Approksimation af løsninger til systemer af første ordens differentialligninger

Anvendt på tennisbold med topspin



Kort Afgangsprojekt (4. semester, MSC) Majbritt Sloth Thomassen Matematik & Statistik Aalborg Universitet 9. januar 2014

Institut for Matematiske Fag

Fredrik Bajers Vej 7G 9220 Aalborg Ø Telefon 99 40 99 40 Fax 98 15 81 29 http://www.math.aau.dk

Synopsis:

Titel: Approksimation af løsninger til systemer af første ordens differentialligninger - Anvendt på tennisbold med topspin Projektperiode: 01.09.13 til 10.01.14 Studieretning: Cand.scient. i matematik og idræt Veileder: Arne Jensen, Institut for Matematiske Fag **Oplagstal:** 4 Sidetal: 71 **Bilagsantal:** 2 Afsluttet d. 9. januar 2014

Dette projekt omfatter approksimation af løsninger til systemer af ordinære første ordens differentialligninger. Metoderne anvendt er Eulers metode, Forbedret Euler, Heuns metode og Runge-Kutta metoden af orden fire, RK4. Peanos eksistenssætning og Osgoods entydighedssætning bevises. Brugen og effektiviteten af de numeriske metoder, illustreres ved eksempler, hvor den generelle løsning er kendt. Derudover udføres numeriske eksperimenter, med henblik på at sammenligne metoderne. Et større eksempel gennemregnes afslutningsvis. Dette eksempel omhandler hvorledes banekurven for en tennisbold påvirkes, når tennisbolden påføres et topspin.

Forfatter:

Majbritt Sloth Thomassen

Rapportens indhold er frit tilgængeligt, men offentliggørelse (med kildeangivelse) må kun ske efter aftale.

Abstract

This project deals with approximation of solutions for systems of ordinary first-order differential equations. Peano's existence theorem and Osgood's uniqueness theorem will be proved. The following methods; Euler's method, the Improved Euler method and Heun's method, will be explained and supported by underlying theory. A short general introduction to methods based on Taylor expansion will be made. Furthermore Runge-Kutta methods will be introduced, among these the classical Runge-Kutta formula of order 4, RK4. Numerical experiments concerning Euler's method, the Improved Euler method, Heun's method and RK4, will be carried out, to illustrate the underlying theory. Throughout the project the use and effectiveness of the mentioned methods will be illustrated by small examples, in which the general solution is known. Finally an example will be carried out, in which no general solution exists. This example shows the effects a topspin will have on the trajectory of a tennisball.

Forord

Dette projekt er et Kandidatspeciale for studieretning i matematik (hovedfag) og idræt (tilvalgsfag) ved Aalborg Universitet. Projektet beskræftiger sig med sædvanlige differentialligninger, med primær fokus approksimation af løsninger, samt eksistens og entydighed af løsninger. Projektperioden er forløbet fra 01.09.13 til 10.01.14.

Der rettes stor tak til vejleder Arne Jensen for god og konstruktiv kritik under hele forløbet. Derudover tak til Tanja Sloth Thomassen for hjælp til forsidebillede.

Læsevejledning

Efter hvert kapitel vil det fremgå hvilke referencer, der er anvendt i det pågældende kapitel. Bagerst i rapporten forefindes en litteraturliste.

Kildehenvisninger er angivet efter Harvardmetoden, og noteres med [Efternavn, år, sidetal]. I litteraturlisten er kilderne angivet med forfatter, år, titel, forlag og ISBN-nummer og derudover URL for internetsider. Figurer og tabeller er nummereret i henhold til kapitel, således at den første figur i kapitel 2 har nummer 2.1, den næste, nummer 2.2 osv.

Matematiske beviser vil blive afsluttet med \Box og eksempler afsluttet med \triangle .

Alle implementeringer er foretaget i programmet Maple 17.00.

Indholdsfortegnelse

Kapitel 1 Introduktion 1.1 Tennisbold med topspin	1 2
Kapitel 2 Eulers metode 2.1 Lokal beskrivelse af one-step metoder 2.2 Eulers metode 2.3 Implementering	7 7 8 11
Kapitel 3 Eksistens og entydighed 3.1 Peanos eksistenssætning 3.2 Entydighed af løsning	15 15 20
Kapitel 4 Metoder baseret på Taylorudvikling	25
Kapitel 5 Forbedring af Eulers metode 5.1 Forbedret Euler 5.2 Heuns metode 5.3 Two-stage metoder 5.4 Implementering Kapitel 6 Runge-Kutta metoder 6.1 RK4	27 27 28 28 30 37 38
Kapitel 7 Numeriske eksperimenter	· · 38 43
7.1 Algoritmer	43
Kapitel 8Tennisbold med topspin8.1Formen af C_D og C_M 8.2Approksimationsmetoderne8.3Tennisbold i vakuum8.4Tennisbold i luft	53 53 53 54 55 55 57
Kapitel 9 Afrunding	61
Bilag A Appendiks A.1 Definitioner A.2 Partielle afledede	63 63 63
Bilag B Algoritmer B.1 Forholdet mellem fejl i endepunkt B.2 Forholdet mellem maksimal fejl	65 65 67
Litteratur	71

Introduktion

Differentialligninger spiller ofte en vigtig rolle indenfor videnskaben. De differentialligninger, der behandles i dette projekt, er på formen

$$\frac{dy}{dx} = f(x, y) \tag{1.1}$$

hvor y er en funktion fra et interval $I \subseteq \mathbb{R}$ ind i \mathbb{R}^d , hvor *d* er et positivt heltal. En **løsning** til (1.1) er en funktion $\phi(x) : I \to \mathbb{R}^d$, der opfylder at

$$\phi'(x) = f(x,\phi(x))$$
 , $x \in I$

Et simpelt eksempel på en sådan differentialligning er positionen af en partikel. En partikel bevæger sig langs *x*-aksen med hastighed givet ved den kontinuerte funktion f(t) til et givet tidspunkt *t*. Ved tidspunktet t_0 er partiklens position x_0 . Hvis en fysiker ønsker at beskrive partiklens bevægelse, skal han dermed løse differentialligningen

$$\frac{dx}{dt} = f(t) \tag{1.2}$$

hvor løsningen antager værdien x_0 ved tispunktet t_0 . Løsningen er givet ved

$$x(t) = x_0 + \int_{t_0}^t f(\tau) d\tau$$

Dette ses tydeligt ved direkte indsættelse i (1.2). Fysikeren kan således bestemme partiklens position til ethvert tidspunkt *t*. En differentialligning som (1.1) er ikke i sig selv nok til bestemmelse af en entydig løsning. Der er derfor brug for supplerende betingelser, som i eksemplet med partiklen, hvor det var givet at positionen ved tidspunktet t_0 var x_0 . En sådan ekstra betingelse kaldes en **begyndelsesbetingelse**. Eksistens og entydighed af en løsning til en given differentialligning, er yderst relevante emner at tage op. I kapitel 3 behandles disse emner derfor, i form af to vigtige resultater: *Peanos eksistenssætning* og *Osgoods entydighedssætning*.

Det er desværre ikke altid muligt at bestemme en generel løsning til en differentialligning. I sådanne tilfælde er det derfor nødvendigt at anvende en eller flere numeriske metoder til at approksimere en løsning. Det er netop disse "uløselige" differentialligninger og metoder til at approksimere en løsning dertil, der er omdrejningspunktet for dette projekt. Der vil blive gennemgået følgende approksimationsmetoder

- Eulers metode (kapitel 2)
- Metoder baseret på Taylorudvikling (kapitel 4)
- Forbedret Eulers metode (kapitel 5)
- Heuns metode (kapitel 5)
- Runge-Kutta metoder, herunder RK4. (kapitel 6)

Der er naturligvis forskel på, hvor effektive de forskellige metoder er, hvilket derfor berøres undervejs for hver enkelt metode. Brugen og effektiviteten af de forskellige metoder vil undervejs blive illustreret vha. mindre eksempler, hvor den eksakte løsning er kendt, således at der er noget at sammenholde de approksimerede løsninger med. Derudover vil et større eksempel afslutningsvist blive gennemregnet. En introduktion til dette eksempel er givet nedenfor.

1.1 Tennisbold med topspin

Et af de steder, hvor disse "uløselige" differentialligninger opstår, er når der sættes spin på en tennisbold. Når en tennisspiller udfører et slag med topspin tilfører han bolden en hastighed, i form af en vektor v, rettet mod modspillerens banehalvdel. Derudover tilfører han bolden en vinkelhastighed, i form af en vektor w, der får bolden til rotere. Boldens placering under turen fra den ene banehalvdel til den anden, angives som en positionsvektor, r. Der er tre forskellige kræfter, der påvirker tennisbolden under flyvningen, disse er:

- Tyngdekraften F_t
- Luftmodstanden D
- Magnuskraften M

Disse er illustreret på figur 1.1



Figur 1.1. Illustration af de kræfter, der påvirker en roterende tennisbold.

Tyngdekraften

Tyngdekraften virker altid lodret på et legeme og er, når den angives som en vektor i 3 dimensioner, givet ved

$$F_t = mg$$

hvor *m* er tennisboldens masse og *g* er en vektor, der angiver tyngdeaccelerationen, dvs. $g = (0;0;-9,82)^T$.

Luftmodstanden

Luftmodstanden, D, for et legeme virker i den modsatte retning af hastighedsvektoren v, og er givet ved

$$D = -D_L(v)\frac{v}{\|v\|}$$

hvor D_L bestemmer størrelsen af luftmodstanden, og er på formen

$$D_L = C_D \frac{1}{2} A \rho \|v\|^2 = C_D \frac{\pi d^2}{8} \rho \|v\|^2$$

hvor ρ er luftmassefylde, d er tennisboldens diameter og A er tværsnitsarealet af det givne legeme, dvs. for en tennisbold lig med $\pi d^2/4$. Koefficienten C_D kan findes eksperimentielt og afhænger for en tennisbold af v og w. Formen af C_D bliver behandlet i kapitel 8.

Magnuskraften

Magnuskraften, M, er ortogonal på hastighedsvektoren v og vinkelhastighedsvektoren w, og er givet ved

$$M = M_L \frac{w}{\|w\|} \times \frac{v}{\|v\|}$$

hvor M_L bestemmer størrelsen af magnuskraften, og er på formen

$$M_L = C_M \frac{1}{2} A \rho \|v\|^2 = C_M \frac{\pi d^2}{8} \rho \|v\|^2$$

og afhænger således af de samme parametre som luftmodstanden, blot med koefficient C_M i stedet for C_D . C_M afhænger ligeledes af v og w og kan findes eksperimentielt.

Differetialligninger for positionsvektoren

For at få opstillet et system af differentialligninger, der beskriver positionsvektorer *r*, skal anvendes Newtons anden lov. Denne siger at:

Hvis et objekt med masse m bliver påvirket af en ydre kraft F_{net} , vil denne forårsage en acceleration, a, af objektet i samme retning, som den ydre kraft:

$$F_{net} = ma$$

Den anden ordens afledede af tennisboldens positionsvektor, mht. til tiden *t*, er lig med accelerationsvektoren for tennisbolden, således kan Newtons anden lov anvendes til at opstille følgende differentialligning for tennisbolden

$$m\frac{d^2r}{dt^2} = F_t + D + M \tag{1.3}$$

Indsættes formlerne for F_t , D og M, fås

$$m\frac{d^2r(t)}{dt^2} = F_t - D_L \frac{v}{\|v\|} + M_L \frac{w}{\|w\|} \times \frac{v}{\|v\|}$$
$$= F_t - C_D \frac{\pi d^2}{8} \rho \|v\|^2 \frac{v}{\|v\|} + C_M \frac{\pi d^2}{8} \rho \|v\|^2 \frac{w}{\|w\|} \times \frac{v}{\|v\|}$$

så

$$\frac{d^2 r(t)}{dt^2} = g - C_D \frac{\pi d^2 \rho \|v\|}{8m} v + C_M \frac{\pi d^2 \rho \|v\|}{8\|w\|m} w \times v$$
(1.4)

Ligningen (1.4) er et system af tre anden ordens differentialligninger med begyndelsesbetingelserne

$$r(0) = r_0 \quad \text{og} \quad \frac{dr}{dt}(0) = v_0$$
 (1.5)

hvor v_0 er tennisboldens starthastighed. Systemet (1.4) har ingen generel løsning og skal derfor løses numerisk. I praksis kan problemet betragtes i to dimensioner i stedet for 3. Dette gøres ved at antage, at vinkelhastighedsvektoren w ligger i det horisontale plan, parallelt med y-aksen og er vinkelret på hastighedsvektoren v. For en illustration af dette se figur 1.2. Bemærk at angivelse af x- og y-akse er byttet for figur 1.1 og figur 1.2.



Figur 1.2. Illustration af magnuskraft ved topspin og bagspin.

Disse antagelser gør, at magnuskraften vil ligge i det vertikale plan, da den som tidligere nævnt er ortogonal på v og w. Retningen af magnuskraften bestemmes af, hvorvidt tennisbolden påføres et topspin eller et bagspin, som illustreret på figur 1.2, hvor M_{Top} angiver magnuskraftens retning for topspin og M_{Bag} for bagspin. Angives hastighedsvektoren ved

$$v = \left[\begin{array}{c} v_x \\ v_z \end{array} \right]$$

Da kan de to mulige magnuskrafter skrives som

$$M_{Top} = M_L \frac{1}{\|v\|} \begin{bmatrix} v_z \\ -v_x \end{bmatrix} = M_L \frac{v_{Top}}{\|v\|}$$
$$M_{Bag} = M_L \frac{1}{\|v\|} \begin{bmatrix} -v_z \\ v_x \end{bmatrix} = M_L \frac{v_{Bag}}{\|v\|}$$

Tilsvarende til (1.4), fås der i to dimensioner med topspin

$$\frac{d^2 r(t)}{dt^2} = g^T - C_D \frac{\pi d^2 \rho \|v\|}{8m} v + C_M \frac{\pi d^2 \rho \|v\|}{8m} v_{Top}$$
(1.6)

og med bagspin

$$\frac{d^2 r(t)}{dt^2} = g^T - C_D \frac{\pi d^2 \rho \|v\|}{8m} v + C_M \frac{\pi d^2 \rho \|v\|}{8m} v_{Bag}$$
(1.7)

Indfør nu et β , hvor $\beta = 1$ når tennisbolden er udsat for topspin og $\beta = -1$ ved bagspin. Indfør derudover et α givet ved

$$\alpha = \frac{\pi d^2 \rho}{8m}$$

Så kan (1.6) og (1.7) skrives som et system af to anden ordens differentialligninger

$$\frac{d^{2}x(t)}{dt^{2}} = -C_{D}\alpha \|v\| \frac{dx(t)}{dt} + \beta C_{M}\alpha \|v\| \frac{dz(t)}{dt}$$

$$\frac{d^{2}z(t)}{dt^{2}} = -9,82 - C_{D}\alpha \|v\| \frac{dz(t)}{dt} - \beta C_{M}\alpha \|v\| \frac{dx(t)}{dt}$$
(1.8)

hvor $||v|| = \sqrt{x'(t)^2 + z'(t)^2}$. For at definere begyndelsesbetingelser, lad da v_0 angive vektoren for begyndelseshastigheden og θ angive vinklen mellem v_0 og *x*-aksen. Da haves følgende begyndelsesbetingelse hørende til (1.8)

$$x(0) = 0, \quad z(0) = H, \quad \frac{dx}{dt}(0) = ||v_0||\cos(\theta), \quad \frac{dz}{dt}(0) = ||v_0||\sin(\theta)$$

hvor *H* angiver tennisboldens starthøjde. Dermed en introduktion til det problem, der skal løses i denne rapport. Inden dette kan gøres skal teorien bag de forskellige approksimationsmetoder gennemgås.

