int^x g(t) dt + C$$

og trekker vi dem fra hverandre, får vi

$$2\psi(x) = f(x) - \frac{1}{c} \int^x g(t) dt - C$$

Vi setter nå alt sammen igjen

$$\begin{aligned} u(x, t) &= \phi(x + ct) + \psi(x - ct) \\ &= \frac{1}{2} (f(x + ct) + f(x - ct)) \\ &\quad + \frac{1}{2c} \left(\int^{x+ct} g(t) dt - \int^{x-ct} g(t) dt \right) \\ &= \frac{1}{2} (f(x + ct) + f(x - ct)) + \frac{1}{2c} \int_{x-ct}^{x+ct} g(t) dt, \end{aligned}$$

og oppsummerer.

Bølgelikningen

$$u_{tt}(x, t) = c^2 u_{xx}(x, t),$$

på hele x -aksen, med initialkrav

$$u(x, 0) = f(x) \quad u_t(x, 0) = g(x),$$

løses av

$$u(x, t) = \frac{1}{2} (f(x + ct) + f(x - ct)) + \frac{1}{2c} \int_{x-ct}^{x+ct} g(t) dt.$$

Varmelikningen

Varmelikningen beskriver kort og godt varmeglyt i materialer. Vi skal ta for oss den enkleste varianten, der varmeledningsevnen er identisk overalt og i alle retninger i materialet.

Utleddning

I en ideell gass er det et lineært forhold mellom temperatur og opplagret bevegelsesenergi, gitt ved

$$\frac{1}{2}mv^2 = \frac{3}{2}kT$$

der $k = 1.38064852(79) \times 10^{-23}$ J/K er Boltzmanns konstant, v er gjennomsnittsfarten til gassmolekylene, og T er temperaturen målt i Kelvin. Vi kan altså

tenke på temperatur som et mål på opplagret energi av et eller annet slag. I andre materialer enn ideelle gasser er virkeligheten noe mer komplisert, men det skal vi ikke gå inn på.

La oss tenke at denne opplagrede varmeenergien (eller temperatur om du vil), er gitt ved funksjonen $u(x, y, z, t)$, der (x, y, z) er de romlige koordinatene, t er tidspunktet, og u måles i J/m^3 eller liknende. La Ω være et område i \mathbb{R}^3 . Den totale opplagrede energien i Ω er gitt ved

$$\iiint_{\Omega} u(x, y, z, t) dx dy dz.$$

Dersom vi legger noen milde restriksjoner på glattheten til u , er endringen i denne energien over tid gitt ved

$$\begin{aligned} \frac{\partial}{\partial t} \iiint_{\Omega} u(x, y, z, t) dV = \\ \iiint_{\Omega} u_t(x, y, z, t) dV \end{aligned}$$

Fouriers lov sier at endringen i varmeenergi i Ω enten skyldes varme som produseres inne i Ω (tenk at det står et stearinlys og brenner inne i Ω), eller varme som slipper inn og ut gjennom randen $\partial\Omega$, gitt ved

$$\iint_{\partial\Omega} \mathbf{F} dS = \iiint_{\Omega} \operatorname{div} \mathbf{F} dV$$

der F er et vektorfelt som beskriver varmeglyten. Vi skal ikke ta i betraktning tilfeller der varme produseres inne i Ω , så Fouriers lov blir

$$\iiint_{\Omega} u_t dV = \iiint_{\Omega} \operatorname{div} \mathbf{F} dV.$$

For mange materialer er det rimelig å anta at

$$\mathbf{F} = c^2 \nabla u,$$

der c er en konstant relatert til varmeledningsevnen til mediet, slik at fouriers lov blir

$$\begin{aligned} \iiint_{\Omega} u_t dV &= \iiint_{\Omega} \operatorname{div} \nabla u dV \\ &= \iiint_{\Omega} u_{xx} + u_{yy} + u_{zz} dV \end{aligned}$$

og siden dette bør gjelde for et tilfeldig valgt område Ω , bør

$$u_t = c^2(u_{xx} + u_{yy} + u_{zz}),$$

som kalles *varmelikningen* eller *diffusjonslikningen*.

Varmelikningen er der alt startet. Fouriers lov er oppkalt etter Joseph Fourier, og det var i forbindelse med hans utledning og analyse av varmelikningen at fourieranalsen så dagens lys.

Løsning på begrenset intervall

Nå skal vi tenke oss en isolert stang der temperaturen er gitt av $u(x, t)$. Varme kan slippe ut eller tilføres gjennom endepunktene, men ingen andre steder. Vi skal løse varmelikningen

$$u_t = c^2 u_{xx}$$

på stang med lengde L . Randkravene

$$u(0, t) = u(L, t) = 0, \quad (4.5)$$

sier at temperaturen holdes konstant lik 0 i endepunktene, og initialkravet

$$u(x, 0) = f(x) \quad (4.6)$$

sier at temperaturfordelingen er gitt ved $f(x)$ når tiden begynner å gå.

Vi prøver separasjon av variable

$$u(x, t) = F(x)G(t).$$

På samme vis som for bølgeligningen, får vi

$$F''(x) + kF(x) = 0$$

og

$$G'(t) + kc^2G(t) = 0.$$

Å finne F og k blir eksakt repetisjon av argumentet for bølgeligningen, mens

$$G_n(t) = A_n e^{-\left(\frac{cn\pi}{L}\right)^2 t}$$

slik at de generelle løsningene til bølgeligningen med randkrav $u(0, t) = u(L, t) = 0$ blir

$$u_n(x, t) = F(x)G_n(t) = A_n e^{-\left(\frac{cn\pi}{L}\right)^2 t} \sin \frac{n\pi}{L} x.$$

Likeledes kan initialkravet $u(x, 0) = f(x)$ inkorporeres ved å legge sammen alle løsninger

$$u(x, t) = \sum_{n=1}^{\infty} u_n(x, t) = \sum_{n=1}^{\infty} A_n e^{-\left(\frac{cn\pi}{L}\right)^2 t} \sin \frac{n\pi}{L} x,$$

sette $t = 0$

$$f(x) = u(x, 0) = \sum_{n=1}^{\infty} A_n \sin \frac{n\pi}{L} x,$$

og observere at vi får dette til dersom A_n er fourierkoeffisientene til den odde utvidelsen til f

$$A_n = \frac{2}{L} \int_0^L f(x) \sin \frac{n\pi}{L} x dx.$$

Varmelikningen

$$u_t(x, t) = c^2 u_{xx}(x, t),$$

med randkrav

$$u(0, t) = u(L, t) = 0,$$

og initialkrav

$$u(x, 0) = f(x)$$

løses av

$$u(x, t) = \sum_{n=1}^{\infty} A_n e^{-\left(\frac{cn\pi}{L}\right)^2 t} \sin \frac{n\pi}{L} x,$$

der

$$A_n = \frac{2}{L} \int_0^L f(x) \sin \frac{n\pi}{L} x dx.$$

Løsning på hele x -aksen

Vi skal nå løse varmeligningen

$$u_t = c^2 u_{xx}$$

på en uendelig lang stang. Randkravene skal være

$$\lim_{x \rightarrow \pm\infty} u(x, t) = 0, \quad (4.7)$$

mens initialkravet er som før

$$u(x, 0) = f(x), \quad (4.8)$$

der vi antar

$$\lim_{x \rightarrow \pm\infty} f(x) = 0. \quad (4.9)$$

Vi skal løse dette problemet ved å fouriertransformere varmeligningen med hensyn på x . Siden $\lim_{x \rightarrow \pm\infty} u(x, t) = 0$, satser vi på at det går bra, og vi får

$$\mathcal{F}(u_t) = c^2 \mathcal{F}(u_{xx}).$$

Vi tar en kikk på disse to. Hvis vi antar at u er kontinuerlig deriverbar i t , og fouriertransformen er med hensyn på x , kan vi skrive

$$\begin{aligned} \mathcal{F}(u_t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u_t(x, t) e^{-iwx} dx \\ &= \frac{\partial}{\partial t} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(x, t) e^{-iwx} dx = \hat{u}_t(w, t), \end{aligned}$$

og

$$\mathcal{F}(u_{xx}) = -w^2 \mathcal{F}(u) = -w^2 \hat{u}(w, t),$$

slik at

$$\hat{u}_t = -c^2 w^2 \hat{u}.$$

Vi later som om dette er en ordinær differensialligning for \hat{u} , og får

$$\hat{u}(w, t) = A(w) e^{-c^2 w^2 t}.$$

Fouriertransformerer vi initialkravet, ser vi at

$$A(w) = \hat{u}(w, 0) = \mathcal{F}(u(x, 0)) = \mathcal{F}(f(x)) = \hat{f}(w),$$

og da er det egentlig bare å inversfouriertransformere, og så har vi løsningen

$$u(x, t) = \mathcal{F}^{-1}(\hat{u}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(w) e^{-c^2 w^2 t} e^{iwx} dw.$$

Vel og bra, men vi kan komme oss et knepp til. Husk at

$$f * g = \mathcal{F}^{-1}(\sqrt{2\pi} \hat{f} \hat{g}),$$

og observer at

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(w) e^{-c^2 w^2 t} e^{iwx} dw = \mathcal{F}^{-1}(\hat{f} e^{-c^2 w^2 t}).$$

Med andre ord, hvis vi kunne inverstransformere $e^{-c^2 w^2 t}$ hadde det vært bra. Husk fra tabellen at

$$\mathcal{F}(e^{-ax^2}) = \frac{1}{\sqrt{2a}} e^{-\frac{w^2}{4a}}.$$

Setter vi

$$a = \frac{1}{4c^2 t},$$

og ganger og deler litt med konstanter, kan vi si at

$$\mathcal{F}\left(\frac{1}{c\sqrt{2t}}e^{-\frac{x^2}{4c^2t}}\right) = e^{-c^2w^2t},$$

som gir

$$\begin{aligned} u(x,t) &= \mathcal{F}^{-1}(\hat{f}e^{-c^2w^2t}) = \frac{1}{\sqrt{2\pi}} \left(f * \frac{1}{c\sqrt{2t}}e^{-\frac{x^2}{4c^2t}} \right) \\ &= \frac{1}{2c\sqrt{\pi t}} \int_{-\infty}^{\infty} f(v)e^{-\frac{(x-v)^2}{4c^2t}} dv \end{aligned}$$

Onde tunger vil hevde at $u(x,t)$ ikke er definert for $t = 0$. Ikke bra, for vi har jo prøvd å få til at $u(x,0) = f(x)$, som åpenbart ikke er helt riktig. Det kan vises at

$$\lim_{t \rightarrow 0} u(x,t) = f(x)$$

men det skal ikke vi gjøre.

Varmelikningen

$$u_t = c^2 u_{xx},$$

på hele x -aksen med initialkrav

$$u(x,0) = f(x)$$

løses av

$$u(x,t) = \frac{1}{2c\sqrt{\pi t}} \int_{-\infty}^{\infty} f(v)e^{-\frac{(x-v)^2}{4c^2t}} dv,$$

og vi har

$$\lim_{t \rightarrow 0} u(x,t) = f(x).$$

Laplace's ligning

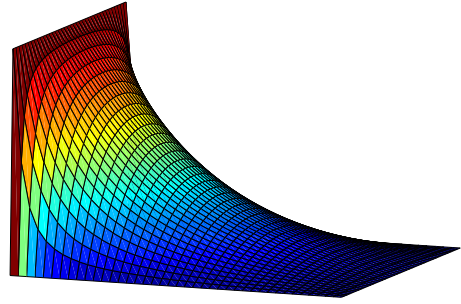
For å forstå hvor Laplace' ligning kommer fra, må man betrakte varmeligningen i romlige dimensjoner

$$u_t = c^2 (u_{xx} + u_{yy}).$$

Det fysiske bildet du kan ha, er en tynn varmeisoleret plate der varmeenergi kan forsvinne ut eller inn gjennom sidekantene. Løsningen til varmeligningen vil da gi temperaturen $u(x,y,t)$ i platen gitt fornuftige rand- og initialkrav. Vi skal ikke løse varmeligningen i to romlige dimensjoner, men derimot betrakte spesialtilfellet der $t \rightarrow \infty$, slik at varmemestrømmen er konstant, eller steady-state flow som vi sier på norsk. Dersom varmemestrømmen er konstant, er $u_t(x,y,t) = 0$, slik at varmeligningen blir

$$u_{xx} + u_{yy} = 0.$$

Dette kalles Laplace' ligning, og den kan vi løse med teknikker vi har lært i kurset. Her er en figur som beskriver en typisk situasjon. Dette er temperaturen i en tynn plate der temperaturen holdes konstant lik null på tre sider, og konstant lik fem eller en eller noe på den siste siden.



Løsning på begrenset intervall - separasjon av variable

Vi setter opp problemet

$$u_{xx} + u_{yy} = 0,$$

på rektangelet $[0, a] \times [0, b]$ med randkrav

$$u(x,0) = u(0,y) = u(a,y) = 0 \quad \text{og} \quad u(x,b) = f(x).$$

Vi separerer i vei

$$u(x,y) = F(x)G(y).$$

På samme vis som for bølgeligningen og varmeligningen, får vi

$$F''(x) + kF(x) = 0$$

og

$$G''(y) - kG(y) = 0.$$

Å finne F og k blir eksakt repetisjon av argumentet for bølgeligningen, mens

$$G_n(y) = A_n e^{-\frac{n\pi}{a}y} + B_n e^{\frac{n\pi}{a}y}$$

slik at de generelle løsningene som tilfredsstiller $u(0,y) = u(b,y) = 0$ blir

$$u_n(x,y) = F(x)G_n(y) = (A_n e^{-\frac{n\pi}{a}y} + B_n e^{\frac{n\pi}{a}y}) \sin \frac{n\pi}{a}x.$$

Krever vi $u(x,0) = 0$, får vi

$$A_n = -B_n,$$

slik at

$$\begin{aligned} u_n(x,y) &= A_n (e^{-\frac{n\pi}{a}y} - e^{\frac{n\pi}{a}y}) \sin \frac{n\pi}{a}x \\ &= A_n \sinh \frac{n\pi}{a}y \sin \frac{n\pi}{a}x, \end{aligned}$$

og krever vi $u(x,b) = f(x)$, får vi det til ved å legge sammen alle løsninger

$$u(x,y) = \sum_{n=1}^{\infty} u_n(x,y) = \sum_{n=1}^{\infty} A_n \sinh \frac{n\pi}{a}y \sin \frac{n\pi}{a}x,$$

og kreve

$$f(x) = u(x,b) = \sum_{n=1}^{\infty} A_n \sinh \frac{n\pi b}{a} \sin \frac{n\pi}{a}x,$$

slik at

$$A_n = \frac{2}{a \sinh \frac{n\pi b}{a}} \int_0^a f(x) \sin \frac{n\pi x}{a} dx.$$

Løsning i halvplanet

Vi skal se på Laplaces ligning

$$u_{xx} + u_{yy} = 0,$$

i halvplanet. Randkrav skal være

$$\lim_{x \rightarrow \pm\infty} u(x, y) = \lim_{y \rightarrow \infty} u(x, y) = 0$$

og

$$u(x, 0) = f(x),$$

der

$$\lim_{x \rightarrow \pm\infty} f(x) = 0.$$

Siden $\lim_{x \rightarrow \pm\infty} u(x, y) = 0$ satser vi på at u er absolutt integrerbar i x , slik at vi kan fouriertransformere med hensyn på x . Da får vi

$$\mathcal{F}(u_{xx}) + \mathcal{F}(u_{yy}) = \mathcal{F}(0)$$

eller

$$-w^2 \hat{u}(w, y) + \hat{u}_{yy}(w, y) = 0.$$

Akkurat som med varmeligningen, later vi som om dette er en ordinær differensalligning i y , og får

$$\hat{u}(w, y) = A(w)e^{-|w|y} + B(w)e^{|w|y}.$$

Nå må vi fouriertransformere randkravet $\lim_{y \rightarrow \infty} u(x, y) = 0$, og da får vi

$$\lim_{y \rightarrow \infty} \hat{u}(w, y) = 0.$$

Bruker vi dette, ser vi at $B(w) = 0$, og følgelig er

$$\hat{u}(w, y) = A(w)e^{-|w|y}.$$

Transformerer vi $u(x, 0) = f(x)$, får vi

$$\hat{u}(w, 0) = \hat{f}(w),$$

og bruker vi denne, får vi

$$\hat{u}(w, y) = \hat{f}(w)e^{-|w|y}.$$

Nå kan vi inverstransformere og få

$$\begin{aligned} u(x, y) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(w)e^{-|w|y} e^{iwx} dw \\ &= \mathcal{F}^{-1}(\hat{f}(w)e^{-|w|y}). \end{aligned}$$

Nå bruker vi samme trikset som for varmeligningen. Siden dette er inversfouriertransformen til produktet mellom $\hat{f}(w)$ og $e^{-|w|y}$, og

$$\mathcal{F}\left(\sqrt{\frac{2}{\pi}} \frac{y}{x^2 + y^2}\right) = e^{-|w|y},$$

ser vi at

$$\begin{aligned} u(x, y) &= \frac{1}{\sqrt{2\pi}} f * \left(\sqrt{\frac{2}{\pi}} \frac{y}{x^2 + y^2}\right) \\ &= \frac{1}{\pi} \int_{-\infty}^{\infty} f(v) \frac{y}{(x-v)^2 + y^2} dv. \end{aligned}$$

Noen artige egenskaper

En funksjon som tilfredsstill

$$u_{xx} + u_{yy} = 0$$

kalles gjerne en *harmonisk funksjon*. Harmoniske funksjoner har morsomme egenskaper.

Teorem 4.1. La u være harmonisk på Ω , og la $\mathcal{C} \in \Omega$ være en sirkelskive med sentrum i (x_0, y_0) og radius a . Da gjelder at

$$u(x_0, y_0) = \frac{1}{2\pi a} \int_{\partial\mathcal{C}} u ds = \frac{1}{\pi a^2} \iint_{\mathcal{C}} u dA.$$

Vi lar denne stå ubevist inntil videre. Beviset kommer ikke på eksamen i år. Ta kontakt om du lur

Teorem 4.2. La u være harmonisk på Ω , og la $\mathcal{C} \in \Omega$ være en sirkelskive med sentrum i (x_0, y_0) og radius a . Da ligger

$$\max_{(x,y) \in \mathcal{C}} u(x, y)$$

og

$$\min_{(x,y) \in \mathcal{C}} u(x, y)$$

på $\partial\mathcal{C}$.

Vi beviser ikke denne heller.