Referencer

Introduktionen er inspireret af [Petrovski, 1973, s. 3-4]. Afsnit 1.1 er primært skrevet på baggrund af [Gander og Hřebíček, 2004, s. 27-29], hvor teorien omhandlende vindmodstand er uddybet vha. [Faber, 1995, s. 264]. Derudover er Newton anden lov skrevet vha. [Bauer og Westfall, 2011, s. 107].

Eulers metode 2

Eulers metode er en forholdsvis simpel metode til at approksimere en løsning til en differentialligning. Eulers metode er en såkaldt *one-step metode*. Indledningsvis indføres derfor relevante definitioner for one-step metoder.

2.1 Lokal beskrivelse af one-step metoder

Det ønskes at approksimere en løsning til begyndelsesværdiproblemet

$$\frac{dy}{dx} = f(x,y), \quad a \le x \le b, \quad y(a) = y_a \tag{2.1}$$

Givet et punkt $x_k \in [a, b]$ og et $y_k \in \mathbb{R}^d$, skrives et enkelt skridt i en one-step metode generelt på formen

$$y_{k+1} = y_k + h\Phi(x_k, y_k; h)$$
(2.2)

hvor h > 0 angiver skridtlængden. Funktionen Φ i (2.2) er altså definerende for approksimationsmetoden og angiver den approksimerede tilvækst per skridt. Udover vektoren y_{k+1} betragtes løsningen u(t) til (2.1), der passerer igennem punktet (x_k, y_k) , således at der haves et lokalt begyndelsesværdiproblem givet ved

$$\frac{du}{dt} = f(t, u), \quad x_k \le t \le x_k + h, \quad u(x_k) = y_k$$
(2.3)

hvor u(t) kaldes *referenceløsningen*. Vektoren y_{k+1} har til formål at approksimere $u(x_k + h)$. For at kunne vurdere hvor god en approksimation er, indføres begrebet trunkeringsfejl.

DEFINITION 2.1

Lad $(x_k, y_k) \in [a, b] \times \mathbb{R}^d$. Den **lokale trunkeringsfejl** for approksiamtionsmetoden Φ i punktet (x_k, y_k) er defineret ved

$$T(x_k, y_k; h) = \frac{1}{h}(y_{k+1} - u(x_k + h))$$
(2.4)

Indsættes udtrykket (2.2) for y_{k+1} i (2.4) fås

$$T(x_k, y_k; h) = \frac{1}{h} (y_k + h\Phi(x_k, y_k; h) - u(x_k + h))$$

$$= \Phi(x_k, y_k; h) + \frac{1}{h} (y_k - u(x_k + h))$$

$$= \Phi(x_k, y_k; h) - \frac{1}{h} (u(x_k + h) - u(x_k))$$
(2.5)

Den lokale trunkeringsfejl, $T(x_k, y_k; h)$, er altså differensen mellem den approksimerede funktionstilvækst og den eksakte tilvækst per skridt. Det vil sige størrelsen af den fejl, der opstår hver gang der foretages et enkelt skridt under approksimationen af løsningen til (2.1). Den næste definition anvendes til at afgøre om en given approksimationsmetode er konsistent.

DEFINITION 2.2 Approksimationsmetoden Φ kaldes **konsistent** hvis

$$T(x_k, y_k; h) \to 0 \tag{2.6}$$

ligeligt for alle $(x_k, y_k) \in [a, b] \times \mathbb{R}^d$ når $h \to 0$.

Bemærk at anvendes (2.3) vil $\frac{1}{h}(u(x_k + h) - u(x_k)) \rightarrow f(x_k, y_k)$ når $h \rightarrow 0$, så betragtes omskrivningen (2.5), er en approksimationsmetode Φ konsistent hvis og kun hvis

$$\Phi(x_k, y_k; 0) = f(x_k, y_k) \text{ for alle } x_k \in [a, b] \text{ og } y_k \in \mathbb{R}^d$$
(2.7)

DEFINITION 2.3 Approksimationsmetoden Φ siges at have **orden** *p*, hvis der eksisterer en konstant *C* således at

$$|T(x_k, y_k; h)|| \le Ch^p \tag{2.8}$$

ligeligt på $[a, b] \times \mathbb{R}^d$. Φ siges at have **eksakt orden** p hvis (2.8) ikke holder for højere værdier end p.

Egenskaben (2.8) i definitionen ovenfor skrives $T(x_k, y_k; h) = O(h^p), h \to 0$. Bemærk derudover at (2.8) og p > 0 vil medføre at metoden Φ er konsistent.

DEFINITION 2.4 En funktion $\sigma : [a, b] \times \mathbb{R}^d \to \mathbb{R}^d$, der opfylder betingelserne

1. $\sigma \neq 0$ 2. $T(x_k, y_k; h) = \sigma(x_k, y_k)h^p + O(h^{p+1})$, $h \rightarrow 0$

kaldes den principielle fejlfunktion.

2.2 Eulers metode

Eulers metode er som sagt en one-step metode. Her er funktionen Φ i (2.2) givet ved

$$\Phi(x,y;h) = f(x,y)$$

Så givet et punkt $x_k \in [a, b]$ og et $y_k \in \mathbb{R}^d$, skrives et enkelt skridt for Eulers metode

$$y_{k+1} = y_k + hf(x_k, y_k)$$
(2.9)

Euler foreslog denne metode i 1768 baseret på simple geometriske overvejelser, hvor hældningen $f(x_k, y_k)$ i et givet punkt (x_k, y_k) følges over et interval med længde *h*. Derefter evalueres hældningen i punktet $(x_{k+1} = x_k + h, y_{k+1})$ og følges til næste punkt (x_{k+2}, y_{k+2}) , osv. Før der tages hul på en generel behandling af Eulers metode, illustreres metoden med et eksempel.

EKSEMPEL 2.5

Det ønskes at approksimere en løsning til begyndelsesværdiproblemet

$$\frac{dy}{dx} = 4x^3y, \qquad 0 \le x \le 1, \qquad y_0 = 1$$
 (2.10)

Approksimationen vælges udført med 6 skridt, så $h = \frac{1-0}{6} = \frac{1}{6}$. Dermed er et skridt i approksimationen på formen

$$y_{k+1} = y_k + \frac{1}{6} \cdot 4 \cdot x_k^3 \cdot y_k$$
$$= y_k + \frac{2}{3} \cdot x_k^3 \cdot y_k$$

hvor

$$x_k = x_0 + hk = hk$$

Så det første skridt i approksimationen bliver

$$y_1 = y_0 + \frac{1}{3} \cdot x_0^3 \cdot y_0 = 1 + \frac{2}{3} \cdot 0^3 \cdot 1 = 1$$

. 3

osv.

$$\begin{array}{rcl} y_2 &= y_1 + \frac{1}{3} \cdot x_1^3 \cdot y_1 &= 1 + \frac{2}{3} \cdot \left(\frac{1}{6}\right)^5 \cdot 1 &= \frac{325}{324} &\approx 1,0031 \\ y_3 &= y_2 + \frac{1}{3} \cdot x_2^3 \cdot y_2 &= \frac{325}{324} + \frac{2}{3} \cdot \left(\frac{2}{6}\right)^3 \cdot \frac{325}{324} &= \frac{26975}{326244} &\approx 1,0279 \\ y_4 &= y_3 + \frac{1}{3} \cdot x_3^3 \cdot y_3 &= \frac{26975}{26244} + \frac{2}{3} \cdot \left(\frac{3}{6}\right)^3 \cdot \frac{26975}{26244} &= \frac{350675}{314928} &\approx 1,1135 \\ y_5 &= y_4 + \frac{1}{3} \cdot x_4^3 \cdot y_4 &= \frac{350675}{314928} + \frac{2}{3} \cdot \left(\frac{4}{6}\right)^3 \cdot \frac{350675}{14928} &= \frac{34015475}{25509168} &\approx 1,3334 \\ y_6 &= y_5 + \frac{1}{3} \cdot x_5^3 \cdot y_5 &= \frac{34015475}{25509168} + \frac{2}{3} \cdot \left(\frac{5}{6}\right)^3 \cdot \frac{34015475}{25509168} &= \frac{15272948275}{8264970432} &\approx 1,8479 \end{array}$$

Approksimationsværdierne for punkterne (x_k , y_k), for k = 0, 1, ..., 6, er plottet på figur 2.1 sammen med den eksakte løsning $y = exp(x^4)$.



Figur **2.1.** Approximation af løsning til begyndelsesværdiproblemet (2.10), foretaget vha. Eulers metode med 6 skridt, samt den eksakte løsning $y = exp(x^4)$.

Dette eksempel illustrerer brugen af Eulers metode. Det er dog nødvendigt at anvende mere end 6 skridt for at opnå en tilfredstillende approksimation af det pågældende begyndelsesværdiproblem. Eksempler med et højere antal skridt vil forekomme senere i rapporten. \triangle

I Eulers metode er $\Phi(x_k, y_k; h) = f(x_k, y_k)$, så da $f(x_k, y_k)$ ikke afhænger af h, må Eulers metode være konsistent jævnfør (2.7). Betragtes omskrivningen af den lokale trunkeringsfejl (2.5) og anvendes det at $u'(x_k) = f(x_k, u(x_k)) = f(x_k, y_k)$, haves

$$T(x_k, y_k; h) = f(x_k, y_k) - \frac{1}{h} (u(x_k + h) - u(x_k))$$

= $u'(x_k) - \frac{1}{h} (u(x_k + h) - u(x_k))$ (2.11)

Antag nu at $f \in C^1[a, b] \times \mathbb{R}^d$. Så må $u \in C^2[x_k, x_k + h]$. Dermed kan Taylors sætning (se sætning A.4) anvendes på $u(x_k + h)$, så

$$u(x_k + h) = u(x_k) + u'(x_k)(x_k + h - x_k) + \frac{u''(\xi)}{2!}(x_k + h - x_k)^2$$
$$= u(x_k) + hu'(x_k) + \frac{h^2}{2}u''(\xi)$$

for et $\xi \in (x_k, x_k + h)$. Ved denne notation bemærkes det at dette ξ kun er et enkelt tal hvis *u* er en funktion ind i \mathbb{R} . Hvis *u* er en vektorfunktion, haves et $\xi \in (x_k, x_k + h)$, for hver af indgangene i *u*. Denne notation vil dog blive anvendt i det følgende. Indsættes dette i (2.11), fås

$$T(x_k, y_k; h) = u'(x_k) - \frac{1}{h} \left(u(x_k) + hu'(x_k) + \frac{h^2}{2} u''(\xi) - u(x_k) \right)$$
(2.12)
= $-\frac{h}{2} u''(\xi)$

for et $\xi \in (x_k, x_k + h)$. Betragt nu igen det lokale begyndelsesværdiproblem (2.3). Differentieres (2.3) vha. kædereglen mht. t, fås

$$\frac{d^2u}{dt} = [f_x + f_y f](t, u(t))$$

hvor f_x er den partielle aflede af f mht. x og f_y er Jacobimatricen af f mht. y. Sæt nu $t = \xi$, så (2.12) bliver

$$T(x_k, y_k; h) = -\frac{1}{2}h[f_x + f_y f](\xi, u(\xi)), \quad \xi \in (x_k, x_k + h)$$
(2.13)

Omskrivningen (2.13) af den lokale trunkeringfejl, viser, at hvis alle de partielle afledede af f er ligeligt begrænsede i $[a, b] \times \mathbb{R}^d$, så findes en konstant *C*, således at

$$||T(x_k, y_k; h)|| \le Ch$$

Hvilket medfører at Eulers metode er af orden 1, jævnfør definition 2.3, såfremt de nævnte antagelser vedrørende de partielle afledede holder. Hvis det ydermere antages, at alle de anden ordens partielle afledede af *f* er ligeligt begrænsede i $[a, b] \times \mathbb{R}^d$, findes et \tilde{C} således at

$$u''(\xi) = u''(x_k) + \widetilde{C}h$$

da $\xi \in (x_k, x_k + h)$. Indsættes dette i (2.12), fås

$$T(x_k, y_k; h) = -\frac{h}{2} (u''(x_k) + \widetilde{C}h)$$

$$= -\frac{h}{2} [f_x + f_y f](x_k, y_k) - \frac{h^2}{2} \widetilde{C}$$

$$= -\frac{h}{2} [f_x + f_y f](x_k, y_k) + O(h^2) , \quad h \to 0$$
(2.14)

Af (2.14) ses det at den principielle fejlfunktion for Eulers metode er givet ved

$$\sigma(x_k, y_k) = -\frac{h}{2}[f_x + f_y f](x_k, y_k)$$

Så Eulers metode har eksakt orden 1, medmindre at $f_x + f_y f \equiv 0$.

2.3 Implementering

Euler metode er implementeret i algoritme 1. Denne algoritme anvender følgened inputs

• *f* - højresiden i differentialligning $\frac{dy}{dx} = f(x, y)$

ALGORITME 1 (Eulers metode med *n* inddelinger).

- *xnul*, *ynul* begyndelsesbetingelserne ynul = y(xnul).
- *a*, *b* det interval [*a*, *b*] løsningen skal approksimeres over.
- *n* antallet af skridt der anvendes.

eulersmetode := proc(f, xnul, ynul, a, b, n) $h := \frac{b-a}{n};$ x[0] := xnul; y[0] := ynul; Lx := [x[0]]; Ly := [y[0]]; for i from 0 to n - 1 do x[i+1] := x[i] + h; $y[i+1] := y[i] + h \cdot f(x[i], y[i]);$ Lx := [op(Lx), x[i+1]]; Ly := [op(Ly), y[i+1]] end do; $return \{Lx, Ly\}$ end proc;

Outputtet af algoritme 1, er en et sæt indeholdende en liste $Lx = [x_0, x_1, ..., x_n]$ og en liste med approksimationerne $Ly = [y_0, y_1, ..., y_n]$. Algoritmen kan nu anvendes til at udføre eksempler med et højere antal skridt *n*.

EKSEMPEL 2.6

Det ønskes at approksimere en løsning til begyndelsesværdiproblemet fra eksempel 2.1

$$\frac{dy}{dx} = 4x^3y, \qquad 0 \le x \le 1, \qquad y_0 = 1$$
 (2.15)

I Maple skal funktionen f(x, y) defineres med kommandoen

$$f := (x, y) \to 4x^3y$$

Derefter kan løsninger med det ønskede antal inddelinger approksimeres vha. Eulers metode vha. algoritme 1. Ved fire inddelinger anvendes kommandoen eulersmetode(f, 0, 1, 0, 1, 4), hvorefter der fås

$$\left\{ \left[0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\right], \left[1, 1, \frac{65}{64}, \frac{585}{512} \frac{53235}{32768}\right] \right\}$$

På figur 2.2 er approksimationer med henholdsvis 4, 8, 16 og 32 inddelinger plottet sammen med den eksakte løsning



Figur 2.2. Approximation af løsning til begyndelsesværdiproblemet (2.15), foretaget vha. Eulers metode med 4,8,16 og 32 skridt, samt den eksakte løsning $y = exp(x^4)$.

På tabel 2.1 ses approksimerede værdier i x = 1, samt afvigelse fra den eksakte i x = 1.

п	Approksimation	Eksakt	Fejl
4	1.6246	2.7183	1.0937
8	1.9955	2.7183	0.7228
16	2.2874	2.7183	0.4309
32	2.4799	2.7183	0.2384

Tabel 2.1. Approksimeret værdi, eksakt værdi og approksimationsfejl i x = 1.

Eksemplet illustrerer hvorledes de approksimerede løsninger lader til at konvergere mod den eksakte løsning, når antallet af skridt n øges.

Det er ikke altid tilfældet, at numeriske approksimationsmetoder fungerer så godt som i eksempel 2.6. I det næste eksempel haves et begyndelsesværdiproblem, hvor Eulers metode bliver sat på prøve.

EKSEMPEL 2.7 Betragt begyndelsesværdiproblemet

$$\frac{dy}{dx} = xy^{\frac{3}{2}}, \qquad 0 \le x \le 3, \qquad y_0 = 1$$
 (2.16)

Den eksakte løsning til begyndelsesværdiproblemet (2.16) er givet ved

$$y = \frac{16}{(x^2 - 4)^2} \tag{2.17}$$

Så i intervallet [0,3] er løsningen altså ikke defineret i x = 2. Dette er problematisk når løsningen ønskes approksimeret, da den approksimerede løsning ikke vil være istand til at passere x = 2 på en hensigtsmæssig måde. Algoritme 1 er anvendt til approksimere løsninger til (2.16) med henholdsvis 4,8 og 16 inddelinger. Disse er plottet på figur 2.3 sammen med den eksakte løsning (2.17).



Figur 2.3. Approksimation af løsning til begyndelsesværdiproblemet (2.16), foretaget vha. Eulers metode med 4,8 og 16 skridt, samt den eksakte løsning (2.17)

Af figur 2.3 ses det, hvordan de approksimerede løsninger ikke formår at passere x = 2 hensigtsmæssigt. I tabel 2.2 er de approksimerede løsninger og den eksakte løsning (2.17) evauleret i x = 3. Derudover er forskellen mellem disse angivet, altså fejlen.

п	Approksimation	Eksakt	Fejl
4	16.0620	0.6400	-15.4220
8	258.172	0.6400	-257.532
16	$6.85592 \cdot 10^5$	0.6400	$-6.85592 \cdot 10^5$

Tabel 2.2. Approksimeret værdi, eksakt værdi og approksimationsfejl i x = 3.

Af tabel 2.2 fremgår det, at størrelsen af fejlen i dette tilfælde ikke bliver mindre, når antallet af skridt øges. \triangle

Referencer

Dette kapitel er skrevet på baggrund af [Gautschi, 1997, s, 272-275].

Eksistens og entydighed B

I dette kapitel behandles eksistens og entydighed af løsninger til differentialligninger. I forhold til eksistensen af en løsning bevises Peanos eksistenssætning. Entydighed af løsning belyses ved at bevise Osgoods entydighedssætning. Det bemærkes at disse resultater omhandlende, eksistens og entydighed af løsninger til differentialligningen y' = f(x, y), kun vises for tilfældet hvor $y \in \mathbb{R}$.

3.1 Peanos eksistenssætning

For at forstå og bevise Peanos eksistenssætning, skal to relevante begreber defineres.

DEFINITION 3.1

En familie af funktioner $\{f(x)\}$ siges at være **ligeligt begrænsede** på intervallet [a, b], hvis der eksisterer en konstant *M* således at

|f(x)| < M

for alle $f(x) \in \{f(x)\}$ og alle $x \in [a, b]$.

DEFINITION 3.2 Lad $U \subseteq \mathbb{R}^n$. En funktion $f : U \to \mathbb{R}^n$ siges at være **uniform ækvikontinuert** på *U* hvis der for alle $\epsilon > 0$ eksisterer et $\eta = \eta(\epsilon) > 0$ således at

 $||x - y|| < \eta$ medfører at $||f(x) - f(y)|| < \epsilon$

for alle $x, y \in U$.

Beviset for Peanos eksistenssætning anvender en anden vigtig sætning af Arzelá og Ascoli, der derfor bevises i det følgende.

SÆTNING 3.3 (Arzelà-Ascoli)

Lad $\{f(x)\}$ være en uendelig familie af funktioner f(x), der for alle $x \in [a, b]$ er uniformt ækvikontinuerte og ligeligt begrænsede. Da findes en uendelig ligelig konvergent delfølge $\{\tilde{f}(x)\} \subseteq \{f(x)\}$.

BEVIS

Da familien $\{f(x)\}$ er ligeligt begrænset, findes en konstant *M* således at for enhver funktion $f(x) \in \{f(x)\}$ er |f(x)| < M for alle $x \in [a, b]$. Dermed må enhver funktion $f(x) \in \{f(x)\}$ ligge i rektanglet med bredde b - a og højde 2*M*, se eventuelt figur 3.1.

Lad α være et ikkenegativt heltal og konstruer den uendelige følge

$$\epsilon_1 = \frac{M}{2^{\alpha+1}}, \ \epsilon_2 = \frac{M}{2^{\alpha+2}}, \ ..., \ \epsilon_k = \frac{M}{2^{\alpha+k}}, \ ...$$

Lad ydermere $\eta_k = \eta(\epsilon_k)$ være tallet tilknyttet ϵ_k , som i definitionen af ækvikontinuitet (se definition 3.2).

Lad nu $\alpha = 1$ og inddel rektanglet i mindre rektangler, hver med højde ϵ_1 og bredde η_1 som på figur 3.1.





Figur 3.2. Illustration, hvor det skraverede område i kolonne 1 vil indeholde uendelig mange funktioner $f(x) \in \{f(x)\}$.

Figur 3.1. Inddeling af rektanglet med bredde b - a og højde 2*M*.

På grund den uniforme ækvikontinuitet, gælder det for alle $x, y \in [a, b]$ og $f(x) \in \{f(x)\}$, at

hvis
$$|x - y| < \eta_1$$
 er $|f(x) - f(y)| < \epsilon_1$ (3.1)

Dermed kan ingen af funktionerne i $\{f(x)\}$ strække sig over mere end to rektangler i hver lodrette kolonne. Se f.eks. figur 3.2, hvor det skraverede område i kolonne 1 vil indeholde uendelig mange funktioner $f(x) \in \{f(x)\}$.

Når disse funktioner forlader kolonne 1, kan de kun passere fire rektangler i kolonne 2 (se figur 3.2) og hver enkelt funktion kun igennem 2 regtangler i kolonne 2.



Figur 3.3. Konstruktion af et bånd, b_1 , med højde $2\epsilon_1$.



Figur 3.4. Konstruktion af et bånd, b_2 , med højde $2\epsilon_2$.

Fortsættes denne fremgangsmåde, kan et bånd med højde $2\epsilon_1$ konstrueres, som det på figur 3.3, der indeholder uendelig mange funktioner $f(x) \in \{f(x)\}$. Angiv dette bånd b_1 , og lad $\{f_1(x)\} \subseteq \{f(x)\}$ angive den uendelige familie af funktioner i bånd b_1 .

 ${f_1(x)}$ kan behandles med samme metode som ${f(x)}$, med ϵ_2 og η_2 istedet for henholdsvis ϵ_1 og η_1 . Således konstrueres et bånd b_2 med højde $2\epsilon_2$, som på figur 3.4.

Lad $\{f_2(x)\} \subseteq \{f_1(x)\}$ angive den uendelige familie af funktioner i bånd b_2 . Fortsættes denne konstruktion haves en uendelig følge af funktioner

$$f_1(x), f_2(x), ..., f_k(x), ...$$
 (3.2)

hvor $f_1(x) \in \{f_1(x)\}, f_2(x) \in \{f_2(x)\}, ..., f_k(x) \in \{f_k(x)\}$ og hvor alle funktioner efter $f_k(x)$ er indeholdt i et bånd med højde $2\epsilon_k = \frac{M}{2^{k+\alpha-1}}$. Dermed er følgen (3.2) ligelig konvergent og sætningen er bevist.

SÆTNING 3.4 (Peanos eksistenssætning)

Lad funktionen f(x, y) være begrænset og kontinuert på definitionsmængden G. Da gælder det for ethvert punkt $(x_0, y_0) \in G$ at mindst en integralekurve hørende til differentialligningen

$$\frac{dy}{dx} = f(x, y) \tag{3.3}$$

passerer igennem punktet $(x_0, y_0) \in G$.

BEVIS

Da f(x, y) er begrænset, findes en konstant M således at $|f(x, y)| \le M$. Tegn to linjer med hældning M og -M, der begge passerer igennem punktet $(x_0, y_0) \in G$. Tegn dernæst to linjer x = aog x = b, således at der konstrueres to trekanter indeholdt i G, som illustreret på figur 3.5



Figur 3.5. Konstruktion af to trekanter indeholdt i G.

Konstruer nu en uendelig følge af polygonlinjer, der alle passerer igennem punktet (x_0, y_0) , vha. Eulers metode, som beskrevet i afsnit 2.2. Angiv denne følge

$$L_1, L_2, ..., L_k, ...$$
 (3.4)

hvor længden af linjestykkerne hørende til L_k går mod nul når $k \to \infty$. Polygonlinjerne (3.4) skærer alle linjer parallelle med *y*-aksen en enkelt gang. Dermed er polygonlinjen L_k , k = 1, 2, ..., en graf for en kontinuert funktion af *x*. Angiv denne følge af kontinuerte funktioner

$$\phi_1(x), \phi_2(x), ..., \phi_k(x), ...$$
 (3.5)

funktionerne (3.5) kan kun forlade trekanterne *ABC* og *ADE* på figur 3.5 gennem siderne *BC* og *DE*, da den absolutte værdi for hældningen af L_k er begrænset af *M*, for alle k = 1, 2, ... Dermed er funktionerne (3.5) alle defineret på intervallet [a, b].

Derudover er følgen (3.5) ligeligt begrænset på [a, b], da alle funktionerne indenfor dette interval ligger i trekanterne *ABC* og *ADE*.

Da hældningen af $\phi_1(x)$, $\phi_2(x)$, ..., $\phi_k(x)$, ... er begrænset, er

$$\left|\frac{\phi_k(x_2) - \phi_k(x_1)}{x_2 - x_1}\right| \le M$$

for alle k = 1, 2, ... og $x_1, x_2 \in [a, b]$, så

$$|\phi_k(x_2) - \phi_k(x_1)| \le M |x_2 - x_1|$$

sæt $\eta = \frac{\epsilon}{M}$, så vil $|x_2 - x_1| < \eta$ medføre at

$$|\phi_k(x_2) - \phi_k(x_1)| < \epsilon$$

for alle k = 1, 2, ... og $x_1, x_2 \in [a, b]$. Dermed er følgen (3.5) ækvikontinuert. Dermed opfylder følgen { $\phi_k(x)$ } alle betingelser i sætning 3.2, så der eksisterer en uendelig delfølge

$$\{\widetilde{\phi_k}(x)\}\subseteq\{\phi_k(x)\}$$

der er ligelig konvergent på intervallet [a, b]. Angiv den tilhørende grænsefunktion

$$\phi(x) = \lim_{k \to \infty} \tilde{\phi}_k(x) \tag{3.6}$$

Da alle funktioner i følgen { $\tilde{\phi}_k(x)$ } passerer punktet (x_0, y_0) , opfylder $\phi(x)$ tydeligevis begyndelsesbetingelsen $\phi(x_0) = y_0$.

Det skal nu vises at (3.6) er en løsning til differentialligningen (3.3) i intervallet [a, b]. Betragtes intervallet $[x_0, b]$, skal det altså vises, at for et vilkårligt $x_1 \in [x_0, b]$ og vilkårligt $\epsilon > 0$ er

$$\left|\frac{\phi(x_2) - \phi(x_1)}{x_2 - x_1} - f(x_1, \phi(x_1))\right| \le \epsilon$$

givet at $|x_2 - x_1|$ er tilstrækkelig lille. Da det haves at

$$\lim_{k \to \infty} \widetilde{\phi}_k(x_1) = \phi(x_1) \quad \text{og} \quad \lim_{k \to \infty} \widetilde{\phi}_k(x_2) = \phi(x_2)$$

er det tilstrækkeligt at vise

$$\left|\frac{\widetilde{\phi}_k(x_2) - \widetilde{\phi}_k(x_1)}{x_2 - x_1} - f(x_1, \phi(x_1))\right| < \epsilon$$
(3.7)

for tilstrækkelig stort *k* og tilstrækkelig lille $|x_2 - x_1|$. f(x, y) er kontinuert på *G*, så givet et $\epsilon > 0$ og punkter (x, y), $(x_1, y_1) \in G$, findes et $\eta > 0$ således at

$$|x - x_1| < 2\eta \text{ og } |y - y_1| < 4M\eta$$
 (3.8)

medfører at

$$f(x_1, y_1) - \epsilon < f(x, y) < f(x_1, y_1) + \epsilon$$
(3.9)

Sættet af punkter $(x, y) \in G$, der opfylder ulighederne (3.8) og (3.9) er indeholdt i et regtangel Q med bredde 4η , højde $8M\eta$ og centrum i (x_1, y_1) , som illustreret på figur 3.6.



Figur 3.6. Regtangel *Q*, der indeholder de funktioner der opfylder ulighederne (3.8) og (3.9).



Figur 3.7. Skraveret område indeholdende $\widetilde{\phi}_k(x)$.

Vælg nu et *K*, således at k > K medfører at længden af linjestykkerne i Eulerlinjen L_k er mindre end η og

$$|\phi(x) - \widetilde{\phi}_k(x)| < M\eta$$

for alle $x \in [a, b]$. Hvis $|x - x_1| < 2\eta$ og k > K vil $\tilde{\phi}_k(x)$ dermed være indeholdt i Q, da den absolutte hældning af $\tilde{\phi}_k(x)$ er begrænset af M og afstanden mellem $\tilde{\phi}_k(x)$ og $\phi(x)$ maksimalt er $M\eta$. Dermed ligger $\tilde{\phi}_k(x)$ indenfor det skraverede område på figur 3.7 og dermed i Q. Så hvis $|x_2 - x_1| < \eta$, er

$$[f(x_1, y_1) - \epsilon](x_2 - x_1) < \widetilde{\phi}_k(x_2) - \widetilde{\phi}_k(x_1) < [f(x_1, y_1) + \epsilon](x_2 - x_1)$$
(3.10)

Det er ikke nok at bevise at der eksisterer en løsning til en given differentialligning. Det er ydermere essentielt at give betingelser for, hvornår en given løsning er entydig.

3.2 Entydighed af løsning

Den typiske betingelse der kræves for entydighed af en løsning er Lipschitz betingelsen

DEFINITION 3.5

Funktionen $f : [a, b] \times \mathbb{R}^d \to \mathbb{R}^d$ opfylder **Lipschitz betingelsen i dens anden variable**, hvis der eksisterer en konstant *L* således at, der for ethvert valg af $x \in [a, b]$ og $y, z \in \mathbb{R}^d$, gælder at

$$\|f(x,y) - f(x,z)\| \le L \|y - z\|$$
(3.11)

Konstanten *L* kaldes en **Lipschitz konstant**.

Der fører til en af mest kendte sætninger vedrørende entydighed af løsning for en differentialligning

SÆTNING 3.6

Lad funktionen $f : [a, b] \times \mathbb{R}^d \to \mathbb{R}^d$ være kontinuert i første variabel og opfylde Lipschitz betingelsen i anden variabel. Da eksisterer der en entydig løsning til begyndelsesværdiproblemet

$$\frac{dy}{dx} = f(x, y)$$
$$y(a) = y_a$$

Denne sætning vil ikke blive bevist her. I stedet bevises en mere generel sætning der kaldes Osgoods entydighedssætning.

Inden Osgoods entydighedssætning behandles, bliver der i det følgende set nærmere på entydigheden af løsninger til to simple typer af differentialligninger, da resultater udledt heraf, skal anvendes for at bevise Osgoods entydighedssætning.