Del II
Numerikk

5 Numeriske likningslødere

Andregradslikningen

$$ax^2 + bx + c = 0$$

kan vi løse med formelen

$$x = \frac{b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Men i mange anvendelser dukker det opp likninger ikke kan løses analytisk. Et klassisk eksempel er $x = \cos x$. I disse tilfellene må vi velge blant forskjellige numeriske lødere istedet. En løsning r av en likning kalles gjerne en *rot*.

Halveringsmetoden

Vi søker r slik at

$$f(r) = 0.$$

Dersom f er kontinuerlig og bytter fortegn i r , kan vi benytte denne informasjonen til å finne en approksimasjon x_n til r .

La oss anta at vi kjenner a og b slik at $r \in [a, b]$ er den eneste roten til f i $[a, b]$. Vi definerer nå $x_1 = a$, $x_2 = b$, og $x_3 = \frac{a+b}{2}$. Fortegnet til $f(x_3)$ vil fortelle oss om $x_3 < r$ eller $x_3 > r$, og følgelig vil vi kunne si med sikkerhet om $r \in [x_1, x_3]$ eller $r \in [x_3, x_2]$.

Intervallene $[x_1, x_3]$ og $[x_3, x_2]$ er akkurat halvparten så lange som $[x_1, x_2]$, så etter å ha utført denne prosessen, har vi en mer nøyaktig ide om hvor r befinner seg. Vi setter så $x_2 = x_3$ eller $x_1 = x_3$, alt etter fortegnet til $f(x_3)$, og repeterer prosedyren. Dersom vi gjør dette n ganger, ender vi opp med et intervall med lengde

$$\frac{b-a}{2^n},$$

og hvis vi til slutt setter $r_n = \frac{x_1+x_2}{2}$, vet vi at

$$|r - x_n| \leq \frac{b-a}{2^{n+1}}.$$

Teorem 5.1. Dersom f er en kontinuerlig funksjon med en rot i intervallet (a, b) , produserer n steg med halveringsmetoden et estimat som tilfredsstillter

$$|r - x_n| \leq \frac{b-a}{2^{n+1}}.$$

Fikspunktmetoden

Anta at vi har en likning på formen

$$r = g(r).$$

Iterasjonen

$$x_{n+1} = g(x_n).$$

vil konvergere mot r under noen omstendigheter. Det kommer litt an på g . Vi begynner med et eksempel.

Eksempel 5.2. Vi løser polynomlikningen

$$x^3 + x^2 - 3x - 3 = 0.$$

Dette polynomet kan spaltes i

$$x^3 + x^2 - 3x - 3 = (x+1)(x-\sqrt{3})(x+\sqrt{3}),$$

og vi skal se hvordan fikspunktmetoden leter etter de forskjellige løsningene. Likningen kan skrives om til $x = g(x)$ på flere måter, men vi skal begynne med å skrive

$$x = \frac{1}{3}(x^3 + x^2 - 3),$$

slik at

$$g(x) = \frac{1}{3}(x^3 + x^2 - 3),$$

og

$$x_{n+1} = \frac{1}{3}(x_n^3 + x_n^2 - 3).$$

Vi prøver å finne løsningen $r = \sqrt{3} \approx 1.732050807568877$, og starter derfor en kjøring i $x_0 = 1.5$. Vi får:

x_0	1.5000000000000000
x_5	-0.995705356719772
x_{10}	-0.999982551541273
x_{15}	-0.999999928199386
x_{20}	-0.99999999704524
x_{25}	-0.99999999998784
x_{30}	-0.99999999999995
x_{35}	-1.0000000000000000

Også nå ønsker metoden heller å finne $r = -1$. \triangle

Eksemplet over er kjørt med en kodesnutt som ser slik ut:

```
import numpy as np
import time

#initialgjetning
x=-1.5

#denne koden er basert paa at vi overskriver ...
#en enkelt variabel istedet for aa lagre ...
#hele rekken av iterasjoner. derfor ...
#trenger vi en ekstra variabel for aa ...
#holde styr paa avstanden mellom ...
#iterasjonene.
y=x-1

#hvor tett skal det vaere mellom ...
#iterasjonene foer vi er fornøyde?
tol=10**(-16)

#fikspunkt for x-cos x = 0
while np.abs(y-x) > tol: # vi holder paa saa ...
    lenge avstanden mellom iterasjonene er ...
    større en tol
    y=x # y er naa forrige iterasjon
    x=(x**3+x**2-3)/3 #dette er ...
    fikspunktiterasjonen
    time.sleep(.5) #en liten pause, slik at ...
    vi skal faa tid til aa nyte synet av ...
    den nye iterasjonen
print x, np.abs(x-y) #som printes her
```

Fikspunktmetoden i forrige eksempel viste en sterk preferanse på hvilken rot den hadde lyst til å finne. Men man kan skrive om likningen til $x = g(x)$ på mange måter.

Eksempel 5.3. Vi skriver nå om likningen

$$x^3 + x^2 - 3x - 3 = 0.$$

til

$$x = \frac{3 + 3x - x^2}{x^2}.$$

Starter vi i $x_0 = -1.5$, får vi

x_0	-1.5000000000000000
x_5	-1.732004423011461
x_{10}	-1.732050803458349
x_{15}	-1.732050807568513
x_{20}	-1.732050807568878
x_{25}	-1.732050807568877

Denne fikspunktmetoden klarte fint å finne roten $r = -\sqrt{3}$. \triangle

Som vi ser av de to foregående eksemplene, kan fikspunktmetoden konvergere mot forskjellige røtter avhengig av hvordan vi skriver om likningen. Den kan også ikke konvergere i det hele tatt.

Eksempel 5.4. Vi prøver igjen

$$x = \frac{3 + 3x - x^2}{x^2}.$$

Starter vi i $x_0 = 1.5$, i håp om å finne $r = \sqrt{3}$, får vi

x_0	-1.5000000000000000
x_1	2.3333333333333333
x_2	0.836734693877551
x_3	6.870315288518744
x_4	-0.499781154362809
x_5	5.007884193672099
x_6	-0.281322161800267
x_7	26.242541136990940
x_8	-0.881325585764740
x_9	-0.541641177723142
x_{10}	3.687092259260734

Denne fikspunktmetoden klarte ikke å finne noe som helst når vi startet i $x_0 = 1.5$. \triangle

For å skjønne hva som skjer, kan vi Taylorutvikle g om r :

$$g(x_n) = g(r) + g'(r)(x_n - r) + \frac{g''(r)}{2}(x_n - r)^2 + \frac{g'''(r)}{3!}(x_n - r)^3 + \dots$$

og skrive

$$\begin{aligned} x_{n+1} - r &= g(x_n) - g(r) \\ &= g'(r)(x_n - r) + \frac{g''(r)}{2}(x_n - r)^2 + \dots \end{aligned}$$

og da har vi en likning som sier noe om størrelsen på $x_{n+1} - r$ som en funksjon av $x_n - r$, altså hvor mye feilen minker fra iterasjon til iterasjon. For fikspunktiterasjonen ser vi at det blir best konvergens om g' er så liten som mulig. Vi tar med et konvergensteorem.

Teorem 5.5. La g være en kontinuerlig deriverbar funksjon. Dersom både $a < g(x) < b$ og $|g'(x)| \leq L < 1$ på $[a, b]$, finnes et entydig punkt r slik at

$$r = g(r).$$

Fikspunktmetoden

$$x_{n+1} = g(x_n)$$

konvergerer mot r dersom $x_0 \in [a, b]$.

Eksempel 5.6. I eksemplene over er

$$g(x) = \frac{3 + 3x - x^2}{x^2}.$$

og

$$g(x) = \frac{x^3 + x^2 - 3}{3}.$$

Hvis du deriverer disse og evaluerer i røttene til polynomet $x^3 + x^2 - 3x - 3$, vil du se et tydelig mønster. Fikspunktmetoden greier ikke finne r dersom $|g'(r)| > 1$. \triangle

Til slutt kan nevnes at fikspunktmetoden også kan brukes på systemer på formen

$$\mathbf{x} = \mathbf{G}(\mathbf{x})$$

der $\mathbf{x} \in \mathbb{R}^n$ og $\mathbf{G} : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Iterasjonen er

$$\mathbf{x}_{n+1} = \mathbf{G}(\mathbf{x}_n).$$

Å analysere konvergensen til denne iterasjonen blir for hardt for oss, men vi skal få bruk for selve iterasjonen i kapitlet om numeriske metoder for ordinære differensiallikninger.

Newton's metode

Anta at likningen er på formen

$$f(r) = 0.$$

Prosedyren for Newtons metode er som følger: slå tangenten til funksjonen f i iterasjonen x_n . Punktet der tangenten skjærer x -aksen er den nye iterasjonen x_n . Hvis vi setter opp likningen for tangenten

$$y - f(x_n) = f'(x_n)(x - x_n),$$

og sier at iterasjonen x_{n+1} er den x -verdien slik at $y = 0$:

$$-f(x_n) = f'(x_n)(x_{n+1} - x_n),$$

blir iterasjonen

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Eksempel 5.7. Vi søker løsningene til

$$x^4 - 10x^3 + 35x^2 - 50x + 24 = 0$$

Som Ingrid Espelid Hovig har vi jukset litt, og valgt et polynom som faktoriseres pent:

$$x^4 - 10x^3 + 35x^2 - 50x + 24 = (x-1)(x-2)(x-3)(x-4)$$

Newtons metode blir:

$$x_{n+1} = x_n - \frac{x_n^4 - 10x_n^3 + 35x_n^2 - 50x_n + 24}{4x_n^3 - 30x_n^2 + 70x_n - 50}$$

La oss lete etter $r = 1$. Vi starter i $x_0 = 0.5$, og får:

x_0	0.5000000000000000
x_1	0.7982954545454545
x_2	0.950817599863883
x_3	0.996063283034122
x_4	0.999971872651984
x_5	0.99999998549667
x_6	1.0000000000000000

Konvergerer fort dette her. \triangle

Som du ser i eksemplet over, doubles antall korrekte desimaler etter hver iterasjon. For å analysere konvergensen til Newtons metode må vi se den som en fikspunktiterasjon, der

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

Vi beregner

$$g'(x) = 1 - \frac{(f'(x))^2 - f(x)f''(x)}{(f'(x))^2} = \frac{f(x)f''(x)}{(f'(x))^2}$$

Husk at $f(r) = 0$. Dersom $f'(r) \neq 0$, får vi

$$g'(r) = \frac{f(r)f''(r)}{(f'(r))^2} = 0,$$

slik at

$$\begin{aligned} x_{n+1} - r &= g(x_n) - g(r) \\ &= \frac{g''(r)}{2}(x_n - r)^2 + \frac{g'''(r)}{3!}(x_n - r)^3 + \dots \end{aligned}$$

Dette kalles kvadratisk konvergens - feilen etter $n+1$ iterasjoner er proporsjonal med kvadratet av feilen etter n iterasjoner. Da konvergerer det fort.

Eksempel 5.8. Nå prøver vi å finne løsningene til

$$x^4 - 9x^3 + 27x^2 - 31x + 12 = 0.$$

Nok en gang er det et lettfaktorisert polynom:

$$x^4 - 9x^3 + 27x^2 - 31x + 12 = (x-1)^2(x-3)(x-4)$$

Newtons metode blir:

$$x_{n+1} = x_n - \frac{x_n^4 - 9x_n^3 + 27x_n^2 - 31x_n + 12}{4x_n^3 - 27x_n^2 + 54x_n - 31}$$

Nok en gang leter vi etter $r = 1$, ved å starte i $x_0 = 0.5$:

x_0	0.5000000000000000
x_1	0.713414634146341
x_2	0.842942878437971
x_3	0.916937117337936
x_4	0.957125910632705
x_5	0.978193460613942
x_6	0.988999465124112
x_7	0.994474755305802
x_8	0.997231047313292
x_9	0.998613930094898
x_{10}	0.999306565270834

Det ser ut til å konvergere, men mye saktere enn i sted. Hva skjedde? \triangle

I eksemplet over er $f'(1) = 0$. Dette betyr at

$$g'(1) = -\frac{f(1)f''(1)}{(f'(1))^2}$$

ikke er bestemt, og grensen

$$\lim_{x \rightarrow 1} g'(x)$$

trenger ikke være null. Eksemplet demonstrerer tydelig at den kvadratiske konvergensen ikke kan garanteres dersom $f'(r) = 0$.

Sekantmetoden

Denne ligner på Newton, men vi slår sekanten til f gjennom iterasjonene x_n og x_{i-1} , istedet for tangenten i x_i . Punktet der sekanten skjærer x -aksen er den nye iterasjonen x_n . Iterasjonen blir

$$x_{n+1} = x_n - f(x_n) \frac{x_n - x_{i-1}}{f(x_n) - f(x_{i-1})}.$$

Merk at her må man ha to initialgjetninger x_1 og x_0 , siden man trenger to punkter for å slå en sekant.

Når det gjelder sekantmetoden, kan det vises at

$$x_{n+1} - r = C(x_n - r)^{1.62},$$

men det er et svineri, du må taylorutvikle noe ut av det hinsidige. Konvergens kalles superlineær, altså kjappere enn fikspunkt, men treigere enn Newton.

Newtons metode i to dimensjoner

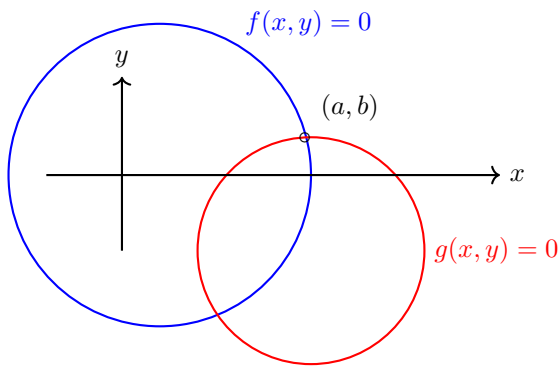
Anta at vi har to funksjoner $f(x, y)$ og $g(x, y)$. Vi leter etter et punkt (a, b) slik at både

$$f(a, b) = 0 \quad \text{og} \quad g(a, b) = 0.$$

likningene

$$f(x, y) = 0 \quad \text{og} \quad g(x, y) = 0$$

angir nivåkurver for funksjonene f og g . Punktet (a, b) må ligge på skjæringspunktet mellom disse.



Anta at du har en iterasjon (x_n, y_n) . Vi setter opp tangentplanene til f og g i (x_n, y_n)

$$z - f(x_n, y_n) = f_x(x_n, y_n)(x - x_n) + f_y(x_n, y_n)(y - y_n)$$

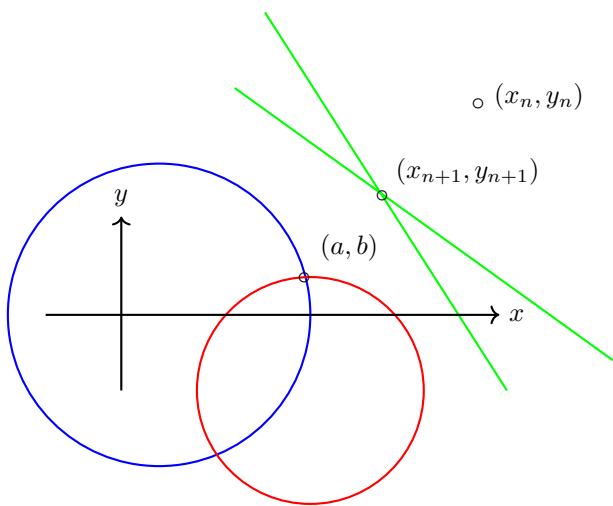
$$z - g(x_n, y_n) = g_x(x_n, y_n)(x - x_n) + g_y(x_n, y_n)(y - y_n).$$

Hvis vi krever at $z = 0$, får vi likninger for skjæringslinjene mellom disse tangentplanene og (x, y) -planet

$$-f(x_n, y_n) = f_x(x_n, y_n)(x - x_n) + f_y(x_n, y_n)(y - y_n)$$

$$-g(x_n, y_n) = g_x(x_n, y_n)(x - x_n) + g_y(x_n, y_n)(y - y_n).$$

Iterasjonen (x_{n+1}, y_{n+1}) defineres som skjæringen mellom disse linjene.



Altså er

$$-f(x_n, y_n) = f_x(x_n, y_n)(x_{n+1} - x_n) + f_y(x_n, y_n)(y_{n+1} - y_n)$$

$$-g(x_n, y_n) = g_x(x_n, y_n)(x_{n+1} - x_n) + g_y(x_n, y_n)(y_{n+1} - y_n),$$

et lineært likningssystem som definerer (x_{n+1}, y_{n+1}) . Matrisen er

$$\begin{pmatrix} f_x(x_n, y_n) & f_y(x_n, y_n) \\ g_x(x_n, y_n) & g_y(x_n, y_n) \end{pmatrix}$$

og vi skriver

$$-\begin{pmatrix} f(x_n, y_n) \\ g(x_n, y_n) \end{pmatrix} = \begin{pmatrix} f_x(x_n, y_n) & f_y(x_n, y_n) \\ g_x(x_n, y_n) & g_y(x_n, y_n) \end{pmatrix} \begin{pmatrix} x_{n+1} - x_n \\ y_{n+1} - y_n \end{pmatrix}.$$

Nå kan vi gange med

$$\begin{pmatrix} f_x(x_n, y_n) & f_y(x_n, y_n) \\ g_x(x_n, y_n) & g_y(x_n, y_n) \end{pmatrix}^{-1}$$

fra venstre, legge til (x_n, y_n) på begge sider, og få Newtons metode for systemer

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} - \begin{pmatrix} f_x(x_n, y_n) & f_y(x_n, y_n) \\ g_x(x_n, y_n) & g_y(x_n, y_n) \end{pmatrix}^{-1} \begin{pmatrix} f(x_n, y_n) \\ g(x_n, y_n) \end{pmatrix}.$$

Jeg gadd ikke skrive ut hva

$$\begin{pmatrix} f_x(x_n, y_n) & f_y(x_n, y_n) \\ g_x(x_n, y_n) & g_y(x_n, y_n) \end{pmatrix}^{-1}$$

er, men du kan regne det ut ved å huske fra lineær-algebraen at

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Eksempel 5.9. Vi leter etter løsninger til systemet

$$\begin{aligned} x^2 + y^2 &= 1 \\ \frac{x^2}{4} + 2y^2 &= 1 \end{aligned}$$

Som du husker fra M2 er den første likningen for enhetssirkelen, mens den andre er likningen for en ellipse med halvaksler 2 og $\frac{1}{2}$. Trekker man to ganger den første fra den andre, kan man regne ut at skjæringspunktet mellom disse kurvene i første kvadrant er

$$\left(\frac{2}{\sqrt{7}}, \sqrt{\frac{3}{7}} \right) \approx (0.755928946018454, 0.654653670707977).$$

La oss se om Newtons metode finner dette punktet. Siden

$$f(x, y) = x^2 + y^2 - 1$$

og

$$g(x, y) = \frac{x^2}{4} + 2y^2 - 1$$

blir

$$\begin{pmatrix} f_x(x_n, y_n) & f_y(x_n, y_n) \\ g_x(x_n, y_n) & g_y(x_n, y_n) \end{pmatrix} = \begin{pmatrix} 2x & 2y \\ \frac{x}{2} & 4y \end{pmatrix}$$

som har invers

$$\frac{1}{7xy} \begin{pmatrix} 4y & -2y \\ -\frac{x}{2} & 2x \end{pmatrix}.$$

Metoden blir

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} - \frac{1}{7xy} \begin{pmatrix} 4y & -2y \\ -\frac{x}{2} & 2x \end{pmatrix} \begin{pmatrix} f(x_n, y_n) \\ g(x_n, y_n) \end{pmatrix},$$

og starter vi i $(1, 1)$, får vi:

n	x_n	y_n
1	1.000000000000000	1.000000000000000
2	0.785714285714286	0.714285714285714
3	0.756493506493507	0.657142857142857
4	0.755929156680230	0.654658385093168
5	0.755928946018484	0.654653670724952
6	0.755928946018454	0.654653670707977
7	0.755928946018455	0.654653670707977

Maskinpresisjon etter syv iterasjoner. Bra greier det her. \triangle

6 Numerisk derivasjon

Hva om du ønsker å finne stigningen til en funksjon som er tabulert? Vi skal gå gjennom en sentral teknikk for numerisk derivasjon som kalles *endelige differanser*. Det er to grunner til dette: for det første trenger vi det for å lage numeriske skjema for differensiallikninger, og for det andre gir det en fryktelig transparent illustrasjon av dette med feil.