Differentialligningen $\frac{dy}{dx} = f(x)$ **:**

Betragtes den simple differentialligning

$$\frac{dy}{dx} = f(x) \tag{3.12}$$

og antages det, at f(x) er kontinuert på det åbne interval (a, b), så er en af løsningerne til (3.12) tydeligvis givet ved

$$y(x) = \int_{x_0}^x f(z)dz$$
, hvor $x_0, x \in (a, b)$ (3.13)

Resten af løsningerne til (3.12) vil være lig (3.13) plus en vilkårlig konstant C. den generelle løsning til (3.12) kan altså skrives som

$$y(x) = \int_{x_0}^x f(z)dz + C$$
, hvor $x_0, x \in (a, b)$ (3.14)

Hvis det kræves at løsningen til (3.12) skal passere igennem et bestemt punkt (x_0 , y_0), ses det ved indsættelse af x_0 i (3.14) at $y_0 = y(x_0) = C$. Det vil altså sige at der kun eksisterer en løsning til (3.12), der passerer igennem punktet (x_0 , y_0), som er givet ved

$$y(x) = \int_{x_0}^x f(z)dz + y_0 \quad \text{, hvor} \quad x_0, x \in (a, b)$$

Differentialligningen $\frac{dy}{dx} = f(y)$:

Betragt nu den simple differentialligning på formen

$$\frac{dy}{dx} = f(y) \tag{3.15}$$

Denne er næsten på samme form som (3.12), blot med en højreside lig f(y) istedet for f(x). Dog kan (3.15) omskrives, såfremt $f(y) \neq 0$, til

$$\frac{dx}{dy} = \frac{1}{f(y)}$$

Antages det at f(y) er kontinuert på det åbne interval (a, b), kan resultaterne udledt ovenfor anvendes. Dermed findes der til ethvert punkt (x_0, y_0) , hvor $y_0 \in (a, b)$, en entydig løsning der passerer igennem, givet ved

$$x = \int_{y_0}^{y} \frac{1}{f(z)} dz + x_0 \quad \text{, hvor} \quad y_0, y \in (a, b)$$
(3.16)

SÆTNING 3.7 (Osgoods entydighedssætning)

Lad f(x, y) være en funktion defineret på en mængde G og $\phi(u)$ være positiv og kontinuert på (0, a]. Lad ydermere f(x, y) opfylde

$$|f(x, y_2) - f(x, y_1)| \le \phi \left(|y_2 - y_1|\right) \tag{3.17}$$

for ethvert par af punkter $(x, y_1), (x, y_2) \in G$, og lad $\phi(u)$ opfylde at

$$\lim_{\epsilon \to 0^+} \int_{\epsilon}^{a} \frac{1}{\phi(u)} du = \infty$$
(3.18)

Da eksisterer der ikke mere end en løsning til $\frac{dy}{dx} = f(x, y)$, der passerer gennem hvert punkt $(x_0, y_0) \in G$.

BEVIS

Lad funktionen f(x, y) opfylde betingelserne givet i sætning 3.7. For et vilkårligt punkt $(x_0, y_0) \in G$, antag at der eksisterer to forskellige løsninger, $y_1(x)$ og $y_2(x)$, til differentialligningen $\frac{dy}{dx} = f(x, y)$, således at

$$y_1(x_0) = y_2(x_0) = y_0 \tag{3.19}$$

Det kan antages at $x_0 = 0$, da x blot kan udskiftes med $x + x_0$. Definer nu en funktion z(x), givet ved

$$z(x) = y_1(x) - y_2(x)$$

Da $y_1(x)$ og $y_2(x)$ er forskellige, eksisterer der et punkt x_1 , således at $z(x_1) \neq 0$. Antag nu at $z(x_1) \geq 0$. Dette medfører ikke tab af generalitet, da et modsat tilfælde kan håndteres ved at

sætte $z(x) = y_2(x) - y_1(x)$. Antag ydermere at $x_1 > 0$. Her kan det modsatte tilfælde håndteres ved at udskifte $x \mod -x$. Da f(x, y) opfylder (3.17), gælder det for z(x), at

$$\frac{dz}{dx} = \frac{d(y_2 - y_1)}{dx}$$

$$= f(x, y_2) - f(x, y_1)$$

$$\leq \phi (|y_2 - y_1|)$$

$$< 2\phi (|y_2 - y_1|)$$
(3.20)

Betragt nu differentialligningen givet ved

$$\frac{dy}{dz} = 2\phi(y) \tag{3.21}$$

Det ønskes at finde en løsning hertil, der opfylder at $y(x_1) = z(x_1) = z_1$. Løsningen skal altså passere gennem punktet (x_1, z_1) . Sådan en løsning eksisterer og er entydig, da (3.21) er på samme form som (3.15), og dermed har en løsning på formen (3.16). Da funktionen ϕ opfylder (3.18) vil grafen for løsningen, y(x), til (3.21), gå mod den negative *x*-akse uden nogensinde at skære *x*-aksen (se evt. figur 3.8).



Figur 3.8. Illustration af y(x) og z(x).

Da løsningen
$$y(x)$$
 opfylder at $y(x_1) = z(x_1)$ og $z(x)$ opfylder (3.20), gælder det at

$$z'(x_1) < 2\phi(z_1) = 2\phi(y(x_1)) = y'(x_1)$$

Der eksisterer altså et $\epsilon > 0$ således at der på intervallet $(x_1 - \epsilon, x_1)$, gælder at

$$z(x) > y(x)$$

Dette gælder faktisk for alle $0 < \epsilon \le x_1$. Dette ses ved at lade x_2 være lig $x_1 - \epsilon_{max}$, hvor ϵ_{max} er den største værdi, for hvilken det gælder at z(x) > y(x). Derved fås

$$z'(x_2) \ge y'(x_2) = 2\phi(y(x_2)) = 2\phi(z(x_2))$$

Hvilket er i modstrid med (3.20), der siger at

$$z'(x_2) < 2\phi(z(x_2)) \tag{3.22}$$

Dermed er

$$z(x) > y(x) \quad \forall x \in [0, x_1] \tag{3.23}$$

Det blev antaget tidligere at $x_0 = 0$, så (3.23) medfører at $z(0) = z(x_0) > 0$, hvilket er i modstrid med (3.19). Dermed kan der ikke findes to forskellige løsninger til differentialligningen $\frac{dy}{dx} = f(x, y)$, og sætningen er dermed bevist.

Eksempler på funktioner der opfylder betingelserne for phi-funktionen i sætning 3.7 er

Ku, Ku|ln(u)|, Ku|ln(u)|ln(|ln(u)|)

hvor *K* er en positiv konstant. Tilfældet hvor funktionen er lig *Ku* er ækvivalent med Lipschitz betingelsen defineret i definition 3.5, så sætning 3.6 følger af Osgoods entydighedssætning.

Referencer

Definition 3.1 er skrevet på baggrund af [Petrovski, 1973, s. 28]. Definition 3.2 er fra [Møller, 2012, s. 329], dog tilføjet egenskaben uniform. Sætning og bevis er skrevet på baggrund af [Petrovski, 1973, s. 28-29]. Sætning med bevis er skrevet på baggrund af [Petrovski, 1973, s. 29-31]. Definition 3.5 af Lipschitz betingelsen [Butcher, 2008, s. 22]. Sætning 3.6 stammer fra [Butcher, 2008, s. 23]. Teorien vedrørende entydighed af løsninger til de to simple differentialligninger, er skrevet på baggrund af [Petrovski, 1973, s. 10-13]. Osgoods entydighedssætning med bevis er skrevet på baggrund af [Petrovski, 1973, s. 34-35].

Metoder baseret på Taylorudvikling 4

Betragtes Eulers metode beskrevet i afsnit 2.2, ses det at hvert skridt (2.9) approksimeres vha. Taylorudvikling med to led

$$y_{k+1} = y(x_{k+1})$$

= $y(x_k + h)$
= $y(x_k) + hy'(x_k)$
= $y(x_k) + hf(x_k, y_k)$

Det er derfor oplagt at overveje yderligere approksimationsmetoder baseret på Taylorudvikling med mere end to led. Således at hvert enkelt skridt er givet ved

$$y_{k+1} = y_k + h\left(f^{[0]}(x_k, y_k) + \frac{h}{2}f^{[1]}(x_k, y_k) + \dots + \frac{h^{p-1}}{p!}f^{[p-1]}(x_k, y_k)\right)$$
(4.1)

Hvor de totale afledede af f er givet ved

$$f^{[0]}(x,y) = f(x,y)$$

$$f^{[1]}(x,y) = f_x(x,y) + f_y(x,y)f(x,y)$$

$$\vdots$$

$$f^{[p+1]}(x,y) = f_x^{[p]}(x,y) + f_y^{[p]}(x,y)f(x,y)$$

Formen af (4.1) kan altså hurtigt kræve en hel del udledninger af partielle afledede af funktionen f, hvilket tidligere udgjorde en stor arbejdsbyrde. I dag kan dette dog klares vha. computere, så metoden er igen en mulighed.

Funktionen Φ for en metode baseret på Taylorudvikling, er altså givet ved

$$\Phi(x_k y_k; h) = f^{[0]}(x_k, y_k) + \frac{h}{2} f^{[1]}(x_k, y_k) + \dots + \frac{h^{p-1}}{p!} f^{[p-1]}(x_k, y_k)$$
$$= \sum_{i=0}^{p-1} \frac{h^i}{(i+1)!} f^{[i]}(x_k, y_k)$$

For at vurdere metodens orden, betragtes den lokale trunkeringsfejl

$$T(x_k, y_k; h) = \Phi(x_k, y_k; h) - \frac{1}{h} \left(u(x_k + h) - u(x_k) \right)$$
(4.2)

Det er nødvendigt at omskrive funktionen Φ . Dette gøres ved at udnytte at

$$u^{(p+1)}(t) = f^{[p]}(t, u(t))$$

som for t = x bliver

$$u^{(p+1)}(x) = f^{[p]}(x, y)$$

så

$$\Phi(x_k, y_k; h) = \sum_{i=0}^{p-1} \frac{h^i}{(i+1)!} u^{(i+1)}(x_k)$$
(4.3)

Antag nu at $f \in C^p$ på $[a, b] \times \mathbb{R}^d$, så kan Taylors sætning (A.4) anvendes på $\frac{1}{h}(u(x_k + h) - u(x_k))$, så (4.2) bliver

$$T(x_k, y_k; h) = \Phi(x_k, y_k; h) - \sum_{i=0}^{p-1} \frac{h^i}{(i+1)!} u^{(i+1)}(x_k) - u^{(p+1)}(\xi) \frac{h^p}{(p+1)!}$$

$$= -u^{(p+1)}(\xi) \frac{h^p}{(p+1)!} , \quad x_k < \xi < x_k + h$$
(4.4)

Bemærk at der igen med notationen $u^{(p+1)}(\xi)$ forståes, at der eksisterer et $x_k < \xi < x_k + h$ for hver af komponenterne i funktionen *u*. Antages det, at den *p*'te totale afledede af *f* er ligeligt begrænset i $[a, b] \times \mathbb{R}^d$, da eksisterer der en konstant C_p således at

$$\|u^{(p+1)}(\xi)\| \le C_p$$

hvilket medfører at

$$||T(x_k, y_k; h)|| \le \frac{C_p}{(p+1)!}h^p$$

Så ordenen af en metode defineret ved (4.1) har altså orden p jævnfør definition 2.3. Den eksakte orden kan findes ved yderligere at antage at den (p + 1)'te totale afledede af f er ligeligt begrænset i $[a, b] \times \mathbb{R}^d$, så der findes en konstant C_{p+1} således at

$$||u^{(p+1)}(x_k) - u^{(p+1)}(\xi)|| \le C_{p+1}h$$

Hvilket medfører at der eksisterer et $\widetilde{C}_{p+1} \leq C_{p+1}$ således at

$$u^{(p+1)}(\xi) = u^{(p+1)}(x_k) + \widetilde{C}_{p+1}h$$

= $u^{(p+1)}(x_k) + O(h)$

Indsættes dette i (4.4) fås

$$T(x_k, y_k; h) = -\left(u^{(p+1)}(x_k) + O(h)\right) \frac{h^p}{(p+1)!}$$

= $-u^{(p+1)}(x_k) \frac{h^p}{(p+1)!} + O(h^{p+2}) , \quad h \to 0$
= $-f^{[p]}(x_k, y_k) \frac{h^p}{(p+1)!} + O(h^{p+2}) , \quad h \to 0$

Hvilket viser at metoden er har eksakt orden p, medmindre $f^{[p]} \equiv 0$, og at den principielle fejlfunktion, jævnfør definition 2.4, er givet ved

$$\sigma(x_k, y_k) = -f^{[p]}(x_k, y_k) \frac{1}{(p+1)!}$$

Referencer

Dette kapitel er skrevet på baggrund af [Gautschi, 1997, s. 275-276].
Forbedring af Eulers metode

I dette kapitel betragtes to forbedringer af Eulers metode; Forbedret Euler og Heuns metode. Disse bygger på den samme idé, som Eulers metode, men i stedet for at følge den samme hælding over en skridtlængde *h*, foreslår disse metoder andre mere effektive hældinger.

5.1 Forbedret Euler

Istedet for at evaulere hældningen i det punkt hvor et givet linjestykke starter, kan hældningen evauleres halvejs gennem linjestykket. Først evauleres hældningen altså i linjestykkets startpunkt (x_k, y_k) , derefter følges denne hældning, $f(x_k, y_k)$, over en halv skridtlængde h, hvorefter hældningen reevalueres. Den reevaluerede hældning $f(x_k + \frac{1}{2}h, y + \frac{1}{2}h(f(x_k, y_k)))$ anvendes herefter over hele intervallet $[x_k, x_{k+1}]$. Dette er illustreret på figur 5.1.



Figur 5.1. Illustration af idéen bag Forbedret Euler.

Et skridt med denne forbedring af Eulers metode kan altså skrives som

$$y_{k+1} = y_k + hf\left(x_k + \frac{1}{2}h, y_k + \frac{1}{2}hf(x_k, y_k)\right)$$
(5.1)

så

$$\Phi(x_k, y_k; h) = f\left(x_k + \frac{1}{2}h, y_k + \frac{1}{2}hf(x_k, y_k)\right)$$
(5.2)

Bemærk at der forekommer en indlejring af $f(x_k, y_k)$ i (5.1). Når dette skal programmeres kan det

derfor være en fordel at anvende nedenstående notation istedet.

$$k_1(x_k, y_k) = f(x_k, y_k)$$

$$k_2(x_k, y_k; h) = f\left(x_k + \frac{1}{2}h, y_k + \frac{1}{2}hk_1\right)$$

så (5.1) bliver

$$y_{k+1} = y_k + hk_2$$

Bemærk ydermere at funktionen Φ i (5.2) opfylder at $\Phi(x_k, y_k; 0) = f(x_k, y_k)$, og er dermed konsistent jævnfør (2.7).

En anden tilgang til at forbedre Eulers metode er Heuns metode, som beskrives i det følgende.

5.2 Heuns metode

Ved Heuns metode findes hvert linjestykkes hældning ved at tage gennemsnittet af hældningen i linjestykkes startpunkt og hældningen i linjestykkets endepunkt. Først evauleres hældningen altså i punktet (x_k, y_k) , hvorefter denne hældning følges over skridtlængden h, hvor hældningen reevalueres i punktet $(x_k + h, y_k + hf(x_k, y_k))$. Til sidst tages gennemsnittet af de to hældninger $f(x_k, y_k)$ og $f(x_k + h, y_k + hf(x_k, y_k))$. Så et skridt i Heuns metode er altså defineret ved

$$y_{k+1} = y_k + \frac{1}{2}h\left(f(x_k, y_k) + f(x_k + h, y_k + hf(x_k, y_k))\right)$$
(5.3)

så

$$\Phi(x_k, y_k; h) = \frac{1}{2} \left(f(x_k, y_k) + f(x_k + h, y_k + hf(x_k, y_k)) \right)$$
(5.4)

Bemærk igen at funktionen Φ opfylder at $\Phi(x_k, y_k; 0) = f(x_k, y_k)$, så Heuns metode er altså konsistent. Der sker igen en indlejring af $f(x_k, y_k)$ i (5.3), så der anvendes igen en omskrivning af (5.3), givet ved

$$k_1(x_k, y_k) = f(x_k, y_k)$$

$$k_2(x_k, y_k; h) = f(x_k + h, y_k + hk_1)$$

$$y_{k+1} = y_k + \frac{1}{2}h(k_1 + k_2)$$

Det er nu relevant at se på effekten de to ovenstående ændringer af Eulers metode. I afsnit 2.2, blev det vist, at Eulers metode er af orden 1. De to metoder i dette afsnit er begge af orden 2, hvilket vil blive vist i næste afsnit.

5.3 Two-stage metoder

De to forbedringer af Eulers metode ovenfor er begge *two-stage metoder*. Generelt har disse metoder en Φ -funktion på formen

 $\Phi(x_k, y_k; h) = \alpha_1 k_1 + \alpha_2 k_2$

hvor

$$k_1(x_k, y_k) = f(x_k, y_k)$$

$$k_2(x_k, y_k; h) = f(x_k + \mu h, y_k + \mu h k_1)$$
(5.5)

Parametrene α_1 , α_2 og μ ovenfor, er alle reelle tal, der kan vælges således at ordenen af metoden maksimeres. Det vises nu, hvorledes parametrene kan vælges, så den givne metode bliver af orden 2.

Først udvides funktionerne Φ og $u(x_k + h)$, hvorefter disse anvendes til at udlede et udtryk for den lokale trunkeringsfejl $T(x_k, y_k; h)$, der kan omskrives, så det fremgår hvilket valg af α_1, α_2 og μ , der er optimalt. For at lette notationen, angives det punkt, der tidligere er angivet (x_k, y_k) , nu med (x, y).

For at udvide funktionen Φ anvendes Taylor udviklingen for en vektorfunktion af flere variable (se evt. sætning A.6)

$$f(x + \Delta x, y + \Delta y) = f + f_x \Delta x + f_y \Delta y + \frac{1}{2} \left(f_{xx} (\Delta x)^2 + 2f_{xy} \Delta x \Delta y + (\Delta y)^T f_{yy} (\Delta y) \right) + \dots$$
(5.6)

Hvor f_x er den partielle afledede af f mht. x, f_y er Jacobimatricen for f mht. y-variablene og f_{yy} en vektor med Hesse matricer af f. Ydermere bemærkes det at alle funktionerne for de partielle afledede ovenfor evalueres i (x, y).

Formlen (5.6) kan nu anvendes på k_2 i (5.5) ved at sætte $\Delta x = \mu h$ og $\Delta y = \mu h f$, så

$$k_{2}(x,y;h) = f + f_{x}\mu h + f_{y}\mu hf + \frac{1}{2}\left(f_{xx}(\mu h)^{2} + 2f_{xy}\mu h\mu hf + (\mu hf)^{T}f_{yy}\mu hf\right) + O(h^{3})$$
(5.7)
= $f + \mu h(f_{x} + f_{y}f) + \frac{1}{2}\mu^{2}h^{2}\left(f_{xx} + 2f_{xy}f + f^{T}f_{yy}f\right) + O(h^{3})$

Anvendes Taylor udvikling ligeledes på u(x+h), fås

$$u(x+h) = u(x) + hu'(x) + \frac{h^2}{2}u''(x) + \frac{h^3}{6}u'''(x) + O(h^4)$$

hvilket medfører at

$$\frac{1}{h}(u(x+h) - u(x)) = u'(x) + \frac{h}{2}u''(x) + \frac{h^2}{6}u'''(x) + O(h^3)$$
(5.8)

Her er funktionerne u'(x), u''(x) og u'''(x) givet ved

$$u'(x) = f$$

$$u''(x) = f_x + f_y f$$

$$u'''(x) = \frac{d(f_x + f_y f)(t, u(t))}{dt}$$

$$= \frac{\partial f_x + f_y f}{\partial x} \frac{dt}{dt} + \frac{\partial f_x + f_y f}{\partial y} \frac{du}{dt}$$

$$= (f_x + f_y f)_x + (f_x + f_y f)_y f$$

$$= f_{xx} + (f_y f)_x + (f_{xy} + (f_y f)_y) f$$

$$= f_{xx} + f_{yx} f + f_y f_x + f_{xy} f + (f_y f)_y f$$
(5.9)

For at omskrive u'''(x) yderligere, anvendes det at $f_{yx} = f_{xy}$ jf. korollar A.3 i Appendiks. Derudover anvendes det at

$$(f_y f)_y f = (f_{yy} f + f_y f_y)_y f = f^T f_{yy} f + f_y^2 f$$

så

$$u'''(x) = f_{xx} + 2f_{xy}f + f_yf_x + f^Tf_{yy}f + f_y^2f = f_{xx} + 2f_{xy} + f^Tf_{yy}f + f_y(f_x + f_yf)$$
(5.10)

Betragt nu den lokale trunkeringsfejl (se evt. definition 2.1) for en two-stage metode

$$T(x,y;h) = \alpha_1 k_1 + \alpha_2 k_2 - \frac{1}{h} [u(x+h) - u(x)]$$
(5.11)

Indsættes $k_1 = f$, (5.8) og (5.7) i (5.11), fås

$$\begin{split} T(x,y;h) &= \alpha_1 f + \alpha_2 \left[f + \mu h(f_x + f_y f) + \frac{1}{2} \mu^2 h^2 \left(f_{xx} + 2f_{xy} f + f^T f_{yy} f \right) + O(h^3) \right] \\ &- \left[f + \frac{h}{2} (f_x + f_y f) + \frac{h^2}{6} \left(f_{xx} + 2f_{xy} + f^T f_{yy} f + f_y \left(f_x + f_y f \right) \right) + O(h^3) \right] \\ &= \alpha_1 f + \alpha_2 f + \alpha_2 \mu h(f_x + f_y f) + \frac{1}{2} \alpha_2 \mu^2 h^2 \left(f_{xx} + 2f_{xy} f + f^T f_{yy} f \right) \\ &- f - \frac{h^2}{6} \left(f_{xx} + 2f_{xy} + f^T f_{yy} f \right) - \frac{h^2}{6} \left(f_y \left(f_x + f_y f \right) \right) + O(h^3) \\ &= (\alpha_1 + \alpha_2 - 1)f + \left(\alpha_2 \mu - \frac{1}{2} \right) h(f_x + f_y f) \\ &+ \frac{1}{2} h^2 \left[\left(\alpha_2 \mu^2 - \frac{1}{3} \right) \left(f_{xx} + 2f_{xy} f + f^T f_{yy} f \right) - \frac{1}{3} f_y(f_x + f_y f) \right] + O(h^3) \end{split}$$

Betragtes ovenstående omskrivning af den lokale trunkeringsfejl, ses det, at de første to led kan give nul vha. et passende valg af α_1, α_2 og μ . For at få det tredje led til at give nul, skal der indføres strenge betingelser for funktionen f og dennes afledede. Den maksimale orden er derfor 2. En two-stage metode er altså af orden 2 hvis følgende ligningssystem er opfyldt

$$\alpha_1 + \alpha_2 - 1 = 0$$
$$\alpha_2 \mu - \frac{1}{2} = 0$$

Der kan altså vælges et vilkårligt $\alpha_2 \neq 0$, hvorefter α_1 og μ er givet ved

$$\alpha_1 = 1 - \alpha_2 \tag{5.12}$$
$$\mu = \frac{1}{2\alpha_2} \quad , \quad \alpha_2 \neq 0$$

Forbedret Euler, beskrevet i sektion 5.1, er en two stage metode, hvor $\alpha_1 = 0$, $\alpha_2 = 1$ og $\mu = \frac{1}{2}$. Heuns metode, beskrevet i sektion 5.2, anvender parametrene $\alpha_1 = \frac{1}{2}$, $\alpha_2 = \frac{1}{2}$ og $\mu = 1$. Både Forbedret Euler og Heuns metode opfylder dermed (5.12), så disse to modifikationer af Eulers metode har altså kunne hæve metodens orden fra 1 til 2.

5.4 Implementering

Forbedret Euler er implementeret i algoritme 2 og Heuns metode i algoritme 3. Begge algoritmer tager samme parametre son input, som algoritme 1 fra tidligere. Disse inputs er

- *f* højresiden i differentialligning $\frac{dy}{dx} = f(x, y)$
- xnul, ynul begyndelsesbetingelserne ynul = y(xnul).
- *a*, *b* det interval [*a*, *b*] løsningen skal approksimeres over.
- *n* antallet af skridt der anvendes.

Outputtet er ligeledes som ved algoritme 1; et sæt indeholdende en liste $Lx = [x_0, x_1, ..., x_n]$ og en liste med approksimationerne $Ly = [y_0, y_1, ..., y_n]$.

ALGORITME 2 (Forbedret Euler med *n* inddelinger).

$$forbedreteuler_1 := proc(f, xnul, ynul, a, b, n)$$

$$h := \frac{b-a}{n};$$

$$x[0] := xnul;$$

$$y[0] := ynul;$$

$$Lx := [x[0]]; Ly := [y[0]];$$

$$for i from 0 to n - 1 do$$

$$x[i+1] := x[i] + h;$$

$$k1 := f(x[i], y[i]);$$

$$k2 := f\left(x[i] + \frac{1}{2} \cdot h, y[i] + \frac{1}{2} \cdot h \cdot k1\right);$$

$$y[i+1] := y[i] + h \cdot k2;$$

$$Lx := [op(Lx), x[i+1]];$$

$$Ly := [op(Ly), y[i+1]]$$
end do;
return {Lx, Ly}
end proc;

ALGORITME 3 (Heuns metode med *n* inddelinger). $heun_1 := proc(f, xnul, ynul, a, b, n)$ $h := \frac{b-a}{n};$ x[0] := xnul; y[0] := ynul; Lx := [x[0]]; Ly := [y[0]];for *i* from 0 to *n* - 1 do x[i+1] := x[i] + h; k1 := f(x[i], y[i]); $k2 := f(x[i] + h, y[i] + h \cdot k1);$ $y[i+1] := y[i] + \frac{h}{2}(k1 + k2);$ Lx := [op(Lx), x[i+1]]; Ly := [op(Ly), y[i+1]]end do;

return {Lx, Ly}

end proc;

Eksempel 5.1

I dette eksempel approksimeres en løsning til begyndelsesværdiproblemet fra eksempel 2.1 og 2.6, vha. Heuns metode.

$$\frac{dy}{dx} = 4x^3y, \qquad 0 \le x \le 1, \qquad y_0 = 1 \tag{5.13}$$

Løsningen er approksimeret ved anvendelse af algoritme 3 med 4, 8 og 16 skridt. De tre approksimerede løsninger er plottet på figur 5.2. Derudover kan den eksakte løsning til (5.13) ses på figur 5.3.



Umiddelbart virker Heuns metode bedre til at approksimere en løsning til (5.13) end Eulers metode i eksempel 2.6. Den absolutte afvigelse fra den eksakte løsning for de tre appoksimationer foretaget vha. Heuns metode er plottet på figur 5.4. Derudover er den absolutte afvigelse fra den eksakte løsning for de fire approksimationer foretaget vha. Eulers metode fra eksempel 2.6 plottet på figur 5.5.



Figur 5.4. Plot af den absolutte afvigelse fra den eksakte løsning for løsninger approksimeret vha. Heuns metode.



Figur 5.5. Plot af den absolutte afvigelse fra den eksakte løsning for løsninger approksimeret vha. Eulers metode.

Af figur 5.4 og 5.5, ses det at den absolutte afvigelse ved Heuns approksimationer er væsenlig mindre end ved Eulers metode. Et helt tydeligt billede af denne forskel, ses ved at plotte den absolutte fejl for Eulers metode med 32 skridt og Heuns metode med 16 skridt, i det samme koordinatsystem. Dette plot kan ses på figur 5.6



Figur 5.6. Plot af den absolutte afvigelse fra den eksakte løsning for løsninger approksimeret vha. Eulers metode med 32 skridt og Heuns metode med 16 skridt.

EKSEMPEL 5.2 I dette eksempel betragtes begyndelsesværdiproblemet

$$\frac{dy}{dx} = -y\sin(x), \qquad 0 \le x \le 4\pi, \qquad y_0 = 1$$
 (5.14)

Den eksakte løsning til (5.14) er givet ved

$$y = \frac{\exp(\cos(x))}{\exp(1)}$$

Et plot af den eksakte løsning kan ses på figur 5.7



Figur 5.7. Eksakt løsning til (5.14).

Forbedret Euler (5.1) anvendes til at approksimere løsninger til (5.14). For at gøre dette anvendes algoritme 2, hvor n = 4, 8, 16, 32. De approksimerede løsninger er plottet på figur 5.8, 5.9, 5.10 og 5.11, for hendholsvis n = 4, n = 8, n = 16 og n = 32.





Figur 5.9. Approksimation af løsning til (5.14) foretaget vha. Forbedret Euler med 8 skridt.



Figur 5.10. Approksimation af løsning til (5.14) foretaget vha. Forbedret Euler med 16 skridt.



Figur 5.11. Approksimation af løsning til (5.14) foretaget vha. Forbedret Euler med 32 skridt.

Af figur 5.8 og 5.9 ses det, at en approksimation af løsningen til (5.14), foretaget vha. forbedret Eulers metode, langt fra er tilfredstillende når der kun anvendes 4 og 8 skridt. På figur 5.10 og 5.11, hvor n er henholdsvis 16 og 32, ses det, at de approksimerede løsninger i højere grad opfører sig som den eksakte løsning på figur 5.7.For at sammenligne de fire approksimerede løsninger ovenfor er disse plottet sammen med den eksakte løsning på figur 5.12.



Figur 5.12. Samlet plot af approksimerede løsninger sammen med den eksakte løsning.

Af figur 5.12 fremgår det tydeligt at en forøgelse af antallet af skridt, n, medfører at den approksimerede løsninger er tættere på den eksakte løsning. \triangle

Referencer

Afsnit 5.1 og 5.2 er skrevet vha. af [Gautschi, 1997, s. 276-278]. Afsnit 5.3 er skrevet på baggrund af [Gautschi, 1997, s. 278-280].

Runge-Kutta metoder

I kapitel 5 blev Eulers metode forbedret i form af to forskellige two-stage metoder, henholdsvis forbedret Euler og Heuns metode. Ideen bag two-stage metoder kan udvides til *r*-stage metoder. Ved en *r*-stage metode er Φ -funktionen givet ved

$$\Phi(x_k, y_k; h) = \sum_{s=1}^r \alpha_s k_s(x_k, y_k; h)$$
(6.1)

hvor

$$k_{1} = f(x_{k}, y_{k})$$

$$k_{2} = f(x_{k} + \mu_{2}h, y_{k} + h\lambda_{11}k_{1})$$

$$\vdots$$

$$k_{s} = f\left(x_{k} + \mu_{s}h, y_{k} + h\sum_{j=1}^{s-1}\lambda_{sj}k_{j}\right) , \quad s = 3, 4, 5, ..., r$$
(6.2)

Her indføres der følgende betingelser for μ_s , λ_{sj} og α_s

$$\mu_s = \sum_{j=1}^{s-1} \lambda_{sj}$$
$$\sum_{s=1}^r \alpha_s = 1$$

Med k-funktioner som i (6.2) kaldes metoden for en *eksplicit r-stage Runge-Kutta* metode. Denne metode kræver r evalueringer af funktionen f. En *implicit r-stage Runge-Kutta* metode har k-funktioner defineret ved

$$k_{s} = f\left(x_{k} + \mu_{s}h, y_{k} + h\sum_{j=1}^{r}\lambda_{sj}k_{j}\right) \quad , \quad s = 1, 2, 3, ..., r$$
(6.3)

Implicit Runge-Kutta er mere generel end den eksplicitte. Den producerer dog *r* ligninger med *r* ubekendte i form af de *r k*-funktioner. Hver af disse *k*-funktioner en en vektor i \mathbb{R}^d , så det er reelt set *rd* ligninger med *rd* ubekendte der skal løses for at bestemme et udtryk for Φ -funktionen (6.1). Der findes ydermere *semi-implicit r-stage Runge-Kutta* metoder, hvor

$$k_{s} = f\left(x_{k} + \mu_{s}h, y_{k} + h\sum_{j=1}^{s}\lambda_{sj}k_{j}\right) \quad , \quad s = 1, 2, 3, .., r$$
(6.4)

Her går summen i (6.4) fra j = 1 til j = s i stedet for fra j = 1 til j = r som i (6.3). Dermed produceres der r ligninger med hver d ubekendte.