Skanter og tangenter

I M1 definerte vi den deriverte til en funksjon f som

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}.$$

Uttrykket

$$\frac{f(x+h) - f(x)}{h}$$

er stigningstallet til sekanten til f mellom punktene x og $x+h$. For små h er denne sekanten en grei tilnærming til stigningstallet $f'(x)$:

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}.$$

På papiret er det slik at jo mindre h , desto bedre tilnærming.

Eksempel 6.1. La $f(x) = e^x$, og la $h = 0.1$. Vi beregner

$$f'(1.5) \approx \frac{e^{1.6} - e^{1.5}}{0.1} = 4.7134.$$

Merk at

$$f'(1.5) = e^{1.5} = 4.4817,$$

så denne tilnærmingen bommer med rundt $2 \cdot 10^{-1}$. Vi kan også prøve $h = 0.01$. Da får vi

$$f'(1.5) \approx \frac{e^{1.51} - e^{1.5}}{0.01} = 4.5042.$$

som er noe bedre, nå er feilen på rundt $2 \cdot 10^{-2}$. Vi knekker til med enda en:

$$f'(1.5) \approx \frac{e^{1.501} - e^{1.5}}{0.001} = 4.4839,$$

og får en feil på rundt $2 \cdot 10^{-3}$. \triangle

Merk. Feilen i forrige eksempel er tydelig proporsjonal med h - deler du h på 10, deler du feilen på 10. En god illustrasjon av lineær feil.

Taylorutvikling

I M1 lærte du også taylorutvikling. Anta f er analytisk funksjon. Det går an å utlede formelen

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}.$$

ved å bruke taylorutviklingen til f i punktet x :

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2}h^2 + \dots$$

Stigningen til sekanten kan skrives

$$\frac{f(x+h) - f(x)}{h} = f'(x) + \frac{f''(x)}{2}h + \dots$$

og vi ser at dette stigningstallet består av den eksakte verdien for $f'(x)$ pluss resten av taylorrekken til f

$$\frac{f''(x)}{2}h + \frac{f'''(x)}{6}h^2 + \dots$$

Denne halen forteller oss noe om feilen. Dersom h er liten nok, vil h være mye større enn h^2 , og vi skriver

$$\frac{f''(x)}{2}h + \frac{f'''(x)}{6}h^2 + \dots = O(h)$$

for å signalisere at feilen er proporsjonal med h .

Høyere ordens tilnærminger

Vi kan relativt lett forbedre den lineære tilnærmingen fra forrige avsnitt ved å skrive

$$f(x-h) = f(x) - f'(x)h + \frac{f''(x)}{2}h^2 + \dots$$

og sette opp tilnærmingen

$$\frac{f(x+h) - f(x-h)}{2h} = f'(x) + \frac{f'''(x)}{3!}h^2 + \frac{f^{(5)}(x)}{5!}h^4 + \dots$$

Feilen i denne tilnærmingen er

$$\frac{f'''(x)}{3!}h^2 + \frac{f^{(5)}(x)}{5!}h^4 + \dots$$

Hvis h er liten, er det rimelig å anta denne feilen er mye mindre enn for den første tilnærmingen, siden $h^2 \ll h$.

Eksempel 6.2. La $f(x) = e^x$, og la $h = 0.1$. Vi beregner

$$f'(1.5) \approx \frac{e^{1.6} - e^{1.4}}{0.2} = 4.489162287752202,$$

som gir en feil på rundt $7.5 \cdot 10^{-3}$. Mye bedre enn i sted. Vi kan prøver også $h = 0.01$:

$$f'(1.5) \approx \frac{e^{1.51} - e^{1.49}}{0.02} = 4.481763765529401$$

som er noe bedre, nå er feilen på rundt $7.5 \cdot 10^{-5}$. Knekker vi til med $h = 0.001$, får vi

$$f'(1.5) \approx \frac{e^{1.501} - e^{1.499}}{0.002} = 4.481689817286139$$

som gir en feil på $7.5 \cdot 10^{-7}$. \triangle

Merk. Legg merke til hvordan feilen i forrige eksempel er nærmest perfekt kvadratisk - vi får to nye desimaler hver gang vi deler h på 10. Feilen deles altså på 100 når h deles på 10.

Hvis du virkelig vil slå på stortrommen, kan du bruke formelen

$$f'(x) = \frac{f(x-2h) - 8f(x-h) + 8f(x+h) - f(x+2h)}{12h}$$

Eksempel 6.3. Vi bruker den store formelen med $h = 0.1$, $h = 0.01$ og $h = 0.001$. Då får vi feil på 10^{-5} , 10^{-9} , og 10^{-13} . Fire nye desimaler hver gang h deles på 10. Prøv selv. \triangle

Høyere ordens deriverte

Dersom du trenger en tilnærming for $f''(x)$, kan du bruke den andre ordens sentraldifferansen

$$\frac{f(x+h) - 2f(x) + f(x-h)}{h^2} = f''(x) + O(h^2)$$

Nå bør det etterhvert være klart hvordan en slik derivasjonsformel konstrueres - man søker en lineærkombinasjon av funksjonsverdier $f(x)$, $f(x-h)$ og $f(x+h)$ og lignende ledd, for å oppnå to ting:

- Korrekt tilnærming av den n -te deriverte.
- Så høy orden som mulig.

Richardsonekstrapolasjon

Det går an å kombinere tilnærminger av forskjellig orden til å oppnå høyere ordens tilnærminger. Vi definerer

$$\phi(h) = \frac{f(x+h) - f(x-h)}{2h}.$$

Da har vi at

$$\phi(h) = f'(x) + h^2 \frac{f'''(x)}{6} + h^4 \frac{f^{(5)}(x)}{120} + \dots$$

og

$$\phi\left(\frac{h}{2}\right) = f'(x) + \left(\frac{h}{2}\right)^2 \frac{f'''(x)}{6} + \left(\frac{h}{2}\right)^4 \frac{f^{(5)}(x)}{120} + \dots$$

Her er trikset.

$$\frac{4\phi\left(\frac{h}{2}\right) - \phi(h)}{3} = f'(x) - \left(\frac{h}{2}\right)^4 \frac{f^{(5)}(x)}{480} + \dots$$

Med andre ord: den rette lineærkombinasjonen av to estimater med forskjellige gitterfinheter kan få et (eller flere) ledd i feilutviklingen til å forsvinne, og da får vi en høyere ordens tilnærming.

Eksempel 6.4. Hvis vi setter $h = 0.1$, og tar to tidligere approksimasjoner,

$$\phi(0.1) = \frac{e^{1.6} - e^{1.4}}{0.2} = 4.489162287752202,$$

med en feil på rundt $7.5 \cdot 10^{-3}$, og

$$\phi(0.01) = \frac{e^{1.51} - e^{1.49}}{0.02} = 4.481763765529401$$

med en feil på rundt $7.5 \cdot 10^{-5}$, og og kombinerer dem, får vi

$$\frac{100\phi\left(\frac{h}{10}\right) - \phi(h)}{99} = 4.481689032981695$$

som gir en feil på $-3.73 \cdot 10^{-8} \approx \left(\frac{h}{10}\right)^4$. \triangle

Det går an å formulere presise teoremer som forteller hvordan man skal lineærkombinere tilnærminger til høyere ordens tilnærminger, men vi nøyer oss med denne lille smakebiten.

7 Interpolasjon

Fra M3 husker du at dersom x_i er $n + 1$ forskjellige punkter på x -aksen med korresponderende y -verdier y_i , finnes det et entydig polynom av maksimal grad n som interpolerer punktene (x_i, y_i) . I dette kapitlet skal vi sette opp to forskjellige formler for dette polynomet, og ta en litt grundigere analyse. Til slutt skal vi også ta en titt på hvordan dette kan gjøres med trigonometriske funksjoner.

Lagranges interpolasjon

La x_i være $n + 1$ forskjellige punkter på intervallet $[a, b]$, med $x_0 = a$ og $x_n = b$. For hvert punkt x_i , definerer vi et polynom:

$$l_i(x) = \prod_{\substack{k=0 \\ k \neq i}}^n \frac{(x - x_k)}{(x_i - x_k)}.$$

Polynomet $l_i(x)$ har orden n , og tilfredsstill

$$l_i(x_k) = \begin{cases} 1 & \text{for } i = k \\ 0 & \text{for } i \neq k \end{cases}.$$

La f være en funksjon, med funksjonsverdier $f(x_i) = f_i$. Det er lett å se at

$$p_n(x) = \sum_{i=0}^n f_i l_i(x).$$

tilfredsstill $p_n(x_i) = f(x_i)$ for alle i .

Teorem 7.1. La x_i være $n + 1$ forskjellige punkter på intervallet $[a, b]$, og f en funksjon med funksjonsverdier $f(x_i) = f_i$. Det finnes et entydig polynom som tilfredsstill $p_n(x_i) = f(x_i)$ for alle i .

Bevis. Konstruksjonen av Lagranges interpolasjon viser at det for en tabell med $n + 1$ punkter eksisterer et interpolasjonspolynom av maksimal grad n ; vi har jo nettopp konstruert det. Hvis vi antar at det finnes to forskjellige polynomer p_n og q_n av grad n som interpolerer den samme tabellen, og evaluerer differansen $p_n - q_n$ i punktene x_i , ser vi at

$$p_n(x_i) - q_n(x_i) = 0 \quad 0 \leq i \leq n.$$

Men polynomet $p - q$ har maksimal grad n , og kan følgelig ha maksimalt n nullpunkter, så den eneste muligheten her er $p = q$, som betyr at interpolasjonspolynomet er entydig. \square

Eksempel 7.2. En funksjon har følgende verdier:

i	0	1	2
x_i	1	2	3
f_i	4	5	6

Vi setter opp

$$l_0(x) = \frac{(x-2)(x-3)}{(1-2)(1-3)} = \frac{1}{2}(x-2)(x-3)$$

$$l_1(x) = \frac{(x-1)(x-3)}{(2-1)(2-3)} = -(x-1)(x-3)$$

og

$$l_2(x) = \frac{(x-1)(x-2)}{(3-1)(3-2)} = \frac{1}{2}(x-1)(x-2).$$

Det andre ordens polynomet som interpolerer denne tabellen er:

$$\begin{aligned} p(x) &= 4l_0(x) + 5l_1(x) + 6l_2(x) \\ &= 2(x-2)(x-3) - 5(x-1)(x-3) \\ &\quad + 3(x-1)(x-2). \end{aligned} \quad \triangle$$

I gamle dager sto det i numerikkbøkene at lagrangepolynomene ikke måtte brukes i numeriske beregninger, fordi det koster for mange flyttalsoperasjoner å evaluere dem. Dette er bare tull dersom man bruker de rette interpolasjonspunktene og evaluerer polynomene på rett måte, men dette er dessverre utenfor vårt pensum.

Newtons interpolasjon

For å konstruere Newtons interpolasjon trenger vi å beregne de *dividerte differansene*. De defineres rekursivt:

$$f[x_i] = f(x_i)$$

$$f[x_i, x_{i-1}] = \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}$$

$$f[x_i, \dots, x_{i-k}] = \frac{f[x_i, \dots, x_{i-k+1}] - f[x_{i-1}, \dots, x_{i-k}]}{x_i - x_{i-k}}$$

Newtons interpolasjonspolynom er:

$$p_n(x) = f_0 + \sum_{i=1}^n f[x_i, \dots, x_0] \prod_{k=0}^{i-1} (x - x_k).$$

Merk også at Lagranges og Newtons interpolasjon er bare to forskjellige formler for å sette opp det samme polynomet, siden interpolasjonspolynomet er entydig.

Eksempel 7.3. Polynomet

$$\begin{aligned} p_2(x) &= \\ &= f_0 + \frac{f_1 - f_0}{x_1 - x_0} (x - x_0) \\ &\quad + \frac{\frac{f_2 - f_1}{x_2 - x_1} - \frac{f_1 - f_0}{x_1 - x_0}}{x_2 - x_0} (x - x_0)(x - x_1) \end{aligned}$$

interpolerer tabellen

x_0	x_1	x_2
f_0	f_1	f_2

så lenge $x_0 \neq x_1 \neq x_2$. \triangle

Det er vanlig å sette opp følgende tableau for å illustrere de dividerte differansene:

x_0	$f[x_0]$			
	$f[x_0, x_1]$			
x_1	$f[x_1]$	$f[x_0, x_1, x_2]$		
	$f[x_1, x_2]$		$f[x_0, x_1, x_2, x_3]$	
x_2	$f[x_2]$	$f[x_1, x_2, x_3]$		
	$f[x_2, x_3]$			
x_3	$f[x_3]$			

Eksempel 7.4. La igjen

i	0	1	2
x_i	1	2	3
f_i	4	5	6

Tableauet blir

1	4		
		1	
2	5		0
		1	
3	6		

og interpolasjonspolynomet blir

$$p(x) = 4 + x - 1 = 3 + x.$$

Dette er selvfølgelig det samme polynomet som i forrige eksempel. Merk at polynomet er av første grad, siden punktene tilfeldigvis ligger på en rett linje. \triangle

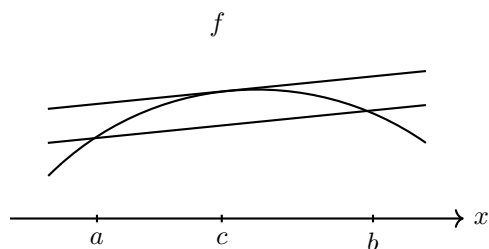
Interpolasjonsfeilen

Det sier seg selv at et interpolasjonspolynom ikke kan være lik funksjonen som interpoleres med mindre denne funksjonen er et polynom av ikke høyere grad enn interpolanten. Nå er vi kommet til steg to i den generelle oppskriften for numeriske metoder, nemlig analyse av feilen. Da må vi begynne med en generalisering av middelverdisatsen.

Middelverdisatsen sier at

$$f'(c) = \frac{f(b) - f(a)}{b - a} = f[a, b]$$

for en funksjon som er deriverbar på $[a, b]$.



En generalisert variant for dividerte differenser går som følger.

Teorem 7.5. Dersom f er $n + 1$ ganger deriverbar på $[a, b]$, og alle punktene x_i er forskjellige, er

$$f[x_n, \dots, x_0] = \frac{f^n(s)}{n!}$$

for en eller annen s i intervallet $[a, b]$.

Bevis. Siden p_n interpolerer f , må funksjonen $g = f - p_n$ ha minst $n + 1$ nullpunkt på intervallet $[a, b]$. Gjentatt anvendelse av middelverdisatsen forteller oss at g' har minst n nullpunkt, at g'' har minst $n - 1$ nullpunkt, og videre at g^{n+1} har minst ett nullpunkt på $[a, b]$. Vi kaller dette s . Siden

$$\frac{d^n}{dx^n} p_n(x) = n! f[x_n, \dots, x_0],$$

må

$$f^n(s) = n! f[x_n, \dots, x_0]. \quad \square$$

Vi kan nå utlede et uttrykk for interpolasjonsfeilen.

Teorem 7.6. La f være en $n + 1$ ganger deriverbar funksjon på $[a, b]$, interpolert i punktene x_i . Interpolasjonsfeilen er

$$f(x) - p_n(x) = \frac{f^{n+1}(s)}{(n+1)!} \prod_{k=0}^n (x - x_k).$$

Bevis. Vi skriver opp Newtons interpolasjonspolynom

$$p_n(x) = f_0 + \sum_{i=1}^{n-1} f[x_i, \dots, x_0] \prod_{k=0}^{i-1} (x - x_k) + f[x_n, \dots, x_0] \prod_{k=0}^{n-1} (x - x_k),$$

det siste leddet er skrevet ut kun av pedagogiske hensyn. Nå bytter vi ut x_n med x i uttrykket over (tenk på x som et nytt interpolasjonspunkt), og får

$$f(x) = f_0 + \sum_{i=1}^{n-1} f[x_i, \dots, x_0] \prod_{k=0}^{i-1} (x - x_k) + f[x, \dots, x_0] \prod_{k=0}^{n-1} (x - x_k).$$

Merk den snedige måten å skrive om f på. Trekker vi de to foregående uttrykkene fra hverandre, får vi

$$\begin{aligned} f(x) - p_n(x) &= (f[x, x_{n-1}, \dots, x_0] - f[x_n, x_{n-1}, \dots, x_0]) \prod_{k=0}^{n-1} (x - x_k) \\ &= f[x, x_n, \dots, x_0] (x - x_n) \prod_{k=0}^{n-1} (x - x_k) \\ &= f[x, x_n, \dots, x_0] \prod_{k=0}^n (x - x_k), \end{aligned}$$

og bruker vi den generaliserte middelverdisatsen over, får vi

$$f(x) - p_n(x) = \frac{f^{n+1}(s)}{(n+1)!} \prod_{k=0}^n (x - x_k),$$

for en eller annen $s \in [a, b]$. Merk at s avhenger av x , akkurat som i Taylors teorem fra M1. \square

Punktfordeling

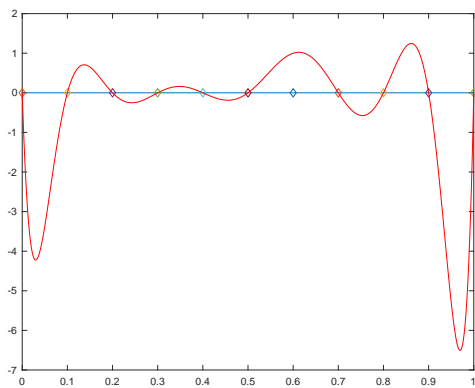
Steg tre i oppskriften på numeriskemetoder, er å finne ut om metoden kan ha noen gode egenskaper utover høy presisjon. I interpolasjonsfaget er det en relativt kjapp måte å avgjøre om en interpolasjonsmetode er god eller ikke: Det er avgjørende for kvaliteten på interpolasjonen at punktene står riktig fordelt på intervallet $[a, b]$.