Fordelene ved implicit og semiimplicit, er at metodens orden kan gøres højere end ved eksplicitte metoder. Den større arbejdsbyrde der er forbundet med implicitte og semiimplicitte metoder gør dog, at disse kun anvendes ved specielle typer af problemer, som der ikke vil blive gennemgået her.

Parametrene α_s og λ_{sj} i (6.1) og (6.2) kan vælges så metoden får den højst mulige orden. Fremgangsmåden til at finde disse parametre er illustreret i afsnit 5.3, hvor det bedste valg af α_1 , α_2 og μ for en two-stage metode blev udledt. Analoge udledninger for en *r*-stage metode er langt mere krævende, grundet det store antal partielle afledede af *f*, der fremkommer.

Kutta viste i 1901 at den maksimale orden for eksplicitte Runge-Kutta metoder med r = 1, 2, 3, 4 er lig med r. Butcher viste i 1960'erne de maksimale ordener for r > 4. Lad p(r) angive den maksimale orden af en eksplicit r-stage Runge-Kutta metode, så har Kutta og Butcher vist at

$$\begin{array}{ll} p(r) &= r & \text{for} & 1 \le r \le 4 \\ p(r) &= r - 1 & \text{for} & 5 \le r \le 7 \\ p(r) &= r - 2 & \text{for} & 8 \le r \le 9 \\ p(r) &< r - 2 & \text{for} & r \ge 10 \end{array}$$

6.1 RK4

I denne rapport anvendes Runge-Kutta formlen af orden 4, ofte angivet *RK4*, hvor Φ - og *k*-funktionerne er defineret ved

$$\Phi(x_k, y_k; h) = \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

$$k_1(x_k, y_k) = f(x_k, y_k)$$

$$k_2(x_k, y_k; h) = f\left(x_k + \frac{1}{2}h, y_k + \frac{1}{2}hk_1\right)$$

$$k_3(x_k, y_k; h) = f\left(x_k + \frac{1}{2}h, y_k + \frac{1}{2}hk_2\right)$$

$$k_4(x_k, y_k; h) = f(x_k + h, y_k + hk_3)$$
(6.5)

Så et skridt ved RK4 er altså defineret ved

$$y_{k+1} = y_k + \frac{h}{6}(k_1(x_k, y_k) + 2k_2(x_k, y_k) + 2k_3(x_k, y_k) + k_4(x_k, y_k))$$
(6.6)

6.2 Implementering

RK4 er implementeret i algoritme 4, efter samme princip som implementeringen af Eulers metode (algoritme 1), med samme inputs og form af output.

ALGORITME 4 (RK4 med n inddelinger).

 $RK4_1 := proc(f, xnul, ynul, a, b, n)$ $h := \frac{b-a}{n};$ x[0] := xnul;y[0] := ynul;Lx := [x[0]];Ly := [y[0]];for i from 0 to n - 1 do x[i+1] := x[i] + h;k1 := f(x[i], y[i]); $k2 := f\left(x[i] + \frac{1}{2} \cdot h, y[i] + \frac{1}{2} \cdot h \cdot k1\right);$ $k3 := f\left(x[i] + \frac{1}{2} \cdot h, y[i] + \frac{1}{2} \cdot h \cdot k2\right);$ $k4 := f(x[i] + h, h \cdot k3 + y[i]);$ $y[i+1] := y[i] + \frac{1}{6} \cdot h \cdot (k1 + 2 \cdot k2 + 2 \cdot k3 + k4);$ Lx := [op(Lx), x[i+1]];Ly := [op(Ly), y[i+1]];end do; *return* {Lx, Ly} end proc;

Eksempel 6.1

I dette eksempel anvendes algoritme 4 til at approksimere en løsning til begyndelsesværdi problemet fra eksempel 5.2, givet ved

$$\frac{dy}{dx} = -y\sin(x), \qquad 0 \le x \le 4\pi, \qquad y_0 = 1$$
 (6.7)

hvor den eksakte løsning er givet ved

$$y = \frac{\exp(\cos(x))}{\exp(1)}$$

Et plot af den eksakte løsning kan ses på figur 5.7 i afsnit 5.4. Algortime 4 er anvendt til at approksimere en løsning til (6.7) med 4, 8, 16 og 32 inddelinger. Disse løsninger er plottet på figur 6.1, 6.2, 6.3 og 6.4.



Figur 6.1. Approksimation af løsning til (6.7) foretaget vha. RK4 med 4 skridt.



Figur 6.2. Approksimation af løsning til (6.7) foretaget vha. RK4 med 8 skridt.



Figur 6.3. Approksimation af løsning til (6.7) foretaget vha. RK4 med 16 skridt.



Figur 6.4. Approksimation af løsning til (6.7) foretaget vha. RK4 med 32 skridt.

På figur 6.2 ses det at den approksimerede løsning foretaget med kun 8 inddelinger, allerede begynder at opføre sig som den eksakte løsning (se figur 5.7 i afsnit 5.4). I eksempel 5.2 blev Forbedret Euler anvendt til at approksimere løsninger til (6.7). For at sammenligne de to metoder, Forbedret Euler og RK4, plottes den absolutte afvigelse fra den eksakte løsning. På figur 6.5 ses den absolutte afvigelse fra den eksakte løsning for approksimationer foretaget med 8 inddelinger vha. henholdsvis, Forbedret Euler og RK4. Samme plot er foretaget for 16 inddelinger på figur 6.6. Af figur 6.5 fremgår det, at approksimationen foretaget vha. RK4 er tydeligt bedre end approksimationen foretaget vha. Forbedret Euler. Anvendes 16 inddelinger, ses det på figur 6.6, at den absolutte afvigelse fra den eksakte løsning, er væsentligt mindre for begge metoder.



Figur 6.5. Den absolutte afvigelse fra den eksakte løsning for approksimationer foretaget med 8 inddelinger vha. henholdsvis, Forbedret Euler og RK4.



Figur 6.6. Den absolutte afvigelse fra den eksakte løsning for approksimationer foretaget med 16 inddelinger vha. henholdsvis, Forbedret Euler og RK4.

Referencer

Dette kapitel er primært skrevet på baggrund af [Gautschi, 1997, s. 280-282]. Derudover er [Butcher, 1965, s. 410] og [Butcher, 1985, s. 521], til at angive de ordenen Butcher har udledt i forbindelse med eksplicitte Runge-Kutta metoder.

Numeriske eksperimenter

Approksimationsmetoderne gennemgået i denne rapport har alle tilknyttet en orden p, som defineret i definition 2.3, hvor en metode siges at have orden p, hvis der eksisterer et C således at

$$\|T(x_k, y_k; h)\| \le Ch^p \tag{7.1}$$

ligeligt på $[a, b] \times \mathbb{R}^d$. Her er $T(x_k, y_k; h)$ den lokale trunkeringsfejl, som defineret i definition 2.1. Fra teorien i de foregående kapitler vides det at

- Euler metode har orden 1
- Forbedret Euler har orden 2
- Heuns metode har orden 2
- Runge-Kutta metoden, RK4, har orden 4

Lad skridtlængden være $h = \frac{b-a}{n}$, hvor *n* er antal skridt på intervallet [a, b]. Fordobles antallet af skridt bliver $h = \frac{b-a}{2n}$, så af (7.1), kan det forventes at den øvre grænse af den lokaletrunkeringsfejl for eulermetode *Ch* halveres. Ligeledes vil en fordobling af antal skridt ved Forbedret Euler og Heuns metode medføre at Ch^2 bliver fire gange mindre. Ordenen for RK4 er lig 4, så Ch^4 , bliver altså 16 gange mindre, ved en fordobling af antallet af skridt *n*.

I de følgende eksempler betragtes forholdet

$$\frac{E^n}{E^{2n}} \tag{7.2}$$

hvor E^n angiver fejlen for *n* inddelinger. Forholdet (7.2) kan efter teorien forventes at gå mod 2 for Eulers metode, 4 for Forbedret Euler og Heuns metode, og 16 for RK4. Der eksperimenteres med to typer af fejl. Den første type fejl er den fejl der opstår i intervallets endepunktet *b*, dvs.

$$E_{endepunkt}^n = |y(b) - app(b)|$$

hvor *y* er den eksakte løsning og *app* er den approksimerede løsning. den anden type fejl er den maksimale approksimationsfejl der opstår over hele intervallet, dvs.

$$E_{max}^{n} = max_{i \in [0,1,\dots,n]}(|y(x_{i}) - app(x_{i})|)$$

7.1 Algoritmer

Til at beregne forholdet mellem $E_{endepunkt}^{n}$ og $E_{endepunkt}^{2n}$ for Eulers metode anvendes algoritme 5. Algoritmens input er

• *f* - højresiden af differentialligning $\frac{dy}{dx} = f(x, y)$.

- *y* den eksakte løsning til differentialligningen.
- xnul, ynul begyndelsesbetingelser hørende til differentialligningen
- [*a*, *b*] intervallet løsningerne approksimeres over.
- *m* antallet af gange algoritmen kører.
- ABC antal betydende cifre algoritmen anvender.

Bemærk at algoritmen undervejs kalder algoritme 1 fra afsnit 2.3. Algoritme 1 har dog fået tilføjet inputtet *ABC*, samt kommandoen evalf(..), der evaluerer funktionsværdier undervejs, hvilket ikke fremgår her.

```
ALGORITME 5 (Forholdet mellem E_{endevunkt}^{n} og E_{endevunkt}^{2n} for Eulers metode).
    fejlendepunkt\_euler := proc(f, y, xnul, ynul, a, b, m, ABC)
                Digits := ABC;
                eksakt := y(b);
                app := [];
                fejl := [];
                forhold := [1];
                for i from 1 to m do
                   app := [op(app), eulersmetode(f, xnul, ynul, a, b, 2<sup>i</sup>, ABC)];
                   feil := [op(feil), abs(evalf(eksakt - app[i][2][2^i + 1]))];
                end do;
                for i from 2 to m do
                  forhold := \left[ op(forhold), \frac{fejl[i-1]}{fejl[i]} \right];
                end do;
                for i from 2 to m do
printf("\%4d\% + 1.5f\% + 1.5f\% 1.5e\% 1.5fn", 2^{i}, app[i][2][2^{i} + 1], eksakt, fejl[i], forhold[i])
              end do;
   end proc;
```

Outputtet af algoritme 5 er en tabel med 5 kolonner, indeholdende antal indelinger, approksimeret værdi app(b), eksakt værdi y(b), fejl $E_{endepunkt}^n = |y(b) - app(b)|$ samt forholdet mellem $E_{endepunkt}^{n/2}$ og $E_{endepunkt}^n$.

Algoritmer til at beregne forholdet mellem $E_{endepunkt}^n$ og $E_{endepunkt}^{2n}$ for Forbedret Euler, Heuns metode og RK4 findes i appendiks B.1.

Til at beregne forholdet mellem E_{max}^n og E_{max}^{2n} for Eulers metode anvendes algoritme 6. Inputparameterne er de samme som ved algoritme 5.

ALGORITME 6 (Forholdet mellem E_{max}^n og E_{max}^{2n} for Eulers metode). $fejlmax_euler := proc(f, y, xnul, ynul, a, b, m, ABC)$ Digits := ABC; *app* := []; *fejlmax* := []; *forhold* := [1]; for i to m do app := [op(app), eulersmetode(f, xnul, ynul, a, b, 2ⁱ, ABC)];end do; for n from 1 to m do fejl := [];ap := app[n]; $k := 2^n + 1;$ forl from 1 to k do fejl := [op(fejl), abs(evalf(y(ap[1][l]) - ap[2][l]))];end do; fejlmax := [op(fejlmax), max(fejl)];end do; for i from 2 to m do $forhold := \left[op(forhold), \frac{fejlmax[i-1]}{fejlmax[i]} \right];$ end do; for i from 2 to m do printf("%4d%1.5e%1.5f", 2ⁱ, fejlmax[i], forhold[i])end do; end proc;

Outputtet af algoritme 6 er en tabel med tre kollonner, indeholdende antal inddelinger *n*, den maksimale approksimationsfejl $E_{max}^n = max_{i \in [0,1,..,n]}(|y(x_i) - app(x_i)|)$ samt forholdet mellem $E_{max}^{n/2}$ og E_{max}^n . Algoritmer til at beregne forholdet mellem E_{max}^n og E_{max}^{2n} for Forbedret Euler, Heuns metode og RK4 findes i appendiks B.2.

Eksempel 7.1

I det første eksempel betragtes det simple begyndelsesværdiproblemet

$$\frac{dy}{dx} = x^2, \qquad 0 \le x \le 5, \qquad y_0 = 1$$
 (7.3)

hvor den eksakte løsning er givet ved

$$y = \frac{1}{3}x^3 + 1$$

Højresiden og den eksakte løsning defineres i Maple med kommandoerne

$$f := (x, y) \to x^2$$

$$y := x \to \frac{1}{3}x^3 + 1$$

Først betragtes forholdet mellem $E_{endepunkt}^{n}$ og $E_{endepunkt}^{2n}$. For Eulers metode anvendes algoritme 5 med 12 gennemkørsler. For Forbedret Euler, Heuns metode og RK4 anvendes algoritme 9, 10 og 11. Outputs for de fire algoritmer ses på tabel 7.1, 7.2, 7.3 og 7.4.

n	Appr.	Eksakt	Fejl	Forhold
4	28.34375	42.66667	1.43229e+01	1.81818
8	35.17969	42.66667	7.48698e+00	1.91304
16	38.84180	42.66667	3.82487e+00	1.95745
32	40.73389	42.66667	1.93278e+00	1.97895
64	41.69519	42.66667	9.71476e-01	1.98953
128	42.17966	42.66667	4.87010e-01	1.99478
256	42.42284	42.66667	2.43823e-01	1.99739
512	42.54468	42.66667	1.21991e-01	1.99870
1024	42.60565	42.66667	6.10153e-02	1.99935
2048	42.63615	42.66667	3.05126e-02	1.99967
4096	42.65141	42.66667	1.52575e-02	1.99984

Tabel 7.1. Tabel for fejl i endepunkt ved Eulers metode; *fejlendepunkt_euler*(*f*, *y*, 0, 1, 0, 5, 12, 30).

n	Appr.	Eksakt	Fejl	Forhold
4	42.01562	42.66667	6.51042e-01	4.00000
8	42.50391	42.66667	1.62760e-01	4.00000
16	42.62598	42.66667	4.06901e-02	4.00000
32	42.65649	42.66667	1.01725e-02	4.00000
64	42.66412	42.66667	2.54313e-03	4.00000
128	42.66603	42.66667	6.35783e-04	4.00000
256	42.66651	42.66667	1.58946e-04	4.00000
512	42.66663	42.66667	3.97364e-05	4.00000
1024	42.66666	42.66667	9.93411e-06	4.00000
2048	42.66666	42.66667	2.48353e-06	4.00000
4096	42.66667	42.66667	6.20882e-07	4.00000

Tabel 7.2. Tabel for fejl i endepunkt ved Forbedret Euler; *fejlendepunkt_forbeuler*(*f*, *y*, 0, 1, 0, 5, 12, 30).

n	Appr.	Eksakt	Fejl	Forhold
4	43.96875	42.66667	1.30208e+00	4.00000
8	42.99219	42.66667	3.25521e-01	4.00000
16	42.74805	42.66667	8.13802e-02	4.00000
32	42.68701	42.66667	2.03451e-02	4.00000
64	42.67175	42.66667	5.08626e-03	4.00000
128	42.66794	42.66667	1.27157e-03	4.00000
256	42.66698	42.66667	3.17891e-04	4.00000
512	42.66675	42.66667	7.94729e-05	4.00000
1024	42.66669	42.66667	1.98682e-05	4.00000
2048	42.66667	42.66667	4.96705e-06	4.00000
4096	42.66667	42.66667	1.24176e-06	4.00000

Tabel 7.3. Tabel for fejl i endepunkt ved Heuns metode; *fejlendepunkt_heun*(*f*, *y*, 0, 1, 0, 5, 12, 30).

n	Appr.	Eksakt	Fejl	Forhold
4	42.66667	42.66667	0.00000e+00	Inf
8	42.66667	42.66667	2.00000e-28	0.00000
16	42.66667	42.66667	2.00000e-28	1.00000
32	42.66667	42.66667	5.00000e-28	0.40000
64	42.66667	42.66667	9.00000e-28	0.55556
128	42.66667	42.66667	2.00000e-27	0.45000
256	42.66667	42.66667	3.90000e-27	0.51282
512	42.66667	42.66667	7.90000e-27	0.49367
1024	42.66667	42.66667	7.90000e-27	1.00000
2048	42.66667	42.66667	1.20000e-27	6.58333
4096	42.66667	42.66667	4.00000e-28	3.00000
8192	42.66667	42.66667	0.00000e+00	Inf

Tabel 7.4. Tabel for fejl i endepunkt ved RK4; *fejlendepunkt_RK4*(*f*, *y*, 0, 1, 0, 5, 12, 30).