Men hvordan skal vi finne gode punkter for polynominterpolasjon, og hva skiller de gode fra de dårlige punktene? Dette spørsmålet kan besvares på mange måter, og vi skal vende tilbake til spørsmålet i kapitlet om numerisk integrasjon. La oss begynne med å ta for oss en god og en dårlig punktmengde.

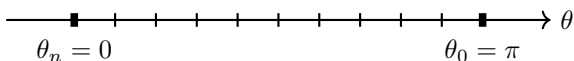
Eksempel 7.7. Den dårlige punktmengden er den kjente og kjære *ekvidistante* punktmengden. Et ekvidistant gitter med n punkter på intervallet $[a, b]$ er gitt ved

$$x_i = a + hi$$

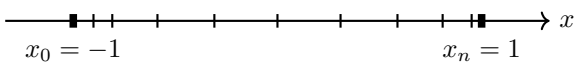
der $0 \leq i \leq n$ og $h = (b - a)/n$. I figuren under er et plot av en lagrangefunksjon på et ekvidistant gitter med elleve punkt. Merk oscillasjonene polynomet gjør mellom interpolasjonspunktene. \triangle



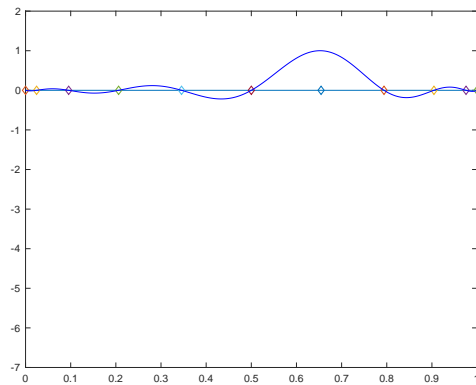
Eksempel 7.8. La θ_i være et ekvidistant gitter på $[0, \pi]$, med $\theta_n = 0$ og $\theta_0 = \pi$.



Nå definerer vi $x_i = \cos \theta_i$. Dette gitteret ligger på $[-1, 1]$.



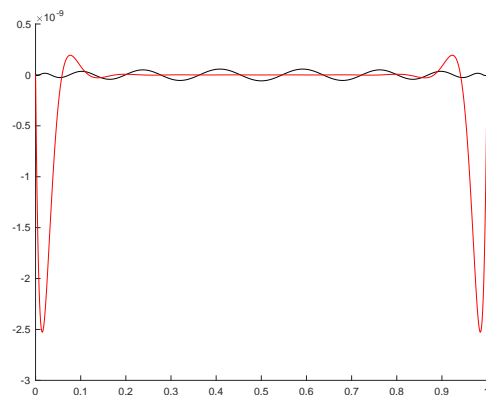
Det gitteret egner seg skikkelig godt for polynominterpolasjon. Under er et tilsvarende plot av en lagrangefunksjon på dette gitteret. Merk hvordan interpolasjonspolynomet ikke oscillerer særlig mellom interpolasjonspunktene. \triangle



Man finner gode punkter for interpolasjon ved å plassere dem slik at de minimerer utslaget til polynomet

$$\prod_{k=0}^n (x - x_k)$$

i interpolasjonsfeilen, og alle gode punktmengder for polynominterpolasjon klumper seg i endene av intervallet. For ekvidistante gitre har feilpolynomet stort utslag i endepunktene, og det forklarer hvorfor lagrangefunksjonen i figuren over har størst utslag fra funksjonsverdiene der. Vi skal nå ta for oss noen forskjellige gode punktmengder.



Vi skal gjøre alt på intervallet $[-1, 1]$. Dersom man har en punktfordeling x_i på dette intervallet, kan man flytte fordelingen til intervallet $[a, b]$ med formelen gitt i eksempel 7.11 under.

Chebyshevpunkter

For å prøve å forklare hva som skjedde i figuren må vi introdusere noen polynomer.

Teorem 7.9. Funksjonen

$$T_n(x) = \cos(n \arccos x)$$

er et polynom når n er et naturlig tall.

Bevis. La $T_n(\theta) = \cos(n\theta)$. Da har vi

$$\begin{aligned} T_0 &= 1 \\ T_1 &= \cos \theta \\ T_2 &= \cos 2\theta = 2 \cos^2 \theta - 1 \end{aligned}$$

Merk at alle disse er polynomer i $\cos \theta$. Dette er en sentral observasjon, for når man gjør variabelskiftet $x = \cos \theta$ vil man få polynomer i x .

Vi fortsetter med et induksjonsbevis for at $\cos n\theta$ er et polynom i $\cos \theta$ for alle naturlige tall n . Legg sammen

$$\cos(n+1)\theta = \cos n\theta \cos \theta - \sin n\theta \sin \theta$$

og

$$\cos(n-1)\theta = \cos n\theta \cos \theta + \sin n\theta \sin \theta$$

slik at

$$\cos(n+1)\theta = 2 \cos n\theta \cos \theta - \cos(n-1)\theta.$$

Dersom vi antar at $\cos n\theta$ og $\cos(n-1)\theta$ er polynomer i $\cos \theta$, må

$$2 \cos n\theta \cos \theta - \cos(n-1)\theta$$

være et polynom i $\cos \theta$, og følgelig må også

$$\cos(n+1)\theta$$

være et polynom i $\cos \theta$. Siden $T_0 = 1$ og $T_1 = \cos \theta$ (og $T_2 = 2 \cos^2 \theta - 1$) er polynomer i $\cos \theta$, må $\cos n\theta$ være det for alle naturlige n . Dersom $\cos n\theta$ er et polynom i $\cos \theta$, må $\cos(n \arccos x)$ være et polynom i x . \square

Polynomene T_n kalles *Chebyshevpolynomer*. Disse kan beregnes ved rekursjonen

$$T_{n+1} = 2xT_n - T_{n-1},$$

som er likningen

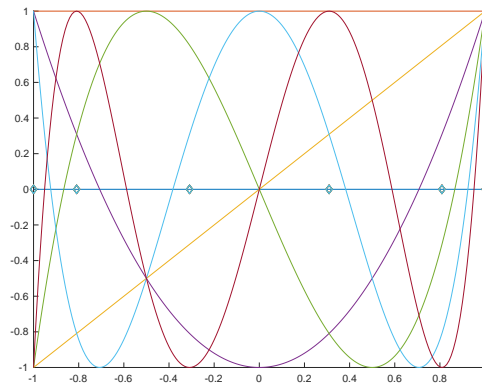
$$\cos(n+1)\theta + \cos(n-1)\theta = 2 \cos n\theta \cos \theta.$$

fra forrige bevis, skrevet ut i form av T_n . De første er

$$\begin{aligned} T_0 &= 1 \\ T_1 &= x \\ T_2 &= 2x^2 - 1 \\ T_3 &= 4x^3 - 3x \\ T_4 &= 8x^4 - 8x^2 + 1 \\ &\vdots \end{aligned}$$

Gitteret i eksempel 7.8 kalles Chebyshevs ekstremal-gitter, for punktene er ekstremalpunktene til et chebyshevpolynom. Det n -te ordens polynom T_n gir opphav til et gitter med $n+1$ punkter, der $n-1$ av dem er T_n s stasjonære punkter, og to er endepunktene i intervallet. De første seks polynomene er plottet under, sammen med 6-punktsgitteret som er ekstremalpunktene til T_5 . En formel for gitterpunktene er:

$$x_i = \cos \frac{\pi i}{n} \quad 0 \leq i \leq n.$$



Eksempel 7.10. Gitteret i figuren over er gitt ved

x_0	-1.0000000000000000
x_1	-0.809016994374947
x_2	-0.309016994374947
x_3	0.309016994374947
x_4	0.809016994374947
x_5	1.0000000000000000

Merk at endepunktene er en annen type ekstremal-punkter enn de andre; det er der T_5 er klippet av intervallgrensene. \triangle

Eksempel 7.11. Hvis vi ønsker å sette opp gitteret fra T_5 på et intervall $[a, b]$, bruker vi bare formelen

$$y_i = a + (b-a) \frac{x_i + 1}{2},$$

der x_i er punktene på $[-1, 1]$ og y_i er punktene på $[a, b]$. På intervallet $[1, 4]$ blir punktene

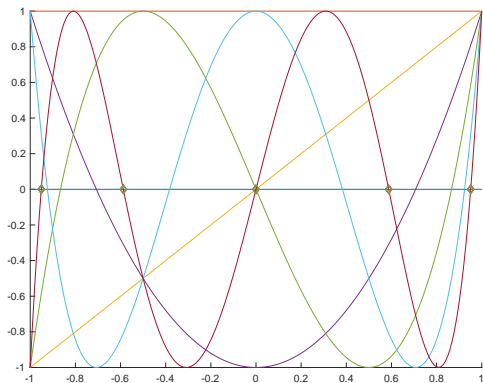
x_0	1.0000000000000000
x_1	1.286474508437579
x_2	2.036474508437579
x_3	2.963525491562421
x_4	3.713525491562421
x_5	4.0000000000000000

Det er viktig å forstå at det hvordan punktene er fordelt på intervallet som har noe å si for kvaliteten på interpolasjonen. \triangle

Chebyshevpolynomenes nullpunkter gir opphav til en annen punktmengde som kalles *Chebyshevs nullpunktgitter*. Dette gitteret er på papiret enda bedre enn Chebyshevs ekstremalgitter, men noe mindre praktisk, siden det ikke inneholder endepunktene. Polynom T_{n+1} gir opphav til et gitter med $n+1$ punkter, gitt ved

$$x_i = \cos \frac{\pi(2i+1)}{2n+2} \quad 0 \leq i \leq n.$$

Under er et nok et plot av de første par chebyshevpolynomene, med nullpunktgitteret til T_5 . Vi tar med et teorem om interpolasjonsfeilen til Chebyshevs nullpunktgitter, men lar det stå ubevist.



Teorem 7.12. La f være en $n + 1$ ganger deriverbar funksjon. Interpolasjonsfeil for interpolanten på chebyshevs nullpunktgitter på $[a, b]$ er:

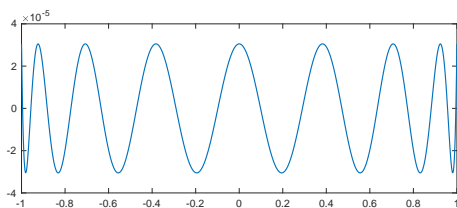
$$\max_{x \in [a, b]} |f(x) - p_n(x)| \leq \frac{(b-a)^{n+1}}{2^{2n+1}} \max_{x \in [a, b]} |f^{n+1}(x)|$$

Under er et plot av feilpolynomet

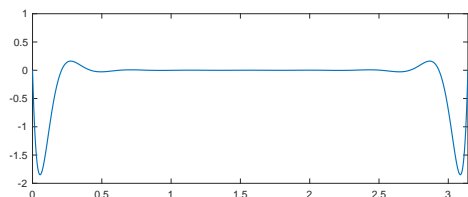
$$\prod_{k=0}^n (x - x_k)$$

for chebyshevs nullpunktgitter og $n = 16$. Dette polynomet har maksimalt utslag

$$\frac{(b-a)^{n+1}}{2^{2n+1}} \approx 1.525878906250000e - 05.$$



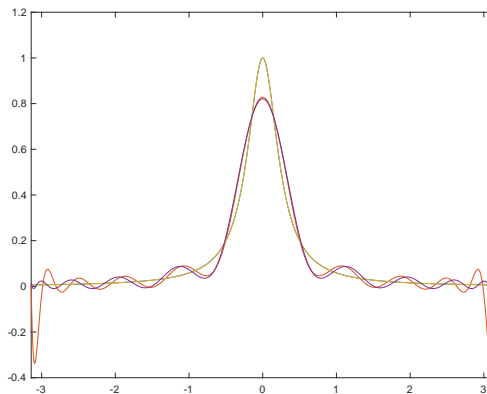
Vi tar med et plot av tilsvarende feilpolynom for ekvidistant gitter med $n = 16$.



Eksempel 7.13. Nedenfor er en figur av 20. ordens interpolanter av Runges funksjon

$$f(x) = \frac{1}{1 + 16x^2}$$

på ekvidistant og chebyshev-gitter. Denne funksjonen er kjent for sine patologisk store n -te deriverte. Dette er en størrelse vi ikke har kontroll på, merk nok en gang hvordan feilpolynomet illustrerer hvorfor ekvidistant tinærmer vesentlig dårligere enn chebyshev i endene. \triangle



Gauss-punkter

En punktmengde som likner på Chebyshev, og klumper seg i endene av intervallet, kalles Gauss-punkter. Akkurat som for Chebyshev, kommer disse i to varianter, som er henholdsvis null- og ekstremalpunkter til en følge av polynomer, som kalles Gauss-Legendre-polynomene. Legendre-polynomene er gitt ved rekursjonen

$$P_0 = 1$$

$$P_1 = x$$

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x).$$

Vi skal ikke utlede disse, men i kapitlet om numerisk integrasjon skal vi vise en metode for å produsere nullpunktene.

Nullpunktgitrene til Legendre-polynomene kalles Gauss-Legendre-punkter. Her er en tabell med et par lavere ordens punktfordelinger på intervallet $[-1, 1]$

n	x_i
2	$\pm \sqrt{\frac{1}{3}}$
3	$0, \pm \sqrt{\frac{3}{5}}$
4	$\pm \sqrt{\frac{3}{7} \pm \frac{2}{7} \sqrt{\frac{6}{5}}}$
5	$0, \pm \frac{1}{3} \sqrt{5 \pm 2 \sqrt{\frac{10}{7}}}$

Gauss-Legendre-punktene er på papiret enda bedre enn chebyshev-punktene, men mindre praktisk i bruk.

Ekstremalpunktene til Legendre-polynomene kalles Gauss-Lobatto-punkter. Vi hoster opp nok en tabell med et par lavere ordens punktfordelinger på intervallet $[-1, 1]$

n	x_i
3	$0, \pm 1$
4	$\pm \sqrt{\frac{1}{5}}, \pm 1$
5	$0, \pm \sqrt{\frac{3}{7}}, \pm 1$
6	$\pm \sqrt{\frac{1}{3}} \pm \sqrt{\frac{2}{3\sqrt{7}}}, \pm 1$
7	$0, \pm \sqrt{\frac{5}{11}} \pm \frac{2}{11} \sqrt{\frac{5}{3}}, \pm 1$

Gauss-Lobatto er på papiret noe dårligere enn Gauss-Legendre, men er noe mer praktisk i bruk, for de inneholder intervallets endepunkter. Fremdeles mindre praktisk enn Chebyshev.

Andre typer polynominterpolasjon

Til slutt kan nevnes at man trenger ikke nøye seg med å kreve at interpolanten p skal ta f sine verdier i interpolasjonspunktene. Man kan også skru opp graden på polynomet, og i tillegg kreve at p skal ha samme stigningstall som f i punktene. I dette tilfellet kalles det *Hermite-interpolasjon*.

Dersom man krever at p skal ha samme verdi som f sine $n + 1$ første deriverte i et enkelt punkt, er vi tilkake i Taylorpolynomene du kjenner fra M1.

Man kan også, istedet for å lage et interpolasjonspolynom som interpolerer f i alle punktene, for eksempel sortere interpolasjonspunktene i grupper på fire og fire etterfølgende punkter, interpolere hver gruppe med et tredjeordens polynom, og så sette sammen en stykkvis kontinuerlig deriverbar polynominterpolant. Dette kalles *spline-interpolasjon*.

DFT - trigonometrisk interpolasjon

Vi skal ta en kjapp vending innom Fourieranalysen igjen. Fourieranalyse og Chebyshevinterpolasjon henger nemlig sammen, og for å se hvordan, må vi lære å interpolere med trigonometriske funksjoner.

I dette avsnittet skal vi jobbe med det ekvidistante gitteret

$$x_k = \frac{\pi k}{N} \quad \text{der} \quad -N \leq k \leq N-1$$

på intervallet $[-\pi, \pi]$. Vi ønsker å finne et trigonometrisk polynom

$$\sum_{n=-N}^N c_n e^{inx}$$

som interpolerer funksjonen f i gitterpunktene, altså at

$$f(x_k) = \sum_{n=-N}^N c_n e^{inx_k} = \sum_{n=-N}^N c_n e^{in \frac{\pi k}{N}}.$$

Merk. Du tenker kanskje at det hadde vært naturlig å ta med gitterpunktet $x = \pi$. Men det trigonometriske polynomet vi skal interpolere med, har fundamentalperiode 2π , så hvis vi bestemmer hva polynomet skal være i $x = -\pi$, bestemmer vi samtidig hva det skal være i $x = \pi$. Derfor er ikke $x = \pi$ med i listen over interpolasjonspunkter.