Betragtes tabel 7.1, ses det at forholdet mellem fejlene i endepunkterne går mod 2, når antallet af inddelinger fordobles, som teorien antyder. Af tabel 7.2 og 7.3, ses det ligeledes at fejlforholdet går mod 4, som forventet. For RK4 ses det på tabel 7.4 at fejlforholdet ikke går mod 16. Dog er RK4 meget hurtig til at ramme den eksakte løsning, hvor fejlen allerede ved 4 inddelinger er lig nul.

Nu behandles forholdet mellem E_{max}^n og E_{max}^{2n} for begyndelsesværdiproblemet (7.3). For Eulers metode anvendes algoritme 6, for Forbedret Euler anvendes algoritme 12, for Heuns metode algoritme 13 og for RK4 algoritme 14. Alle algoritmerne køres igennem 16 gange. Outputs for algoritmerne ses på tabel 7.5, 7.6, 7.7 og 7.8.

n	E_{max}^n	Forhold	n	E_{max}^n	Forhold
4	1.43229e+01	1.81818	4	6.51042e-01	4.00000
8	7.48698e+00	1.91304	8	1.62760e-01	4.00000
16	3.82487e+00	1.95745	16	4.06901e-02	4.00000
32	1.93278e+00	1.97895	32	1.01725e-02	4.00000
64	9.71476e-01	1.98953	64	2.54313e-03	4.00000
128	4.87010e-01	1.99478	128	6.35783e-04	4.00000
256	2.43823e-01	1.99739	256	1.58946e-04	4.00000
512	1.21991e-01	1.99870	512	3.97364e-05	4.00000
1024	6.10153e-02	1.99935	1024	9.93411e-06	4.00000
2048	3.05126e-02	1.99967	2048	2.48353e-06	4.00000
4096	1.52575e-02	1.99984	4096	6.20882e-07	4.00000
8192	7.62908e-03	1.99992	8192	1.55220e-07	4.00000
16384	3.81462e-03	1.99996	16384	3.88051e-08	4.00000
32768	1.90733e-03	1.99998	32768	9.70128e-09	4.00000
65536	9.53669e-04	1.99999	65536	2.42532e-09	4.00000



bedret Euler, der viser output af *fejlmax_forbeuler*(*f*, *y*, 0, 1, 0, 5, 16, 30).

n	E_{max}^n	Forhold		n	E_{max}^n	Forhold	
4	1.30208e+00	4.00000	-	4	1.00000e-29	10.00000	
8	3.25521e-01	4.00000		8	2.00000e-28	0.05000	
16	8.13802e-02	4.00000		16	2.00000e-28	1.00000	
32	2.03451e-02	4.00000		32	5.00000e-28	0.40000	
64	5.08626e-03	4.00000		64	9.00000e-28	0.55556	
128	1.27157e-03	4.00000		128	2.00000e-27	0.45000	
256	3.17891e-04	4.00000		256	3.90000e-27	0.51282	
512	7.94729e-05	4.00000		512	7.90000e-27	0.49367	
1024	1.98682e-05	4.00000		1024	7.90000e-27	1.00000	
2048	4.96705e-06	4.00000		2048	1.30000e-27	6.07692	
4096	1.24176e-06	4.00000		4096	5.00000e-28	2.60000	
8192	3.10441e-07	4.00000		8192	5.00000e-28	1.00000	
16384	7.76102e-08	4.00000		16384	2.30000e-27	0.21739	
32768	1.94026e-08	4.00000		32768	3.40000e-27	0.67647	
65536	4.85064e-09	4.00000		65536	7.30000e-27	0.46575	
Tabel 7.7. Ta	bel for maksima	le fejl ved He	- Tabel	7.8. Tabe	for maks	imale fejl	ved
un	is metode, der v	viser output a	f	RK4,	der vise	er output	af
fejlmax_heun(f,y,0,1,0,5,16,30).				fejln	$ax_RK4(f, y, 0)$,1,0,5,16,30)	

Når fejlforholdet for den maksimale fejl betragtes, ses det at tabel 7.5, 7.6, 7.7, at dette opfører sig som forventet. Igen konvergerer fejlforholdet ikke for RK4 på tabel 7.8, hvor den maksimale fejl igen er meget lille allerede ved 4 skridt.

Eksempel 7.2

I det næste eksempel betragtes begyndelsesværdiproblemet fra eksempel 6.1, givet ved

$$\frac{dy}{dx} = -y\sin(x), \qquad 0 \le x \le 4\pi, \qquad y_0 = 1$$

hvor den eksakte løsning er givet ved

$$y = \frac{\exp(\cos(x))}{\exp(1)}$$

Højresiden og den eksakte løsning defineres i Maple

$$f := (x, y) \to -y \sin(x)$$

$$y := x \to \frac{\exp(\cos(x))}{\exp(1)}$$

Herefter anvendes algoritme 5 til at finde forholdet medllem $E_{endepunkt}^n$ og $E_{endepunkt}^{2n}$ for Eulers metode. Antal gennemkørsler af algoritmen sættes til m = 14. Outputtet kan ses på tabel 7.9. Det samme er gjort for Forbedret Euler, Heuns metode og RK4, vha. algoritme 9, 10 og 11. Disse outputs ses på tabel 7.10, 7.11 og 7.12.

n	Appr.	Eksakt	Fejl	Forhold
4	1.00000	1.00000	0.00000e+00	NaN
8	2.15327	1.00000	1.15327e+00	0.00000
16	0.03358	1.00000	9.66419e-01	1.19334
32	0.26932	1.00000	7.30680e-01	1.32263
64	0.53475	1.00000	4.65254e-01	1.57049
128	0.73378	1.00000	2.66220e-01	1.74763
256	0.85697	1.00000	1.43030e-01	1.86129
512	0.92578	1.00000	7.42247e-02	1.92698
1024	0.96218	1.00000	3.78215e-02	1.96250
2048	0.98091	1.00000	1.90922e-02	1.98099
4096	0.99041	1.00000	9.59202e-03	1.99043
8192	0.99519	1.00000	4.80755e-03	1.99520
16384	0.99759	1.00000	2.40667e-03	1.99759

Tabel 7.9. Tabel for fejl i endepunkt ved Eulers metode, $fejlendepunkt_euler(f, y, 0, 1, 0, 4 \cdot Pi, 14, 30)$.

n	Appr.	Eksakt	Fejl	Forhold
4	78.66988	1.00000	7.76699e+01	0.00000
8	0.28193	1.00000	7.18067e-01	108.16517
16	1.03331	1.00000	3.33127e-02	21.55538
32	1.01013	1.00000	1.01250e-02	3.29014
64	1.00143	1.00000	1.43197e-03	7.07068
128	1.00018	1.00000	1.84090e-04	7.77864
256	1.00002	1.00000	2.31707e-05	7.94494
512	1.00000	1.00000	2.90135e-06	7.98621
1024	1.00000	1.00000	3.62825e-07	7.99655
2048	1.00000	1.00000	4.53580e-08	7.99914
4096	1.00000	1.00000	5.66990e-09	7.99978
8192	1.00000	1.00000	7.08743e-10	7.99995
16384	1.00000	1.00000	8.85930e-11	7.99999

Tabel 7.10. Tabel for fejl i endepunkt ved Forbedret Euler, $fejlendepunkt_forbeuler(f, y, 0, 1, 0, 4 \cdot Pi, 14, 30)$.

n	Appr.	Eksakt	Fejl	Forhold
4	1.00000	1.00000	0.00000e+00	NaN
8	0.02155	1.00000	9.78449e-01	0.00000
16	0.83033	1.00000	1.69672e-01	5.76669
32	0.98521	1.00000	1.47931e-02	11.46967
64	0.99842	1.00000	1.58030e-03	9.36100
128	0.99981	1.00000	1.88758e-04	8.37209
256	0.99998	1.00000	2.33171e-05	8.09523
512	1.00000	1.00000	2.90593e-06	8.02398
1024	1.00000	1.00000	3.62968e-07	8.00601
2048	1.00000	1.00000	4.53625e-08	8.00150
4096	1.00000	1.00000	5.67004e-09	8.00038
8192	1.00000	1.00000	7.08747e-10	8.00009
16384	1.00000	1.00000	8.85931e-11	8.00002

Tabel 7.11. Tabel for fejl i endepunkt ved Heuns metode, $fejlendepunkt_heun(f, y, 0, 1, 0, 4 \cdot Pi, 14, 30)$.

n	Appr.	Eksakt	Fejl	Forhold
4	6.80785	1.00000	5.80785e+00	0.00000
8	0.61305	1.00000	3.86951e-01	15.00927
16	0.99509	1.00000	4.91402e-03	78.74422
32	0.99988	1.00000	1.15626e-04	42.49935
64	1.00000	1.00000	3.29261e-06	35.11674
128	1.00000	1.00000	1.00340e-07	32.81446
256	1.00000	1.00000	3.11559e-09	32.20587
512	1.00000	1.00000	9.72054e-11	32.05161
1024	1.00000	1.00000	3.03644e-12	32.01291
2048	1.00000	1.00000	9.48793e-14	32.00323
4096	1.00000	1.00000	2.96490e-15	32.00081
8192	1.00000	1.00000	9.26527e-17	32.00020
16384	1.00000	1.00000	2.89539e-18	32.00005

Tabel 7.12. Tabel for fejl i endepunkt ved RK4, $fejlendepunkt_RK4(f, y, 0, 1, 0, 4 \cdot Pi, 14, 30)$.

Tabel 7.9 viser at Eulers metode opfører sig som forventet. Tabel 7.10, 7.11 og 7.12, viser at fejlforhorholdet i endepunkterne for de tre andre metoder konvergerer mod det dobbelte, af hvad der forventes. Det er interessant at se om den samme tendens finder sted for forholdet mellem E_{max}^{n} og E_{max}^{2n} . For Eulers metode anvendes algoritme 6, for Forbedret Euler anvendes algoritme 12, for Heuns metode algoritme 13 og for RK4 algoritme 14. Alle algoritmerne køres igennem 14 gange. Outputs for algoritmerne ses på tabel 7.13, 7.14, 7.15 og 7.16.

n	E_{max}^n	Forhold	n	E_{max}^n	Forhold
4	8.64665e-01	0.00000	4	7.76699e+01	0.00000
8	2.46740e+00	0.35044	8	1.53097e+00	50.73236
16	9.66419e-01	2.55314	16	4.66607e-02	32.81079
32	7.30680e-01	1.32263	32	1.01250e-02	4.60846
64	4.65951e-01	1.56815	64	2.17966e-03	4.64523
128	2.68410e-01	1.73597	128	5.07677e-04	4.29340
256	1.44502e-01	1.85749	256	1.22714e-04	4.13705
512	7.50577e-02	1.92521	512	3.02212e-05	4.06054
1024	3.82638e-02	1.96158	1024	7.49591e-06	4.03169
2048	1.93200e-02	1.98053	2048	1.86666e-06	4.01569
4096	9.70765e-03	1.99019	4096	4.65762e-07	4.00775
8192	4.86580e-03	1.99508	8192	1.16328e-07	4.00388
16384	2.43590e-03	1.99753	16384	2.90678e-08	4.00194

Eulers metode, der viser output af $fejlmax_euler(f, y, 0, 1, 0, 4)$ *Pi*, 14, 30).

Tabel 7.13. Tabel for maksimale fejl ved Tabel 7.14. Tabel for maksimale fejl ved Forbedret Euler, der viser output af $fejlmax_forbeuler(f, y, 0, 1, 0, 4)$ *Pi*, 14, 30).

n	E_{max}^n	Forhold			
4	8.64665e-01	0.00000	n	E_{max}^n	Forhold
8	9.78449e-01	0.88371	4	5.80785e+00	0.00000
16	1.69672e-01	5.76669	8	3.86951e-01	15.00927
32	1.47931e-02	11.46967	16	4.91402e-03	78.74422
64	2.81542e-03	5.25432	32	1.15626e-04	42.49935
128	7.07235e-04	3.98088	64	3.29261e-06	35.11674
256	1.77002e-04	3.99564	128	1.57863e-07	20.85738
512	4.42636e-05	3.99881	256	9.96994e-09	15.83391
1024	1.10672e-05	3.99954	512	6.25817e-10	15.93109
2048	2.76688e-06	3.99987	1024	3.91795e-11	15.97305
4096	6.91726e-07	3.99997	2048	2.45064e-12	15.98744
8192	1.72932e-07	3.99998	4096	1.53223e-13	15.99395
16384	4.32331e-08	3.99999	8192	9.57823e-15	15.99703
Tabel 7.15. Tabel for maksimale fejl ved		16384	5.98694e-16	15.99852	
Heuns metode, der viser output		Tabel 7.16. Tal	oel for maks	imale fejl ved	
a	f fejlmax_heu	n(f, y, 0, 1, 0, 4	RK	4, der vise	er output af
<i>Pi</i> , 14, 30).			fej	lmax_RK4(f,y,0	$(1, 1, 0, 4 \cdot Pi, 14, 30).$

Forholdet mellem E_{max}^n og E_{max}^{2n} for de fire metoder på tabel 7.13, 7.14, 7.15 og 7.16, konvergerer altså mod de forventede værdier.

Referencer

Der er ikke anvendt referencer i dette kapitel.

Tennisbold med topspin 8

I dette kapitel bliver eksemplet beskrevet i kapitel 1 regnet igennem. Her ønskes det at beregne hvorledes en tennisbold påvirkes når den bliver påført et spin. I kapitel 1 blev følgende differentialligningssystem for tennisboldens position til tiden *t*, udledt

$$\frac{d^{2}x(t)}{dt^{2}} = -C_{D}\alpha \|v\| \frac{dx(t)}{dt} + \beta C_{M}\alpha \|v\| \frac{dz(t)}{dt}$$

$$\frac{d^{2}z(t)}{dt^{2}} = -9,82 - C_{D}\alpha \|v\| \frac{dz(t)}{dt} - \beta C_{M}\alpha \|v\| \frac{dx(t)}{dt}$$
(8.1)

hvor

$$\|v\| = \sqrt{x'(t)^2 + z'(t)^2}$$
$$\alpha = \frac{\pi d^2 \rho}{8m}$$
$$\beta = 1 \quad \text{for topspin}$$
$$\beta = -1 \quad \text{for bagspin}$$

og begyndelsesbetingelserne er givet ved

$$x(0) = 0, \quad z(0) = h, \quad \frac{dx}{dt}(0) = \|v_0\|\cos(\theta), \quad \frac{dz}{dt}(0) = \|v_0\|\sin(\theta)$$
(8.2)

 C_D og C_M er bestemmende for størrelsen af henholdsvis vindmodstanden og magnuskraften. Disse kan, som tidligere nævnt, bestemmes eksperimentielt.