Teorem 7.14. Det trigonometriske polynomet

$$\sum_{n=-N}^N c_n e^{inx}$$

interpolerer f i punktene x_i dersom koeffisientene er gitt ved

$$c_n = \frac{1}{2N} \sum_{k=-N}^{N-1} f(x_k) e^{-inx}.$$

Bevis. Vi ganger ligningen over med e^{-imx} og summerer over alle gitterpunkter

$$\sum_{k=-N}^{N-1} f(x_k) e^{-imx} = \sum_{k=-N}^{N-1} \sum_{n=-N}^N c_n e^{i \frac{\pi k}{N} (n-m)}.$$

Nå bytter vi summasjonsrekkefølgen på høyre side, og får

$$\sum_{k=-N}^{N-1} f(x_k) e^{-imx} = \sum_{n=-N}^N c_n \sum_{k=-N}^{N-1} e^{i \frac{\pi k}{N} (n-m)}.$$

Merk at den innerste summen er en endelig geometrisk rekke med multiplikasjonsfaktor $e^{i \frac{\pi k}{N}}$. Dersom $n \neq m$ er

$$\begin{aligned} \sum_{k=-N}^{N-1} e^{i \frac{\pi k}{N} (n-m)} &= e^{i \frac{\pi k}{N} (-N-m)} \sum_{k=0}^{2N-1} e^{i \frac{\pi k}{N} (n-m)} \\ &= e^{i \frac{\pi k}{N} (-N-m)} \sum_{k=0}^{2N-1} \left(e^{i \frac{\pi}{N} (n-m)} \right)^k \\ &= e^{i \frac{\pi k}{N} (-N-m)} \frac{1 - \left(e^{i \frac{\pi}{N} (n-m)} \right)^{2N}}{1 - e^{i \frac{\pi}{N} (n-m)}} \\ &= e^{i \frac{\pi k}{N} (-N-m)} \frac{1 - e^{2\pi i (n-m)}}{1 - e^{i \frac{\pi}{N} (n-m)}} = 0, \end{aligned}$$

og dersom $n = m$, er

$$\sum_{k=-N}^{N-1} e^{i \frac{\pi k}{N} k(n-m)} = \sum_{k=-N}^{N-1} 1 = 2N,$$

slik at

$$\sum_{n=-N}^N c_n \sum_{k=-N}^{N-1} e^{\frac{2\pi i}{N} k(n-m)} = 2N c_n.$$

Med andre ord er

$$\begin{aligned} \sum_{k=-N}^{N-1} f(x_k) e^{-imx} &= \sum_{k=-N}^{N-1} \sum_{n=-N}^N c_n e^{i \frac{\pi k}{N} (n-m)} \\ &= \sum_{n=-N}^N c_n \sum_{k=-N}^{N-1} e^{\frac{2\pi i}{N} k(n-m)} \\ &= 2N c_n. \end{aligned}$$

slik at

$$c_n = \frac{1}{2N} \sum_{k=-N}^{N-1} f(x_k) e^{-inx}. \quad \square$$

Det finnes imidlertid en teoretisk enda enklere måte å skrive opp det trigonometriske interpolasjonspolynomet på.

Teorem 7.15. *Dirichletkjernen*

$$\sum_{n=-N}^N e^{in(x-x_k)}$$

tar verdien 1 i x_k og 0 i de andre gitterpunktene. Interpolanten kan skrives

$$\sum_{n=-N}^{N-1} f(x_k) D_n(x - x_k).$$

Merk at interpolasjonspolynomet er skrevet som en diskret konvolusjon. I kapitlene om fourierrekker og fouriertransform har vi skrevet funksjoner som trigonometriske rekker og integraler. Dette er en diskret variant av akkurat det samme, og kalles derfor *diskret fouriertransform*. Koeffisientene c_n kan beregnes ekstremt raskt med en berømt algoritme som går under navnet FFT - *fast fourier transform*. Både Python og Matlab har innebygde rutiner for dette.

I fouriertransformkapitlet var både transformasjonen og inverstransformasjonen gitt ved integraler. I fourierekkekapitlet var fourierkoeffisientene gitt ved integraler, og inverstransformer gitt ved en sum. I dette kapitlet var både transformasjon og inverstransformasjon gitt ved summer. Det finnes en fjerde variant der fourierkoeffisientene er gitt ved summer, mens inverstransformasjonen er et integral, men denne er utenfor vårt pensum.

8 Numerisk integrasjon

I dette kapitlet skal vi ta for oss noen vanlige numeriske integrasjonsmetoder. I M1 har du lært to av dem - trapesregelen og Simpsons metode. Vi skal raskt repetere disse, og så gå videre til mer spennende metoder. Alle er bygget på interpolasjonsteknikken vi lærte i forrige uke.

Kvadraturregler

Vi skal finne tilnærminger til integralet

$$I[f] = \int_a^b f(x) dx$$

ved å interpolere f , og så integrere interpolasjonspolynomet analytisk. Dette kalles kvadratur, og en bestemt metode kalles gjerne *kvadraturregel*. La x_i være interpolasjonspunkter på $[a, b]$, og l_i de korrespondende lagrangefunksjonene, slik at

$$f(x) \approx \sum_{i=0}^n f(x_i) l_i(x).$$

Vi skriver

$$\begin{aligned} \int_a^b f(x) dx &\approx \int_a^b \sum_{i=0}^n f(x_i) l_i(x) dx \\ &= \sum_{i=0}^n f(x_i) \int_a^b l_i(x) dx \\ &= \sum_{i=0}^n f(x_i) A_i = Q[f], \end{aligned}$$

der vi har definert $A_i = \int_a^b l_i(x) dx$. Disse kalles *kvadraturvektene*, eller bare *vektene*.

Eksempel 8.1. En av de aller enkleste kvadraturreglene kjenner du fra M1. Den kalles trapesregelen, er gitt ved

$$\int_a^b f(x) dx \approx (b-a) \left(\frac{f(a) + f(b)}{2} \right)$$

og utledes enkelt ved å interpolere f med et førsteordens polynom i $x_0 = a$ og $x_1 = b$ og integrere dette. Vektene er

$$A_0 = A_1 = \frac{b-a}{2}. \quad \triangle$$

Eksempel 8.2. Gauss-Legendre-punktene for $n = 2$ på intervallet $[-1, 1]$ er $\pm\sqrt{\frac{1}{3}}$, og lagrangefunksjoner er

$$l_0(x) = \frac{\sqrt{3}}{2} \left(x + \sqrt{\frac{1}{3}} \right)$$

og

$$l_1(x) = -\frac{\sqrt{3}}{2} \left(x - \sqrt{\frac{1}{3}} \right).$$

Vi beregner

$$A_0 = \frac{\sqrt{3}}{2} \int_{-1}^1 x + \sqrt{\frac{1}{3}} dx = 1$$

og

$$A_1 = -\frac{\sqrt{3}}{2} \int_{-1}^1 x - \sqrt{\frac{1}{3}} dx = 1. \quad \triangle$$

Eksempel 8.3. Simpsons regel, som du også kjenner fra M1,

$$Q[f] = \frac{b-a}{6} \left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right),$$

har vekter

$$A_0 = A_2 = \frac{b-a}{6} \quad \text{og} \quad A_1 = \frac{2(b-a)}{3}.$$

Disse utledes ved å interpolere f med andre ordens polynomer i a, b og $\frac{a+b}{2}$, men dette er gjort i M1, så vi dropper det. \triangle

Feilestimat

Det første vi kan gjøre, er å sette opp en generell regel for kvadraturfeil. Denne detter rett ut av feilestimatet for interpolasjon.

Teorem 8.4. La Q være en $n+1$ -punkts kvadraturregel på $[a, b]$, og anta at f er $n+1$ ganger kontinuerlig deriverbar på $[a, b]$. En øvre skranke for feilen er

$$|I[f] - Q[f]| = \frac{M}{(n+1)!} \int_a^b \prod_{i=0}^n |x - x_i| dx$$

der $M = \max_{s \in [a, b]} |f^{(n+1)}(s)|$.

Bevis. Integrer feilestimatet

$$f(x) - p_n(x) = \frac{f^{(n+1)}(s)}{(n+1)!} \prod_{k=0}^n (x - x_k).$$

fra a til b og bruk at $|f^{(n+1)}| \leq M$. \square

Fordelen med dette teoremet er at det er lett å bevis. Ulempen er at det stort sett er mulig å finne skarpere estimater, men disse er mer grisete å utlede. Vi tar derfor med noen bedre feilestimater for forskjellige kvadraturregler uten bevis.

Teorem 8.5. For trapesregelen finnes det en $s \in (a, b)$ slik at

$$I[f] - Q[f] = \frac{(b-a)^3}{12} f''(s).$$

Teorem 8.6. For Simpsons regel finnes det en $s \in (a, b)$ slik at

$$I[f] - Q[f] = -\frac{(b-a)^5}{2880} f^4(s).$$

Det går an å skrive opp feilestimater for kvadraturregler basert på Chebyshev-punkter, Gauss-Legendre-punkter og Gauss-Lobatto-punkter, men de er ganske grisete, og bevisen er altfor komplisert for oss. Vi nøyer oss derfor med noen eksempler, der vi sammenlikner med den analytiske verdien til integralet.

Eksempel 8.7. Vi tilnærmer

$$\int_{-1}^1 e^x dx = e - e^{-1} = 2.350402387287603$$

med trapesregelen:

$$Q[f] = e^{-1} + e \approx 3.086161269630488.$$

Feilen er i samme størrelsesorden som svaret. Triste greier. \triangle

Eksempel 8.8. Med Simpsons regel går det litt bedre:

$$Q[f] = \frac{1}{3}(e^{-1} + 4e^0 + e) \approx 2.362053756543496$$

Her er feilen $0.011651369255893 \approx 10^{-2}$. \triangle

Eksempel 8.9. Gauss-Legendre med $n = 2$ gir

$$Q[f] = e^{-\sqrt{1/3}} + e^{\sqrt{1/3}} \approx 2.342696087909730$$

Her er feilen på $-0.007706299377873 \approx 10^{-2}$, og faktisk mindre enn for Simpsons regel, til tross for at kvadraturregelen er basert på en lavere ordens interpolasjon. Dette eksemplet illustrerer at plasseringen av interpolasjonspunktene har mye å si. \triangle

Presisjonsgrad

Har forskjellige kvadraturregler noen andre gode egenskaper enn høy presisjon? Vi sier at en integrasjonsformel er eksakt for funksjonen f dersom

$$\int_a^b f dx = \sum_i f(x_i)A_i.$$

Dersom du interpolerer et polynom av grad n eller lavere med et polynom av grad n , blir f og p identiske. Derfor må det være klart at

$$\int_a^b p(x) dx = \sum_{i=0}^n f(x_i)A_i,$$

for alle polynomer av grad n eller lavere. Dersom kvadraturregelen er intelligent designet, kan man oppnå høyere presisjonsgrad enn som så.

Eksempel 8.10. Trapesregelen har presisjonsgrad 1, for

$$Q[1] = (b-a) \left(\frac{1+1}{2} \right) = b-a = \int_a^b dx$$

og

$$Q[x] = (b-a) \left(\frac{a+b}{2} \right) = \frac{b^2-a^2}{2} = \int_a^b x dx,$$

mens

$$\begin{aligned} Q[x^2] &= (b-a) \left(\frac{a^2+b^2}{2} \right) \\ &= \frac{b^3-ab^2+a^2b-a^3}{2} \neq \int_a^b x^2 dx. \end{aligned}$$

Disse beregningene viser at trapesregelen integrerer alle første ordens polynomer riktig, siden

$$Q[cx+d] = cQ[x] + dQ[1].$$

Ingen andre ordens polynomer integreres riktig, siden

$$Q[cx^2+dx+e] = cQ[x^2] + dQ[x] + eQ[1],$$

og $Q[x^2]$ ikke integreres riktig. \triangle

Eksempel 8.11. Gauss-Lobatto $n = 2$ har presisjonsgrad 3. Dette kan vises på samme måte som for trapesregelen, men vi venter litt med å se på det. Det er nemlig ikke så vanskelig å vise generelt at en $n+1$ -punkts Gauss-Lobatto-regel har presisjonsgrad $2n+1$, og dette skal vi gjøre senere. \triangle

Eksempel 8.12. Simpsons regel har presisjonsgrad 3. En rask titt på feilestimatet forteller at dersom f er et tredjegradspolynom, er $f^4 = 0$. \triangle

Noen forskjellige kvadraturklasser

Kvadraturregler skilles av er punktfordelingen, og hvorvidt man interpolerer med et polynom av høy grad på alle punktene, eller deler opp intervallet og interpolerer med stykkvis kontinuerlige polynombiter, såkalt *sammensatte regler*. Vi skal ta for oss et par klassiske kvadraturregler, og avslutte med en diskusjon rundt sammensatte regler.

Newton-Cotes

Både trapesregelen og Simpsons metode er eksempler på Newton-Cotes-regler. Dette er regler der interpolasjonen er gjort på ekvidistante gitre. En klassisk lærebok i numerisk analyse vil typisk inneholde en lengre utgreining om Newton-Cotes, men vi vet jo at man skal styre unna polynominterpolasjon på ekvidistante gitre, så derfor lar vi Newton-Cotes ligge.

Eksempel 8.13. Trapesregelen og Simpsons regel er Newton-Cotes-regler med henholdsvis $n = 2$ og $n = 3$. \triangle

Clenshaw-Curtis

Clenshaw-curtiskvadratur baserer seg på å integrere chebyshevinterpolanten. For chebyshevinterpolasjon er det lett å skrive opp formler for interpolasjonspunktene, og pene formler for kvadraturvektene, men disse formlene er vanskelige å utlede.

Teorem 8.14. Dersom n er et partall, og interpolasjonsgitteret er et n -punkts chebyshev ekstremalgitter, er vektene til kvadraturregelen gitt ved

$$w_1 = w_n = \frac{1}{(n-1)^2}$$

og

$$w_i = \frac{2}{(n-1)} \left(1 - \sum_{j=1}^{(n-2)/2} \frac{2}{4j^2-1} \cos \frac{2j(i-1)}{n-1} \pi \right)$$

for $2 \leq i \leq n-1$.

Gauss-Legendre

Vi skal beskrive prosessen som produserer Legendre-polynomene, og vise at denne prosessen produserer kvadraturformler med presisjonsgrad $2n + 1$.

Teorem 8.15. *La q være et polynom av grad $n + 1$ slik at*

$$\int_a^b qp \, dx = 0 \quad (8.1)$$

for alle polynomer p av grad mindre enn eller lik n . En kvadraturregel med disse $n + 1$ nullpunktene som noder, vil være eksakt for alle polynomer av grad mindre enn eller lik $2n + 1$.

Bevis. Vi begynner med å vise at polynomet q har $n + 1$ forskjellige nullpunkter på intervallet $[a, b]$. La x_0, x_1, \dots, x_{r-1} være de r punktene der q bytter fortegn på $[a, b]$. Polynomet

$$\prod_{i=0}^{r-1} (x - x_i)$$

har grad r , og bytter fortegn nøyaktig samtidig som q , slik at enten

$$q \prod_{i=0}^{r-1} (x - x_i) \leq 0$$

eller

$$q \prod_{i=0}^{r-1} (x - x_i) \geq 0$$

på $[a, b]$. Dette betyr at

$$\int_a^b q \prod_{i=0}^{r-1} (x - x_i) \, dx \neq 0,$$

og siden vi har

$$\int_a^b qp \, dx = 0 \quad (8.2)$$

for alle polynomer p av grad mindre enn eller lik n , impliserer dette at $r \geq n + 1$. Siden r åpenbart ikke kan være større enn $n + 1$, må $r = n + 1$. Altså bytter q fortegn $n + 1$ ganger på $[a, b]$, og må følgelig ha $n + 1$ nullpunkter på $[a, b]$.

La nå h være et polynom av grad $2n + 1$ eller lavere, og del h på q med rest r :

$$h = qp + r.$$

der p har grad n , og r har maksimal grad n . Vi evaluerer h i x_i , og ser at

$$h(x_i) = q(x_i)p(x_i) + r(x_i) = r(x_i),$$

siden $q(x_i) = 0$ for alle i . Siden kvadraturregelen er basert på integrasjon av et n -te ordens polyom, er den åpenbart eksakt for polynomer av orden n eller lavere, og vi kan beregne

$$\begin{aligned} \int_a^b h(x) \, dx &= \int_a^b q(x)p(x) + r(x) \, dx = \int_a^b r(x) \, dx \\ &= \sum_{i=0}^n r(x_i)A_i = \sum_{i=0}^n h(x_i)A_i. \end{aligned} \quad \square$$

Eksempel 8.16. Vi tar et eksempel der vi konstruerer q fra grunnen av. La $n = 1$ og $[a, b] = [0, 1]$. Vi må finne

$$q(x) = ax^2 + bx + c.$$

Vi begynner med å kreve at q skal stå ortogonalt på alle polynomer av grad 1 eller lavere. Dersom

$$\int_0^1 q(x) \, dx = \int_0^1 ax^2 + bx + c \, dx = \frac{a}{3} + \frac{b}{2} + c = 0$$

og

$$\int_0^1 xq(x) \, dx = \int_0^1 ax^3 + bx^2 + cx \, dx = \frac{a}{4} + \frac{b}{3} + \frac{c}{2} = 0,$$

vil

$$\int_0^1 (ax + b)q(x) \, dx = 0.$$

Vi ganger den første likningen med seks, den andre med tolv, og får likningssystemet

$$\begin{pmatrix} 2 & 3 & 6 \\ 3 & 4 & 6 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Litt gausseliminasjon gir

$$\begin{pmatrix} 2 & 3 & 6 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Her står det at $a + b = 0$, og at $2a + 3b + 6c = 0$. Løsningsrommet er

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = c \begin{pmatrix} 6 \\ -6 \\ 1 \end{pmatrix},$$

slik at

$$q(x) = 6x^2 - 6x + 1,$$

er et mulig valg for q . Andre valg av konstanten c vil gi andre polynomer, men alle vi ha de samme nullpunktene, og det er dem vi er ute etter. Nullpunktene til q er

$$x_0 = \frac{1}{2} - \frac{1}{2\sqrt{3}} \quad \text{og} \quad x_1 = \frac{1}{2} + \frac{1}{2\sqrt{3}}.$$

Lagrangefunksjonene blir

$$L_0 = \frac{x - x_1}{x_0 - x_1} = \sqrt{3} \left(x - \left(\frac{1}{2} + \frac{1}{2\sqrt{3}} \right) \right)$$

og

$$L_1 = \frac{x - x_0}{x_1 - x_0} = \sqrt{3} \left(x - \left(\frac{1}{2} - \frac{1}{2\sqrt{3}} \right) \right)$$

slik at

$$A_1 = \sqrt{3} \int_0^1 x - \left(\frac{1}{2} - \frac{1}{2\sqrt{3}} \right) \, dx = \frac{1}{2} = A_0.$$

Integrasjonsrutinen vi har laget skal være eksakt for alle polynomer opp til grad 3. Vi tester:

$$\frac{1}{2} \left(\frac{1}{2} - \frac{1}{2\sqrt{3}} \right)^2 + \frac{1}{2} \left(\frac{1}{2} + \frac{1}{2\sqrt{3}} \right)^2 = \frac{1}{3} = \int_0^1 x^2 \, dx$$

$$\frac{1}{2} \left(\frac{1}{2} - \frac{1}{2\sqrt{3}} \right)^3 + \frac{1}{2} \left(\frac{1}{2} + \frac{1}{2\sqrt{3}} \right)^3 = \frac{1}{4} = \int_0^1 x^3 dx.$$

Merk at punktene er bare Gauss-Legendre-punktene flyttet til intervallet $[0, 1]$. Dette eksemplet illustrerer hvordan Gauss-Legendre-tabellen i interpolasjonskapitlet er konstruert. \triangle

Kommentar. Legendrepolynomene er definert på intervallet $[-1, 1]$, så eksemplet over laget ikke et legendrepolytom, men derimot et polynom som har nullpunktene til legendrepolytomet

$$P_2 = \frac{1}{2}(3x^2 - 1)$$

flyttet til intervallet $[0, 1]$.

Vi kan utvide tabellen fra forrige kapittel med vektorer. Husk at tabellen gjelder for $[-1, 1]$, så om du trenger integrasjonsrutine for andre intervaller, må punktene flyttes, og vektene beregnes på nytt.

n	x_i	A_i
2	$\pm\sqrt{\frac{1}{3}}$	1
3	0	$\frac{8}{9}$
	$\pm\sqrt{\frac{3}{5}}$	$\frac{5}{9}$
4	$\pm\sqrt{\frac{3}{7} - \frac{2}{7}\sqrt{\frac{6}{5}}}$	$\frac{18+\sqrt{30}}{36}$
	$\pm\sqrt{\frac{3}{7} + \frac{2}{7}\sqrt{\frac{6}{5}}}$	$\frac{18-\sqrt{30}}{36}$
5	0	$\frac{128}{225}$
5	$0, \pm\frac{1}{3}\sqrt{5 - 2\sqrt{\frac{10}{7}}}$	$\frac{322+13\sqrt{70}}{900}$
5	$0, \pm\frac{1}{3}\sqrt{5 + 2\sqrt{\frac{10}{7}}}$	$\frac{322-13\sqrt{70}}{900}$

Merk at alle vektene er positive. Kvadraturentusiasier regner dette som et kvalitetsstempel.

Gauss-Lobatto

Vi koster på oss en tabell med vektorer for Gauss-Lobatto også.

n	x_i	A_i
3	0	$\frac{4}{3}$
	± 1	$\frac{1}{3}$
4	$\pm\sqrt{\frac{1}{5}}$	$\frac{5}{6}$
	± 1	$\frac{1}{6}$
5	0	$\frac{32}{45}$
	$\pm\sqrt{\frac{3}{7}}$	$\frac{49}{90}$
	± 1	$\frac{1}{10}$
6	$\pm\sqrt{\frac{1}{3} - \sqrt{\frac{2}{3\sqrt{7}}}}$	$\frac{14+\sqrt{7}}{30}$
	$\pm\sqrt{\frac{1}{3} + \sqrt{\frac{2}{3\sqrt{7}}}}$	$\frac{14-\sqrt{7}}{30}$
	± 1	$\frac{1}{15}$
7	0	$\frac{256}{525}$
	$\pm\sqrt{\frac{5}{11} - \frac{2}{11}\sqrt{\frac{5}{3}}}$	$\frac{124+7\sqrt{15}}{350}$
	$\pm\sqrt{\frac{5}{11} + \frac{2}{11}\sqrt{\frac{5}{3}}}$	$\frac{124-7\sqrt{15}}{350}$
	± 1	$\frac{1}{21}$

Eksempel 8.17. Simpsons regel er også en Gauss-Lobatto-regel med $n = 3$. \triangle

9 Metoder for ordinære differensiallikninger

Vi skal lage numeriske metoder for å finne tilnærmede løsninger for initialverdiproblemet

$$y' = f(x, y) \quad y(x_0) = y_0.$$

Dette er et kjempfelt. Vi har bare tid til å skrape så vidt i overflaten, men vi skal prøve å belyse et par momenter.

Runge-Kutta-metoder

En numerisk metode for ordinære differensiallikninger starter med følgende to observasjoner:

- Vi vet hva den analytiske løsningen er i x_0 . Dette vet vi på grunn av initialkravet $y(x_0) = y_0$.
- Vi vet hvilket stigningstall den analytiske løsningen har i x_0 , for evaluerer vi differensiallikningen i x_0 , får vi $y'(x_0) = f(x_0, y_0)$.

La oss lage oss et punkt x_1 litt ut fra x_0 , med avstand $h = x_1 - x_0$. Siden vi har funksjonsverdien og stigningstallet til y i x_0 , kan vi bruke lineær tilnærming, og gjette på $y(x_1)$:

$$y(x_1) \approx y(x_0) + hf'(x_0) = y(x_0) + hf(x_0, y_0).$$

Nå definerer vi $y_1 = y(x_0) + hf(x_0, y_0) \approx y(x_1)$. Dette er den tilnærmede verdien til y i x_1 . Vi tar den for god fisk, lager oss et nytt punkt $x_2 = x_1 + h$, og beregner

$$y_2 = y_1 + hf(x_1, y_1),$$

som er en tilnærming $y_2 \approx y(x_2)$. Nå fortsetter vi i samme stilen, girtrer opp intervallet vi skal løse likningen på med gitterfinhet h , slik at punktene er gitt ved $x_i = ih$. Tilnærmingen til $y(x_i)$ kaller vi y_i , og metoden kalles *Eulers eksplisitte metode*:

$$y_{i+1} = y_i + hf(x_i, y_i),$$

Metoden kalles *eksplisitt*, siden likningen kommer ferdig løst for y_{i+1} .

Vi skriver nå opp et par andre varianter. Alle er basert på å bytte ut stigningen $f(x_i, y_i)$ med et eller annet estimat. Setter vi inn $f(x_{i+1}, y_{i+1})$ istedet for $f(x_i, y_i)$, får vi *Eulers implisitte metode*:

$$y_{i+1} = y_i + hf(x_{i+1}, y_{i+1}),$$

og bytter vi ut med $\frac{1}{2}(f(x_i, y_i) + f(x_{i+1}, y_{i+1}))$, får vi *trapesmetoden*:

$$y_{i+1} = y_i + \frac{h}{2}(f(x_i, y_i) + f(x_{i+1}, y_{i+1})).$$

Disse to metodene kalles *implisitte* fordi likningene ikke er ferdig løst for y_{i+1} . Noen ganger er det lett å finne y_{i+1} , andre ganger ikke.

Hvis vi bytter ut $f(x_{i+1}, y_{i+1})$ i trapesmetoden med en tilnærming basert på et eksplisitt eulersteg, får vi den eksplisitte *Heuns metode*:

$$y_i^* = y_i + hf(x_i, y_i) \\ y_{i+1} = y_i + \frac{h}{2}(f(x_i, y_i) + f(x_{i+1}, y_i^*)),$$

og hvis vi klinker til og bytter ut tilnærmingen til stigningstallet med følgende avanserte opplegg, får vi nok en eksplisitt variant, nemlig Runge-Kuttas klassiske fjerdeordens metode, populært kalt RK4:

$$k_1 = f(x_i, y_i) \\ k_2 = f\left(x_i + \frac{h}{2}, y_i + \frac{h}{2}k_1\right) \\ k_3 = f\left(x_i + \frac{h}{2}, y_i + \frac{h}{2}k_2\right) \\ k_4 = f(x_i + h, y_i + hk_3) \\ y_{i+1} = y_i + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4).$$

Dette er alle metodene vi skal analysere, og alle er eksempler på *Runge-Kutta-metoder*. Nå lurer du sikkert på hvorfor man har så mange forskjellige metoder, og det korte svaret er som ellers i anvendt matematikk: noen metoder eksisterer fordi de er lette å finne opp og forstå, mens andre metoder finnes fordi de er skikkelig bra.

Eksempel 9.1. Vi løser initialverdiproblemet

$$y' = -y \quad y(0) = 1$$

med Eulers eksplisitte metode på intervallet $[0, 1]$. Siden $f(x, y) = -y$, blir metoden

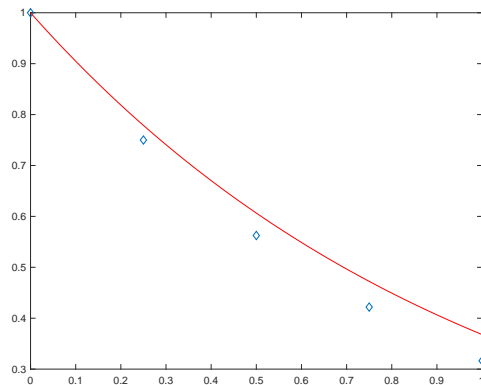
$$y_{i+1} = y_i - hy_i = (1 - h)y_i$$

med

$$y_0 = 1.$$

Løsning for $h = 0.25$ gir figuren under. De blå diamantene er y_1, y_2, y_3, y_4 og y_5 , mens den røde kurven er den analytiske løsningen $y = e^{-x}$. Vi beregner $y(1) = 1/e \approx 0.367879441171442$, som kan sammenliknes med $y_5 = 0.31640625$:

$$y_5 - y(1) = -0.051473191171442. \quad \triangle$$



Eksempel 9.2. Vi løser samme problem som i sted med Eulers implisitte metode. Metoden blir

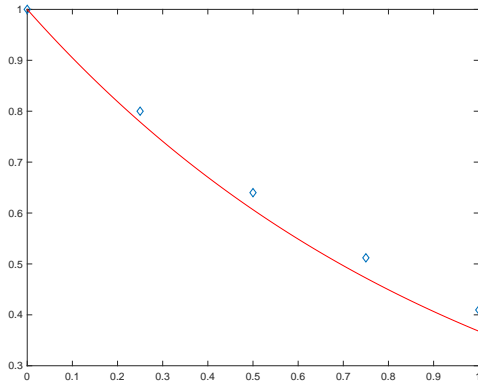
$$y_{i+1} = y_i - hy_{i+1},$$

som vi løser for y_{i+1} , og får

$$y_{i+1} = \frac{y_i}{(1+h)}.$$

Figur under for $h = 0.5$. Vi får $y_5 = 0.4096$, og

$$y_5 - y(1) = 0.041720558828558. \quad \triangle$$



Eksempel 9.3. Trapesmetoden:

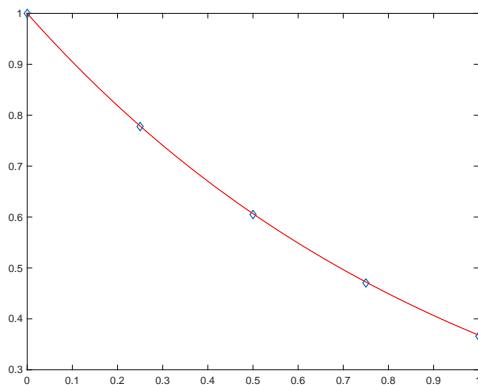
$$y_{i+1} = y_i - \frac{1}{2}h(y_i + y_{i+1}).$$

Vi løser for y_{i+1} , og får

$$y_{i+1} = \frac{2+h}{2-h}y_i.$$

Denne treffer noe bedre:

$$y_5 - y(1) = -0.001929128719072. \quad \triangle$$



Eksempel 9.4. Heuns metode:

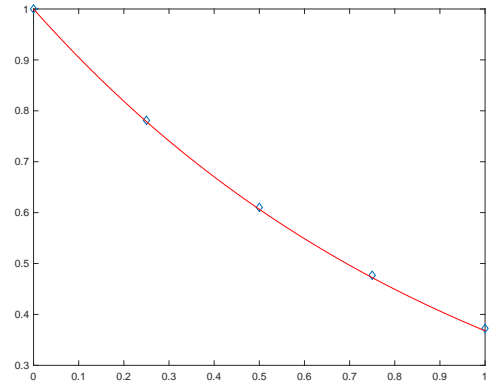
$$y_i^* = y_i - hy_i$$

$$y_{i+1} = y_i - \frac{1}{2}h(y_i + y_i^*)$$

og

$$y_5 - y(1) = 0.004649588674749.$$

Bedre enn Euler, men ikke helt trapesmetoden. \triangle



Eksempel 9.5. Til slutt RK4:

$$k_1 = -y_i$$

$$k_2 = -\left(y_i + \frac{h}{2}k_1\right)$$

$$k_3 = -\left(y_i + \frac{h}{2}k_2\right)$$

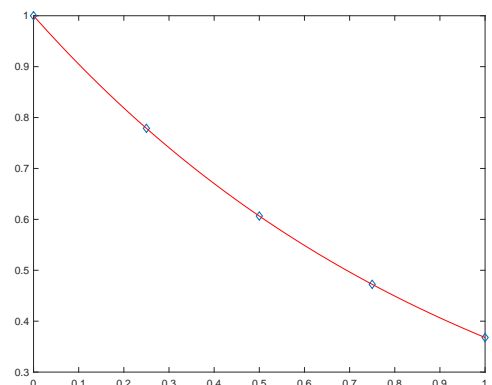
$$k_4 = -(y_i + k_3)$$

$$y_{i+1} = y_i + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4).$$

Denne treffer ganske bra:

$$y_5 - y(1) = 1.475823530627807e - 05.$$

Noe må den ha igjen for å være så komplisert. \triangle



Utleddningsmetoder

I forrige avsnitt utledet vi Eulers eksplisitte metode. For å sette disse metodene i kontekst med tidligere pensum i kurset, og for å indikere hvordan man kan lage flere metoder, skal vi utlede noen av dem med kjente metoder for numerisk integrasjon og derivasjon.

En uledningsteknikk er å gitte opp med $h = x_{i+1} - x_i$, integrere differensiallikningen

$$\begin{aligned} y(x_{i+1}) - y(x_i) &= \int_{x_i}^{x_{i+1}} y'(x) dx \\ &= \int_{x_i}^{x_{i+1}} f(x, y(x)) dx, \end{aligned}$$

bruke $y_{i+1} - y_i \approx y(x_{i+1}) - y(x_i)$ på venstre side, og tilnærme integralet

$$\int_{x_i}^{x_{i+1}} f(x, y(x)) dx,$$

med en kvadraturregel. Gjør man den særdeles enkle tilnærmingen

$$\int_{x_i}^{x_{i+1}} f(x, y(x)) dx \approx \int_{x_i}^{x_{i+1}} f(x_i, y_i) dx = hf(x_i, y_i),$$

får man eksplisitt Euler, og velger man den tilsvarende enkle tilnærmingen

$$\int_{x_i}^{x_{i+1}} f(x, y(x)) dx \approx hf(x_{i+1}, y_{i+1}),$$

får man Eulers implisitte metode. Trapesregelen

$$\int_{x_i}^{x_{i+1}} f(x, y(x)) dx \approx \frac{h}{2}(f(x_i, y_i) + f(x_{i+1}, y_{i+1}))$$

gir trapesmetoden

$$y_{i+1} = y(x_i) + \frac{h}{2}(f(x_i, y_i) + f(x_{i+1}, y_{i+1})).$$

Tilnærmer man trapesregelen med formelen

$$\int_{x_i}^{x_{i+1}} f(x, y(x)) dx \approx \frac{h}{2}(f(x_i, y_i) + f(x_{i+1}, y_{i+1}^*))$$

der $y_{i+1}^* = y_i + hf(x_i, y_i)$ er et eksplisitt eulersteg, får man Heuns metode. Bruker man tilnærmingen

$$\int_{x_i}^{x_{i+1}} f(x, y(x)) dx \approx hf\left(\frac{x_i + x_{i+1}}{2}, \frac{y_i + y_{i+1}}{2}\right)$$

får man *midtpunktmetoden*.

RK4 er avledet fra Simpsons metode. Tilnærmer man

$$\int_{x_i}^{x_{i+1}} f(x, y(x)) dx,$$

med Simpson, får man en implisitt metode som ikke er pensum, og bytter man ut de implisitte verdiene leddene i denne metoden med forskjellige estimater basert på eksplisitt Euler, får man RK4, litt som Heuns metode er avledet fra trapesmetoden.

Metodene vi har utledet til nå, kalles *enstegsmetoder*, for kun y_{i+1} og y_i figurerer i likningene. Grunnen til at alle metodene er på denne formen, er at det

kun er brukt en type endelig differansetilnærming på venstre side av

$$y'(x) = f(x, y(x)),$$

nemlig den første ordens differansen

$$y'(x_{i+1}) \approx \frac{y_{i+1} - y_i}{h}.$$

Nå er det ingenting i veien for å bruke en høyere ordens tilnærming, for eksempel sentraldifferansen

$$y'(x_i) \approx \frac{y_{i+1} - y_{i-1}}{2h}.$$

Setter man denne inn for y' , får man *leap-frog-metoden*

$$y_{i+1} = y_{i-1} + 2hf(x_i, y_i).$$

Her inngår både y_{i+1} , y_i og y_{i-1} , og leap-frog er et eksempel på en *flerstegsmetode*. Flerstegsmetoder er ikke pensum i dette kurset.

Feilanalyse

I dette avsnittet skal vi ta en titt på hvorfor metodene treffer så foreskjellig. Vi skal indikere hvordan analysen får for eksplisitt Euler, og så skrive opp resultatet for de andre metodene.

Lineariseringen som gir det første eulersteget er

$$\begin{aligned} y(x_1) &= y(x_0 + h) \\ &\approx y(x_0) + hf(x_0, y(x_0)) \\ &= y(x_0) + hy'(x_0). \end{aligned}$$

Vi antar at y er en analytisk funksjon, og taylorutvikler:

$$\begin{aligned} y(x_1) &= y(x_0 + h) \\ &= y(x_0) + hy'(x_0) + \frac{y''(x_0)}{2}h^2 + \dots \end{aligned}$$

Sammenlikner vi denne med

$$y_1 = y(x_0) + hy'(x_0),$$

ser vi at feilen i det første eulersteget er gitt ved

$$y(x_1) - y_1 = \frac{y''(x_0)}{2}h^2 + \frac{y'''(x_0)}{6}h^3 + \dots$$

altså taylorrekkehalen til y . Hvis vi antar at h er liten, slik at h^2 er mye større enn h^3 , og leddet

$$\frac{y''(x_0)}{2}h^2$$

dominerer halen på taylorrekken, er det ikke urimelig å hevde at eksplisitt Euler har *lokal feil* av størrelsesorden h^2 .

Feilen etter ett steg er altså av størrelsesorden h^2 . Men hva er feilen etter n steg? I eksemplene i forrige avsnitt, kjørte vi løserne på intervallet $[0, 1]$. La oss si at vi kjører på intervallet $[x_0, x_0 + a]$. Vi velger h slik at

$$x_n = x_0 + hn = x_0 + a$$

og

$$n = \frac{a}{h}.$$

Hvis vi nå gjør n steg med eksplisitt Euler, samler vi i hvert steg opp en lokal feil omtrent lik

$$\frac{y''(x_i)}{2}h^2.$$

Feilen etter n steg blir

$$\sum_{i=1}^n \frac{y''(x_i)}{2}h^2,$$

og hvis vi antar at $y'' \leq M$ på $[x_0, x_0 + a]$, er det rimelig å hevde at

$$\sum_{i=1}^n \frac{y''(x_i)}{2}h^2 \leq Mnh^2 = M\frac{a}{h}h^2 = Mah.$$

Vi sier derfor at eulers metode har *global feil* av størrelsesorden h .

Teorem 9.6. *Lokal og global feil for metode-*
ne:

Metode	Lokal feil	Global feil
Eksplisitt Euler	h^2	h
Implisitt Euler	h^2	h
Trapesmetoden	h^3	h^2
Heuns metode	h^3	h^2
RK4	h^5	h^4

Vi skal ikke bevise dette teoremet, men nevner at bevisteknikken er den samme for alle metodene: taylor-utvikle om x_0 for å finne lokal feil, og så se på hva som skjer etter n steg. Dette teoremet forklarer langt på vei hva som skjedde i eksemplene i forrige avsnitt. Nå tar vi et par eksempler der vi lar $h \rightarrow 0$.

Eksempel 9.7. Vi kjører samme eksempel som i forrige avsnitt, men nå bruker vi Eulers eksplisitte metode for $h = 0.1$, $h = 0.01$ og så videre. Resultatene er oppsummert i følgende tabell:

h	$y_n - y(1)$
10^{-1}	$-1.920100107144218e - 02$
10^{-2}	$-1.847099898213078e - 03$
10^{-3}	$-1.840164004788258e - 04$
10^{-4}	$-1.839473847314865e - 05$
10^{-5}	$-1.839403194148215e - 06$

Dette eksemplet demonstrerer tydelig at feilen etter n steg er proporsjonal med h . På folkemunne sier man gjerne at man får en ekstra korrekt desimal hver gang man tideler h . Eulers implisitte metode oppfører seg omtrent likt, så den hopper vi over. \triangle

Eksempel 9.8. Trapesmetoden for $h = 0.1$, $h = 0.01$ og så videre:

h	$y_n - y(1)$
10^{-1}	$-3.068987885734287e - 04$
10^{-2}	$-3.065695217463471e - 06$
10^{-3}	$-3.065658332745969e - 08$
10^{-4}	$-3.069314802317535e - 10$
10^{-5}	$-5.472844399889709e - 12$

Her er feilen etter n steg proporsjonal med h^2 . På folkemunne sier man gjerne at man får to desimaler hver gang man tideler h . Heuns metode produserer omtrent den samme tabellen. \triangle

Eksempel 9.9. RK4:

h	$y_n - y(1)$
10^{-1}	$3.332410560830112e - 07$
10^{-2}	$3.091293887536040e - 11$
10^{-3}	$3.996802888650564e - 15$
10^{-4}	$2.664535259100376e - 15$
10^{-5}	$-3.330669073875470e - 15$

Hva skjedde her? Feilen etter n steg proporsjonal med h^4 , altså fire desimaler for hver tideling av h , men bare for de første tre tidelingene. Når $h = 10^{-3}$ har vi nådd såkalt *maskinpresisjon*. Matlab regner bare med 16 desimaler, og dette setter en stopper for konvergensen. \triangle

Eksempel 9.10. RK4, men nå har matlab fått beskjed om å regne med 32 desimaler:

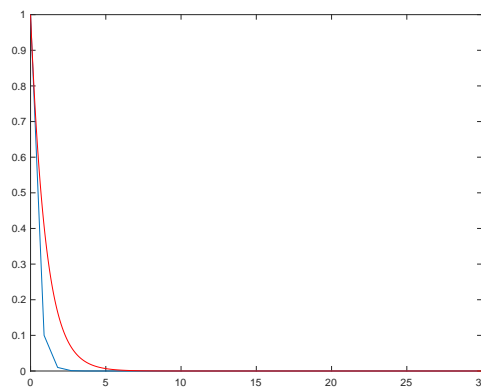
h	$y_n - y(1)$
10^{-1}	$3.332410561118064e - 07$
10^{-2}	$3.091319001284922e - 11$
10^{-3}	$3.068217823302200e - 15$
10^{-4}	$3.065917492545278e - 19$
10^{-5}	$3.065687557054922e - 23$

Tabellene til nå har tatt en brøkdel av et sekund å produsere. Til sammenlikning tok denne her rundt ti minutter, pluss noen timer knoting for å finne ut av hvordan matlab skal regne riktig med 32 desimaler. Presisjon koster! \triangle

Stabilitet

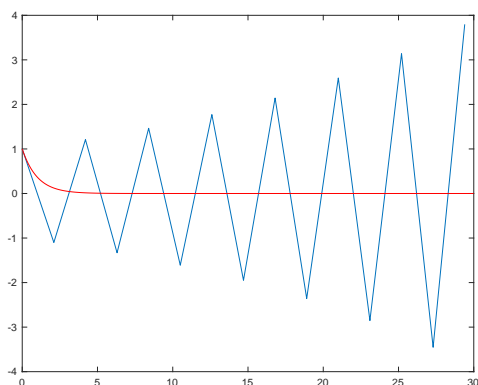
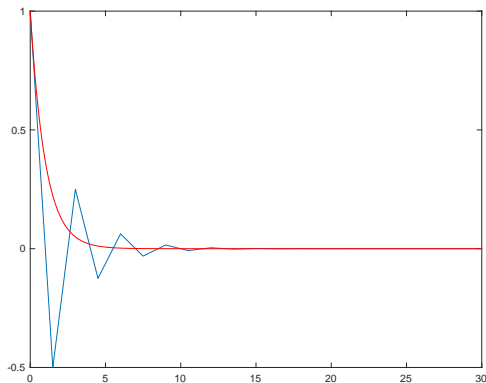
Har metodene noen andre egenskaper? Eulers eksplisitte og implisitte metoder ser til forveksling like ut, og har akkurat samme orden. Men de oppfører seg ganske forskjellig.

Eksempel 9.11. Vi kjører eksplisitt Euler på samme problem som over, men på intervallet $[0, 30]$, og $h = 0.9$. Det er trukket rette linjer mellom eulerstegene, så det skal bli litt enklere å se hva som skjer. \triangle

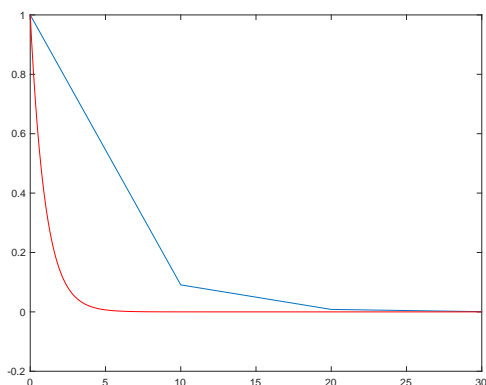


Eksempel 9.12. Vi kjører igjen på intervallet $[0, 30]$, men nå med $h = 1.5$. Den numeriske løsningen ser ut til å virre frem og tilbake en del før den sikter seg inn på rett spor. \triangle

Eksempel 9.13. Intervallet $[0, 30]$, men nå med $h = 2.1$. Hva den numeriske løsningen tenker på, er ikke godt å si, men noe fornuftig er det ihvertfall ikke. \triangle



Eksempel 9.14. Vi kan slå fast at eksplisitt Euler ikke fungerte, og det ser ut som om det går galt fordi h er for stor. Vi prøver Euler implisitt på samme intervall, men med $h = 10$. Det går riktig bra. \triangle



Forklaringen på hva som skjedde her, er ganske enkel. Eulers eksplisitte metode er, for eksemplet vi har studert,

$$\begin{aligned} y_{i+1} &= y_i - hy_i \\ &= (1-h)y_i \\ &= (1-h)^2 y_{i-1} \\ &= (1-h)^{i+1} y_0 = (1-h)^{i+1}, \end{aligned}$$

med andre ord en geometrisk følge. Siden gymnaset har du visst at denne følgen divergerer dersom

$$|1-h| \geq 1$$

og denne ulikheten blir innfridd akkurat i det h blir 2. Kjører vi samme resonnementet på Euler implisitt, får vi

$$y_{i+1} = \frac{1}{(1+h)^{i+1}}.$$

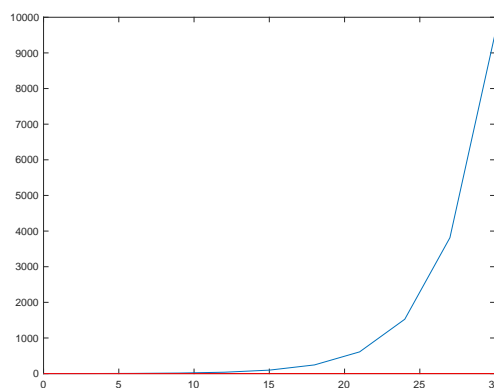
I vårt tilfelle er $h > 0$, så det må være klart at

$$0 < \frac{1}{(1+h)^{i+1}} < 1,$$

og følgelig konvergerer følgen mot 0 for alle valg av h .

Analysen vi har gjort, kalles *stabilitetsanalyse*, og $y' = -y$ er et såkalt *testproblem*. Vi får ikke eksakt informasjon om hvordan Eulers metode kommer til å oppføre seg for alle mulige differensiallikninger, men vi kan få en magesfølelse allikevel. Vi skal ikke gå inn på en lengre diskusjon om stabilitetsanalyse, som er et forskningsfelt i seg selv, men nevne at stabilitetsproblemer er som regel betydelige for eksplisitte metoder, og ikke-eksisterende for implisitte metoder.

Eksempel 9.15. Vi prøver Heuns metode på $[0, 30]$, med $h = 3$. Som du ser, går det ganske dårlig. \triangle



Eksempel 9.16. Vi prøver RK4 på $[0, 30]$, med $h = 2.1$. Det går ikke noe bedre. \triangle

Mer om implisitte metoder

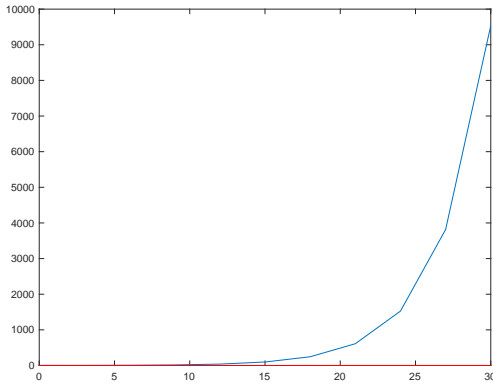
Så hvorfor bør vi ikke alltid bruke implisitte metoder? Det er et komplisert spørsmål å svare på, men vi skal gjøre et forsøk i dette avsnittet. Vi begynner med et eksempel, der vi setter opp de forskjellige numeriske metodene.

Eksempel 9.17. Vi skriver opp de forskjellige metodene for

$$y' = \frac{y - 2xy^2}{1+x}.$$

Eksplisitt Euler:

$$y_{i+1} = y_i + h \frac{y_i - 2x_i y_i^2}{1+x_i}$$



Imspisitt Euler:

$$y_{i+1} = y_i + h \frac{y_{i+1} - 2x_{i+1}y_{i+1}^2}{1 + x_{i+1}}$$

Trapesmetoden

$$y_{i+1} = y_i + \frac{h}{2} \left(\frac{y_i - 2x_i y_i^2}{1 + x_i} + \frac{y_{i+1} - 2x_{i+1} y_{i+1}^2}{1 + x_{i+1}} \right)$$

Heuns metode

$$y_{i+1}^* = y_i + h \frac{y_i - 2x_i y_i^2}{1 + x_i}$$

$$y_{i+1} = y_i + \frac{h}{2} \left(\frac{y_i - 2x_i y_i^2}{1 + x_i} + \frac{y_{i+1}^* - 2x_{i+1} (y_{i+1}^*)^2}{1 + x_{i+1}} \right)$$

RK4

$$k_1 = \frac{y_i - 2x_i y_i^2}{1 + x_i}$$

$$k_2 = \frac{(y_i + \frac{h}{2} k_1) - 2(x_i + \frac{h}{2})(y_i + \frac{h}{2} k_1)^2}{1 + (x_i + \frac{h}{2})}$$

$$k_3 = \frac{(y_i + \frac{h}{2} k_2) - 2(x_i + \frac{h}{2})(y_i + \frac{h}{2} k_2)^2}{1 + (x_i + \frac{h}{2})}$$

$$k_4 = \frac{(y_i + h k_3) - 2(x_i + h)(y_i + h k_3)^2}{1 + (x_i + h)}$$

$$y_{i+1} = y_i + \frac{h}{6} (k_1 + 2k_2 + 2k_3 + k_4).$$

Jeg har for RK4 beholdt parentesene for å prøve å beholde den visuelle likhetene med de generelle likningene som definerer metoden. Men det er strengt tatt ikke nødvendig. \triangle

I dette eksemplet kan y_{i+1} beregnes analytisk for de implisitte metodene, men merk at dette fort kan bli en smule håpløst om likningene ikke er kvadratiske i y_{i+1} , slik som her. Standardteknikken er da å slå til med en numerisk likningsløser. Vanligvis er fikspunktmetoden et greit valg. Iterasjonen

$$y_{i+1} = y_i + h \frac{y_{i+1} - 2x_{i+1}y_{i+1}^2}{1 + x_{i+1}}$$

allerede er på formen $y_{i+1} = g(y_{i+1})$, og dersom h er liten, blir gjerne g' liten, og da husker du fra tidligere at fikspunktmetoden konvergerer ganske kjapt.

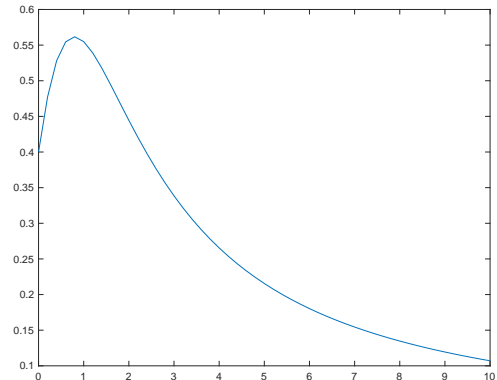
Eksempel 9.18. Vi løser

$$y' = \frac{y - 2xy^2}{1 + x} \quad y(0) = \frac{2}{5}$$

med

$$y_{i+1} = y_i + h \frac{y_{i+1} - 2x_{i+1}y_{i+1}^2}{1 + x_{i+1}}.$$

Under er løsningskurve beregnet med $h = 0.2$ på intervallet $[0, 10]$. Fikspunktmetoden trengte med startgjetning y_i et sted mellom 13 og 21 iterasjoner for å nå maskinpresisjon i hvert steg. Med lavere h vil antall iterasjoner gå ned. \triangle



Eksemplet over illustrerer et viktig moment. Kostet det mange flyttallsoperasjoner å kjøre en metode til en gitt presisjon? Det hjelper ikke å ha en metode som beregner alt til maskinpresisjon om metoden tar ett år å kjøre. Implisitte metoder er ofte robuste og stabile, men de koster også mer å bruke.

Systemer av differensiallikninger

Alle metodene gjennomgått til nå, fungerer like fint på systemer av differensiallikninger. Husk fra M3 at høyere ordens differensiallikninger kan skrives om til førsteordens systemer av differensiallikninger, så vi trenger ikke lage egne metoder for dem.

Eksempel 9.19. Differensiallikningen for en pendel er

$$y'' + \sin y = 0.$$

Vi skriver denne om til et system ved å sette $z = y'$, slik at systemet blir

$$y' = z$$

$$z' = -\sin y$$

Vi skriver nå opp metodene.

Eksplisitt Euler:

$$y_{i+1} = y_i + h z_i$$

$$z_{i+1} = z_i - h \sin y_i$$

Imspisitt Euler:

$$y_{i+1} = y_i + h z_{i+1}$$

$$z_{i+1} = z_i - h \sin y_{i+1}$$

Trapesmetoden:

$$y_{i+1} = y_i + \frac{h}{2}(z_i + z_{i+1})$$

$$z_{i+1} = z_i - \frac{h}{2}(\sin y_i + \sin y_{i+1})$$

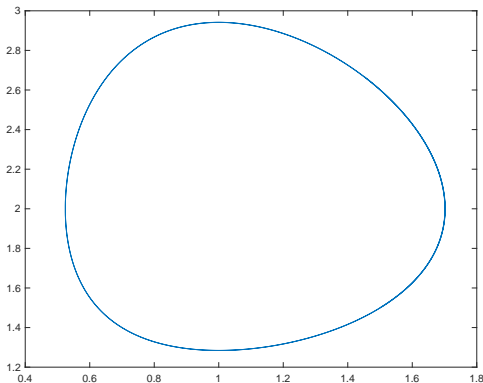
Skjøenner du disse her, klarer du nok Heun og RK4 også. (Alle k -ene blir vektorer med to komponenter.) Merk at man på de implisitte metodene må kjøre en eller annen flerdimensjonal likningsløser for å finne (y_{i+1}, z_{i+1}) i hvert steg. \triangle

Eksempel 9.20. Lotka-Volterra-systemet

$$y' = y(2 - z) \quad y(0) = 1.5$$

$$z' = z(y - 1) \quad z(0) = 1.5$$

beskriver to dyrepopulasjoner, der den ene driver med predasjon på den andre. Dersom det er mange mus (y) i fjellet, får rev og mår (z) rikelig med mat til ungene sine, men er det få mus, vokser ikke så mange unger opp. Vi løser dette systemet med eksplisitt Euler, og får figuren under. \triangle

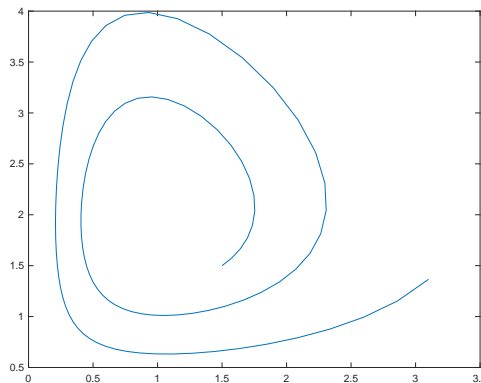


Bevaring av viktige størrelser

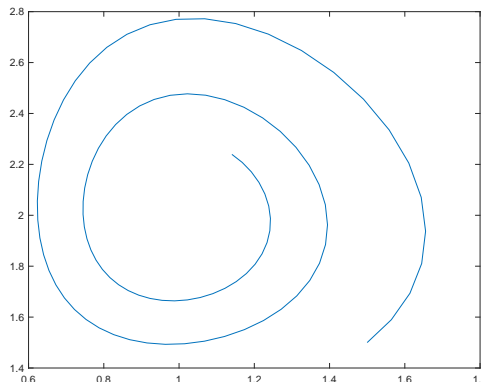
Nå skal vi ta for oss nok en instans av steg tre: har metodene noen andre egenskaper som er verdt å snakke om?

Det kan være lurt å ha følgende eksempel i bakhodet. La oss si at du ønsker å gjøre beregninger på planetenes gang i solsystemet. Du lærte på gymnasen at i solsystemet er energien for alle praktiske formål bevart, ihvertfall på kort sikt. Planetenes bevegelser følger Newtons gravitasjonslov, som for solsystemet vårt blir et differensiallikningssystem med tre likninger per planet. En numerisk metode for å simulere planetenes gang rundt solen, bør sørge for å bevare energien til hver planet, ellers kan planetene kjøre i numeriske spiraler ut i verdensrommet.

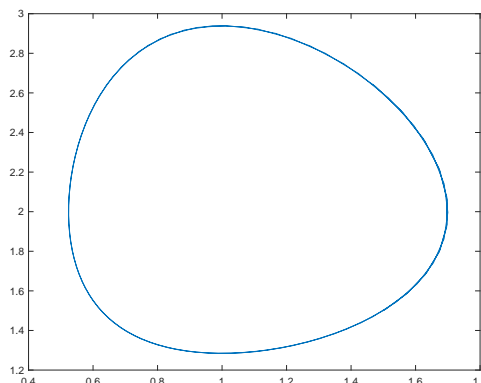
Eksempel 9.21. I forrige eksempel løste vi Lotka-Volterra med eksplisitt euler, og $h = 0.0001$ gav en pen figur. Hvis prøver med $h = 0.1$, får vi figuren under. Systemet starter i $(1.5, 1.5)$, så vi ser at Euler lager en utadgående spiral. \triangle



Eksempel 9.22. Nå kjører vi implisitt euler med $h = 0.1$, og får følgende figur. Siden systemet starter i $(1.5, 1.5)$, ser vi at dette er en innadgående spiral. \triangle



Eksempel 9.23. Trapesmetoden klarer visst å beholde det pene periodiske svingemønsteret, selv med $h = 0.1$. \triangle

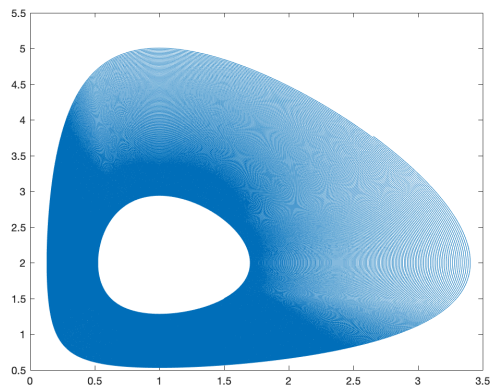
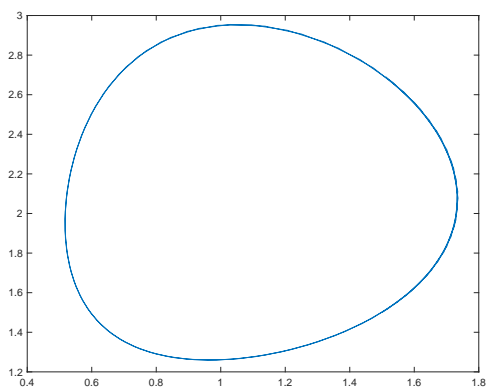


Hvorvidt figurene i de tre eksemplene over ble spiraler eller lukkede kurver, har ingenting med differensiallikningene å gjøre; det er kun forskjeller i de numeriske metodene som slår ut her. Løsningene skal definitivt være lukkede baner, og vi slår oss til ro med at trapesmetoden klarer noe som ikke eksplisitt og implisitt euler klarer. Vi merker oss også at numeriske metoder for differensiallikninger er mer enn bare presisjon og desimaler. Dette er for øvrig også et forskningsfelt som kalles *geometrisk integrasjon* - studiet om artige fenomener i numerisk løsning av differensiallikninger.

Eksempel 9.24. Vi tar et siste eksempel. *Symplektisk Euler* er definert ved følgende skjema

$$\begin{aligned}y_{i+1} &= y_i + hy_i(2 - z_i) \\z_{i+1} &= z_i + hz_i(y_{i+1} - 1),\end{aligned}$$

og produserer en tilsvarende pen figur. Symplektisk euler bevarer nemlig noe som kalles *den symplektiske strukturen*. \triangle



Eksempel 9.25. Som absolutt siste eksempel (jeg lover), tar vi eksplisitt Euler med $h = 0.001$, men for $x \in [0, 1000]$. Dette er litt for å demonstrere at det går galt selv med små h (det spiller ingen rolle hvor liten h er her, det går åt skogen uansett). Og litt fordi figuren var så vakker å skue. I eksempel 9.20 var ikke intervallet langt nok til at spiraleringen kom til uttrykk i figuren. Men den er der. \triangle

10 Metoder for partielle differensiallikninger

Vi skal lage numerisk metoder for varmelikningen

$$u_t = u_{xx}$$

med randkrav

$$u(0, t) = u(1, t) = 0$$

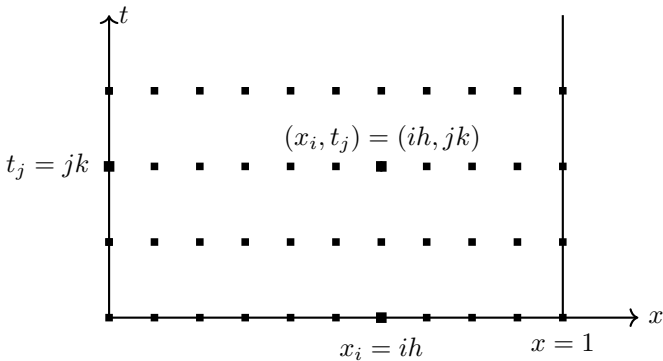
og

$$u(x, 0) = f(x).$$

Metodene vi skal lage er basert på ting vi har gjort tidligere, og varmelikningen er kun en illustrasjon. Det er relativt lett å konstruere liknende metoder for andre likninger med de teknikkene vi skal gå gjennom for varmelikningen.

Gitteret

Når vi skal løse en partiell differensiallikning, må vi holde styr på to gitre - et i x -retningen, og et i t -retningen. Vi skal finne approksimasjoner til løsningen u i alle gitterpunktene. Vi gitrer opp intervallet $[0, 1]$ på x -aksen med gitteravstanden h , og nummererer punktene slik at $x_0 = 0$ og $x_n = 1$. Den positive t -aksen gitrer vi opp med gitteravstanden k , og nummerer slik at $t_0 = 0$.



Diskretisering i x

Vi setter først opp en tilnærming for $u_{xx}(x, t)$, basert på den andre ordens endelige differanseformelen for x :

$$u_{xx}(x, t) \approx \frac{u(x+h, t) - 2u(x, t) + u(x-h, t)}{h^2}$$

Man kan fint bruke høyere ordens differanseformler, men det skal ikke vi gjøre. Vi kan nå tenke at vi erstatter $u(x, t)$ med $n+1$ envariable funksjoner $u_i(t)$, som beskriver temperaturendringen i hvert sitt punkt x_i på stangen:

$$u_{xx}(x_i, t) \approx \frac{u_{i+1}(t) - 2u_i(t) + u_{i-1}(t)}{h^2},$$

Nå setter vi inn dette uttrykket i varmelikningen:

$$u_t(x_i, t) = u_{xx}(x_i, t) \approx \frac{u_{i+1}(t) - 2u_i(t) + u_{i-1}(t)}{h^2},$$

Siden u_i kun er en funksjon av t , passer det å skrive

$$u_i'(t) = \frac{u_{i+1}(t) - 2u_i(t) + u_{i-1}(t)}{h^2},$$

som vi kjenner igjen som et system av ordinære differensiallikninger. Dette systemet har $n-1$ likninger, for $u_0 = 0$ og $u_n = 0$ kjenner vi; disse er gitt av randbetingelsene.

Diskretisering i t

Nå kan vi løse systemet med en ønsket metode fra forrige kapittel, og vi skal ta i betraktning tre valg. Etter diskretiseringen i t , skriver vi approksimasjonen i punktet (x_i, t_j) som u_{ij} :

$$u(x_i, t_j) \approx u_{ij}$$

De tre metodene vi skal bruke, er eksplisitt Euler, implisitt Euler, og trapesmetoden. De korrespondende skjemaene for varmelikningen blir:

Eksplisitt:

$$\frac{u_{i,j+1} - u_{ij}}{k} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2}$$

Implisitt:

$$\frac{u_{i,j+1} - u_{ij}}{k} = \frac{u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}}{h^2}$$

Crank-Nicholson:

$$\frac{u_{i,j+1} - u_{ij}}{k} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{2h^2} + \frac{u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}}{2h^2}$$

Teorem 10.1. *Det eksplisitte og implisitte skjema er av orden 1 i t , og av orden 2 i x . Crank-Nicolson er av orden 2 i begge variable.*

Beviset er en ganske hårete variant av argumentet for ordenen til eksplisitt Euler i forrige kapittel. Hver av disse skjemaene har sine fordeler, og vi skal behandle dem i tur og orden.

Eksplisitt skjema

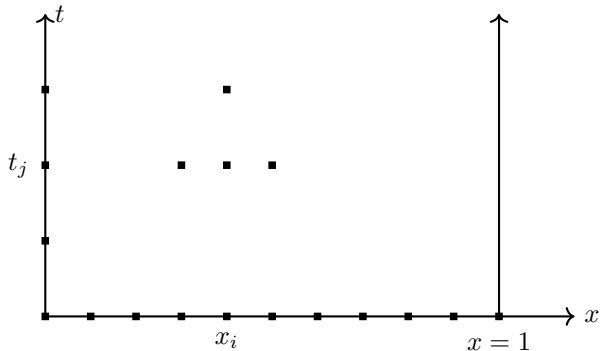
Det eksplisitte skjemaet

$$\frac{u_{i,j+1} - u_{ij}}{k} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2}$$

har egentlig bare en fordel: det er så lett å programmere opp. Vi løser for $u_{i,j+1}$, og får

$$u_{i,j+1} = u_{ij} + \frac{k}{h^2} (u_{i+1,j} - 2u_{i,j} + u_{i-1,j})$$

Nå er det vanlig å tegne opp noe som kalles *stensen*. Dette er en figur som illustrerer hvilke gitterpunkter som er involvert i ligningen man bruker for å beregne nye approksimasjoner.



Anta at du har beregnet u_{ij} for alle i og opp til og med en bestemt j . Vi kan da enkelt beregne $u_{i,j+1}$ for alle i ved å bruke formelen.

For å analysere mer, må vi sette opp et matrisevektorprodukt som beskriver iterasjonen. Vi definerer

$$\mathbf{u}_j = \begin{bmatrix} u_{1j} \\ u_{2j} \\ \vdots \\ u_{n-2,j} \\ u_{n-1,j} \end{bmatrix}$$

Det eksplisitte skjemaet kan skrives kompakt som

$$\mathbf{u}_{j+1} = (I - \frac{k}{h^2}A)\mathbf{u}_j$$

der

$$A = \begin{bmatrix} 2 & -1 & & & & & & & \\ -1 & 2 & -1 & & & & & & \\ & -1 & \ddots & \ddots & & & & & \\ & & \ddots & \ddots & & & & & \\ & & & & -1 & & & & \\ & & & & -1 & 2 & -1 & & \\ & & & & & -1 & 2 & & \end{bmatrix}$$

Det kan vises at denne matrisen har $n-1$ forskjellige egenverdier gitt ved

$$\lambda_k = 4 \sin^2 \frac{k\pi}{2n} \quad \text{for } 1 \leq k \leq n-1.$$

Egenverdiene til

$$I - \frac{k}{h^2}A$$

er dermed gitt ved

$$1 - 4 \frac{k}{h^2} \sin^2 \frac{k\pi}{2n}.$$

Dersom

$$\frac{k}{h^2} < \frac{1}{2}.$$

går dette veldig bra, for alle egenverdiene er mellom 0 og 1. Da vil

$$\mathbf{u}_{j+1} = (I - \frac{k}{h^2}A)\mathbf{u}_j$$

være en kontraksjon, og konvergere mot et fikspunkt.

Dersom

$$\frac{k}{h^2} \geq \frac{1}{2}$$

vil egenverdiene sende \mathbf{u}_i mot uendelig, og dette er ikke oppførsel vi ønsker fra et numerisk skjema for varmelikningen. Vi vet jo at temperaturen skal synke mot 0 i denne situasjonen!

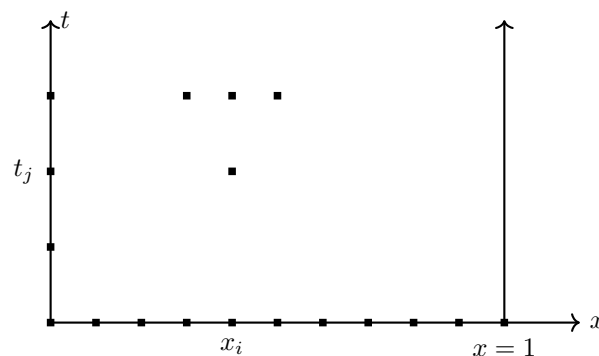
Dette er ganske restriktivt. Kanskje noen andre skjema er enklere å ha med å gjøre.

Implisitt skjema

Problemet med det eksplisitte skjemaet er at det går til $h \dots$ med mindre $k/h^2 < 1/2$, og dette gjør at du må ha veldig tett mellom punktene t -aksen. Akkurat som for ordinære differensiallikninger, kan vi bøte på dette med å bruke implisitt skjema. Det implisitte skjemaet skriver vi

$$\left(1 + 2\frac{k}{h^2}\right)u_{i,j+1} - \frac{k}{h^2}u_{i+1,j+1} - \frac{k}{h^2}u_{i-1,j+1} = u_{ij}.$$

Stensilen er



I hver tidssteg må vi løse et lineært likningssystem

$$\left(I + \frac{k}{h^2}A\right)\mathbf{u}_{j+1} = \mathbf{u}_j$$

der A er den samme matrisen som i sted.

Egenverdiene til $(I + \frac{k}{h^2}A)$ er

$$1 + 4 \frac{k}{h^2} \sin^2 \frac{k\pi}{2n},$$

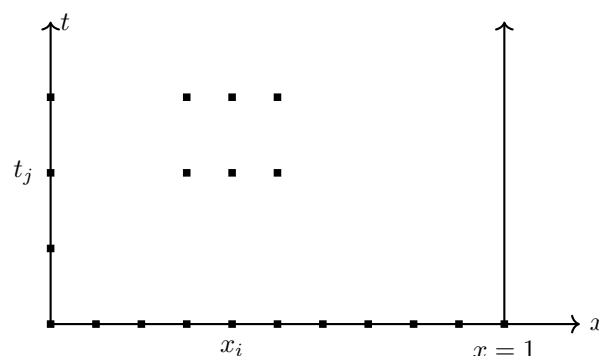
som åpenbart er større enn 1 uansett hva $\frac{k}{h^2}$ er. Dermed vil egenverdiene til $(I + \frac{k}{h^2}A)^{-1}$ alltid være mindre enn 1, og følgelig vil skjemaet konvergere mot 0 uansett.

Crank-Nicolson

Crank-Nicolson skriver vi

$$2u_{i,j+1} + \frac{k}{h^2}(-u_{i+1,j+1} + 2u_{i,j+1} - u_{i-1,j+1}) = 2u_{ij} - \frac{k}{h^2}(-u_{i+1,j} + 2u_{i,j} - u_{i-1,j}).$$

Stensilen er



m . Vi diskretiserer med de samme differanseformlene som før, og får eksplisitt skjema:

$$\frac{u_{i,j,m+1} - u_{i,j,m}}{k} = \frac{u_{i+1,j,m} - 2u_{i,j,m} + u_{i-1,j,m}}{h^2} + \frac{u_{i,j+1,m} - 2u_{i,j,m} + u_{i,j-1,m}}{h^2}$$

implisitt skjema:

$$\frac{u_{i,j,m+1} - u_{i,j,m}}{k} = \frac{u_{i+1,j,m+1} - 2u_{i,j,m+1} + u_{i-1,j,m+1}}{h^2} + \frac{u_{i,j+1,m+1} - 2u_{i,j,m+1} + u_{i,j-1,m+1}}{h^2}$$

og Crank-Nicolson

$$\frac{u_{i,j,m+1} - u_{i,j,m}}{k} = \frac{u_{i+1,j,m} - 2u_{i,j,m} + u_{i-1,j,m}}{h^2} + \frac{u_{i,j+1,m} - 2u_{i,j,m} + u_{i,j-1,m}}{h^2} + \frac{u_{i+1,j,m+1} - 2u_{i,j,m+1} + u_{i-1,j,m+1}}{h^2} + \frac{u_{i,j+1,m+1} - 2u_{i,j,m+1} + u_{i,j-1,m+1}}{h^2}$$

Det går an å skrive disse skjemaene om til matrisevektor-iterasjoner på samme vis som i det endimensjonale tilfellet. For eksplisitt skjema får vi

$$\mathbf{u}_{m+1} = \left(I - \frac{k}{h^2} A \right) \mathbf{u}_m + \frac{k}{h^2} \mathbf{f}$$

der A er matrisen fra diskretiseringen av Laplaces likning, og

$$\mathbf{f} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ f(x_1) \\ \vdots \\ f(x_{n-1}) \end{bmatrix}$$

Likninger for implisitt skjema og Crank-Nicolson blir henholdsvis

$$\left(I + \frac{k}{h^2} A \right) \mathbf{u}_{m+1} = \mathbf{u}_m + \frac{k}{h^2} \mathbf{f}$$

og

$$\left(I + \frac{k}{2h^2} A \right) \mathbf{u}_{m+1} = \left(\left(I - \frac{k}{2h^2} A \right) \mathbf{u}_m + \frac{k}{2h^2} \mathbf{f} \right) + \frac{k}{2h^2} \mathbf{f}.$$