8.1 Formen af C_D og C_M

Formler for C_D og C_M for en tennisbold, er undersøgt eksperimentielt af Štěpánek [1988]. Målinger blev udført for hastigheder $||v|| \in [13.6; 28]m/s$ og antal rotation $n \in [800, 3250]rpm$.

Štěpánek foretog målinger af C_D og C_M og plottede dem mod ||w|| / ||v||, som vist på figur 8.1. Bemærk at Štěpánek anvender notationen C_L for C_M .

På figur 8.1 ses ydermere en regressionslinje. Forskrifter for disse regressionslinjer udledte Štěpánek, til at være

$$\begin{split} C_D &= 0.508 + \left(\frac{1}{22.503 + 4.196 \left(\frac{\|w\|}{\|v\|}\right)^{-\frac{5}{2}}}\right)^{\frac{2}{5}} \\ C_M &= \frac{1}{2.202 + 0.981 \left(\frac{\|w\|}{\|v\|}\right)^{-1}} \end{split}$$



Figur 8.1. Štěpáneks målinger af C_D og $C_L(C_M)$ plottet mod ||w|| / ||v||

For en tennisbold kan det antages at rotationen ikke decelererer, hvilket medfører at vinkelhastigheden ||w|| er konstant.

8.2 Approksimationsmetoderne

Approksimationsmetoderne der er gennemgået i denne rapport, er alle beregnet til første ordens differentialligninger, derfor er det nødvendigt at omskrive (8.1). Systemet (8.1) omskrives til et system af 4 første ordens differentialligninger. Dette gøres ved at sætte $v_x = x'$ og $v_z = z'$, så

$$x' = v_{x}$$

$$v'_{x} = -C_{D}\alpha \sqrt{v_{x}^{2} + v_{z}^{2}} v_{x} + \beta C_{M}\alpha \sqrt{v_{x}^{2} + v_{z}^{2}} v_{z}$$

$$z' = v_{z}$$

$$v'_{z} = -9,82 - C_{D}\alpha \sqrt{v_{x}^{2} + v_{z}^{2}} v_{z} - \beta C_{M}\alpha \sqrt{v_{x}^{2} + v_{z}^{2}} v_{x}$$
(8.3)

Ligeledes omskrives begyndelsesværdierne til

$$x(0) = 0, \quad z(0) = H, \quad v_x(0) = v_0 \cos(\theta), \quad v_z(0) = v_0 \sin(\theta)$$
(8.4)

Algoritmerne til Eulers metode, Forbedret Euler, Heuns metode og RK4, gennemgået i de forrige kapitler, tager alle udgangspunkt i at vektoren y tilhører \mathbb{R} . Det er derfor nødvendigt at foretage et par modifikationer af algoritmerne, således at de kan anvendes for $y \in \mathbb{R}^4$, som skal anvendes i system (8.3). Nærmere betegnet en vektor (x, v_x, z, v_z) .

I algoritme 7 er algoritme 4 for RK4 tilpasset et $y \in \mathbb{R}^4$. Her er indgangene i funktionen f ændret, de steder hvor denne skal evalueres. Det forudsættes at f angives som en vektor. Dette gøres i Maple ved at skrive

 $f(x, y1, y2, y3, y4) := (x, y1, y2, y3, y4) \rightarrow Vector([f_1, f_2, f_3, f_4]);$

hvor f_1 , f_2 , f_3 og f_4 er funktioner af x, y1, y2, y3, y4. Derudover skal ynul angives som en vektor. Denne bliver for begyndelsesværdierne (8.4)

 $ynul := Vector([0, v_0 \cos(theta), H, v_0 \sin(theta)]);$

Udover at tilpasse algoritmen for $y \in \mathbb{R}^4$, er kommandoen evalf(..) anvendt, således at funktionens værdier bliver evalueret undervejs. Hvis evalf(..) ikke anvendes kommer algoritmen hurtigt til at arbejde med meget store eksakte udtryk, hvilket kan forøge tiden anvendt til beregning væsentligt. For at imødekomme det præcissionsproblem, der kan opstå ved at evaluere funktionsværdierne, er der til algoritmen tilføjet inputtet *ABC*, der angiver det antal betydende cifre algoritmen skal anvende. Standard for Maple er 10 betydende cifre.

```
ALGORITME 7 (RK4 med n inddelinger fpr y \in \mathbb{R}^4).
RK4_1 := proc(f, xnul, ynul, a, b, n, ABC)
   Digits := ABC;
  h := \frac{b-a}{n};
   x[0] := xnul;
   y[0] := ynul;
   Lx := [x[0]];
   Ly := [y[0]];
  for i from 0 to n - 1 do
x[i+1] := x[i] + h;
k1 := evalf(f(x[i], y[i][1], y[i][2], y[i][3], y[i][4]));
k2 := evalf(f(x[i] + \frac{h}{2}, y[i][1] + \frac{h}{2}k1[1], y[i][2] + \frac{h}{2}k1[2], y[i][3] + \frac{h}{2}k1[3], y[i][4] + \frac{h}{2}k1[4]))
k3 := evalf(f(x[i] + \frac{h}{2}, y[i][1] + \frac{h}{2}k2[1], y[i][2] + \frac{h}{2}k2[2], y[i][3] + \frac{h}{2}k2[3], y[i][3] + \frac{h}{2}k2[3]))
k4 := evalf(f(x[i] + h, h \cdot k3[1] + y[i][1], h \cdot k3[2] + y[i][2], h \cdot k3[3] + y[i][3], h \cdot k3[4] + y[i][4]))
y[i+1] := y[i] + \frac{1}{6} \cdot h \cdot (k1 + 2 \cdot k2 + 2 \cdot k3 + k4);
Lx := [op(Lx), x[i+1]];
Ly := [op(Ly), y[i+1]];
   end do;
   return {Lx, Ly}
end proc;
```

8.3 Tennisbold i vakuum

Det simpleste tilfælde er kurven for en tennisbold i vakuum uden spin. Det vil sige at hverken vindmodstanden eller magnuskraften påvirker tennisbolden. Sættes *D* og *M* begge lig nul, redu-

ceres (8.1) til det simple system

$$\frac{d^2 x(t)}{dt^2} = 0 \tag{8.5}$$

$$\frac{d^2 z(t)}{dt^2} = -9,82\tag{8.6}$$

med begyndelsesbetingelserne (8.2). Da systemet (8.5) er lineært kan Maple finde en generel løsning. Dette gøres ved at definere systemet (8.5) i Maple, med kommandoen

ode := diff(x(t), t, t) = 0, diff(z(t), t, t) = -g;

$$ode := \frac{d^2}{dt^2} x(t) = 0, \ \frac{d^2}{dt^2} z(t) = -g$$

Derefter defineres begyndelsesbetingelserne (8.2), ved

$$ics := x(0) = 0, z(0) = H, (D(x))(0) = v0 \cdot cos(theta), (D(z))(0) = v0 \cdot sin(theta);$$
$$ics := x(0) = 0, z(0) = H, D(x)(0) = v0 \cos(\theta), D(z)(0) = v0 \sin(\theta)$$

Dernæst anvendes kommandoen dsolve til at løse begyndelsesværdiproblemet

dsolve({ics,ode});

$$\left\{ x(t) = v0\cos(\theta) t, z(t) = -\frac{1}{2}gt^{2} + v0\sin(\theta) t + H \right\}$$

hvor g = 9,82. Vælges følgende begyndelsesværdier

$$v0 = 25m/s, \quad H = 1m, \quad \theta = 15^{\circ} = \frac{15\pi}{180}rad$$

kan tennis
boldens flyvetid beregnes ved at løse $\boldsymbol{z}(t)=0.$ Dette kan gøres i Maple med kommando
en

 $g := -9.82: v0 := 25: H := 1: theta := \frac{15 \cdot Pi}{180}: evalf(solve(-\frac{1}{2} \cdot g \cdot t^2 + v0 \cdot sin(theta) \cdot t + H = 0, t));$

-0.1397319392, 1.457547851

Det første tal -0.1397319392 er ikke interessant, da dette er skæringen med den negative *x*-akse. Flyvetiden for en tennisbold i vakuum uden spin er, med de valgte begyndelsesværdier, således 1,4576 sekunder. Dette anvendes til at plotte tennisboldens banekurve, med kommandoen

 $plot([t \cdot v0 \cdot \cos(theta), -\frac{1}{2} \cdot g \cdot t^2 + v0 \cdot \sin(theta) \cdot t + H, t = 0..1.457547])$



Figur 8.2. Plot af banekurve af tennisbold i vakuum, med begyndelsesværdier: $v_0 = 25m/s$, H = 1m og $\theta = 15^{\circ}$.

8.4 Tennisbold i luft

For en tennisbold i luft behandles to tilfælde; uden spin og med topspin. En tennisbold i luft uden spin påvirkes kun af tyngdekraften og luftmodstanden. Differentialligningssystemet i dette tilfælde svarer altså til systemet (8.3) med $\beta = 0$. Løsninger til de to tilfælde approksimeres i Maple vha. algoritme 7.

Indledningsvist defineres C_D og C_M , som funktioner af v_x og v_z , samt koefficienten *al pha* i Maple

$$C_{D} := (vx, vz) \to 0.508 + \left(\frac{1}{22.503 + 4.196 \left(\frac{\|w\|}{sqrt(vx^{2} + vz^{2})}\right)^{-\frac{5}{2}}}\right)^{\frac{4}{5}} :$$

$$C_{M} := (vx, vz) \to \frac{1}{2.202 + 0.981 \left(\frac{\|w\|}{sqrt(vx^{2} + vz^{2})}\right)^{-1}} :$$

$$alpha := \frac{Pi \cdot d^{2} \cdot rho}{8m} :$$

Herefter defineres et differentialligningssystem for en tennisbold i luft uden spin ($\beta = 0$) og et for en tennisbold med topspin ($\beta = 1$). Disse angives henholdvis *fluft* og *fspin*.

$$\begin{aligned} fluft &:= (t, x, vx, z, vz) \rightarrow Vector([vx, -CD(vx, vz)alpha \cdot sqrt(vx^{2} + vz^{2})vx), \\ vz, -g - CD(vx, vz)alpha \cdot sqrt(vx^{2} + vz^{2})vz]); \\ fspin &:= (t, x, vx, z, vz) \rightarrow Vector([vx, alpha \cdot sqrt(vx^{2} + vz^{2})(-CD(vx, vz)vx + CM(vx, vz)vz), \\ vz, -g - alpha \cdot sqrt(vx^{2} + vz^{2})(CD(vx, vz)vz + CM(vx, vz)vx)]); \end{aligned}$$

Ligeledes angives begyndelsesværdierne som en vektor. Denne angives ics.

$$ics := Vector([0, v0 \cdot \cos(theta), H, v0 \cdot \sin(theta)]);$$

Herefter vælges værdier for starthastigheden v_0 , starthøjden H, startvinklen θ og vinkelhastigheden ||w||. Disse vælges i tråd med værdierne anvendt for tennisbolden i vakuum

$$v0 = 25m/s, \quad H = 1m, \quad \theta = 15^{\circ} = \frac{15\pi}{180} rad, \quad ||w|| = 20m/s$$

Derudover haves følgende kendte værdier for tennisboldens diameter *d* og masse *m*, samt luftmassefylden ρ og tyngdeaccelerationen *g*.

$$d = 0,063m$$
, $m = 0,05kg$, $\rho = 1,29kg/m^3$, $g = 9,82m/s^2$

Herefter kan løsninger til de to tilfælde approksimeres vha. algoritme 7. Approksimationen vælges udført med n = 200 skridt over intervallet $t \in [0; 1, 45]$. Intervallets egentlige endepunkt er ikke kendt, derfor vælges flyvetiden 1,45 for tennisbolden i vakuum beregnet tidligere. Antal betydende cifre sættes til ABC = 30. Der anvendes dermed følgende kommandoer i Maple

 $Loesnluft := RK4_4(fluft, 0, ics, 0, 1.4, 200, 30) :$ $Loesnspin := RK4_4(fspin, 0, ics, 0, 1.4, 200, 30) :$

Outputtet af ovenstående kommandoer er på formen

$$\left\{ \begin{bmatrix} t_0, t_1, \dots, t_n \end{bmatrix}, \begin{bmatrix} x(t_0) \\ v_x(t_0) \\ z(t_0) \\ v_z(t_0) \end{bmatrix}, \begin{bmatrix} x(t_1) \\ v_x(t_1) \\ z(t_1) \\ v_z(t_1) \end{bmatrix}, \dots, \begin{bmatrix} x(t_n) \\ v_x(t_n) \\ z(t_n) \\ v_z(t_n) \end{bmatrix} \right\}$$
(8.7)

Vektorer indeholdende approksimationerne af x(t) og z(t), kan herefter udtrækkes ved at først at definere 4 tomme lister; en x og en z vektor hørende til *Loesnluft*, samt en x og en z vektor hørende til *Loesnluft*, samt en x og en z vektor hørende til *Loesnspin*. Dernæst anvendes en forløkke til at tilføje de relevante elementer fra *Loesnluft* og *Loesnspin* til de tomme lister. Dette kan gøres i Maple med følgende kommandoer

```
\begin{aligned} & Loesnspin\_x := [ ]; Loesnluft\_x := [ ]; Loesnspin\_z := [ ]; Loesnluft\_z := [ ]; \\ & \text{for } i \text{ from 1 to } 201 \text{ do} \\ & Loesnluft\_x := [op(Loesnluft\_x), Loesnluft[2][i][1]]; \\ & Loesnspin\_x := [op(Loesnspin\_x), Loesnspin[2][i][1]]; \\ & Loesnluft\_z := [op(Loesnluft\_z), Loesnluft[2][i][3]] \\ & Loesnspin\_z := [op(Loesnspin\_z), Loesnspin[2][i][3]]; \\ & \text{end do:} \end{aligned}
```

De approksimerede løsninger i de tre tilfælde (i vakuum samt i luft med og uden spin) er plottet på figur 8.3. Udover de approksimerede løsninger er længden af en tennisbane tilføjet samt placering og højde af nettet.



Figur 8.3. Plot af banekurve af tennisbold i vakuum, i luft uden spin og i luft med topspin, med begyndelsesværdier: $v_0 = 25m/s$, H = 1m, $\theta = 15^{\circ}$ og ||w|| = 20m/s.

Beregning af flyvetid og skudlængde

Ved tennisbolden i vakuum blev flyvetiden beregnet til at være 1,4576 sekunder, dermed må flyvelængden være

$$l_vakum = x(1,4576) = v0\cos(\theta)\,1,4576 = 35,2m\tag{8.8}$$

Flyvetiden blev fundet ved at løse ligningen z(t) = 0. Der haves intet udtryk for z(t) ved approksimationerne *Loensluft* og *Loesnspin*. Istedet konstrueres en algoritme der kan finde indekset for det skridt hvor z(t) er tilnærmelsesvis nul. Herefter anvendes det fundne indeks til at udtrække flyvetiden og flyvelængden fra approksimationsoutputtet (8.7). Algoritmen tager en vektor med approksimerede værdier for z(t) som input, f.eks. listerne *Loesnluft_z* og *Loesnspin_z* fra tidligere. Algoritmen kan i Maple skrives som

ALGORITME 8 (Algoritme til at finde indeks n hvor $z(t_n) = 0$). findindeks := proc(L :: list) n := 1; $while L[n] > 10^{-5} do$ n := n + 1 end do; return n;end proc : Algoritme 8 anvendes nu til at bestemme flyvetid og flyvelængde for en tennisbold henholdvis med og uden spin

nluft := *findindeks*(Loesnluft_z) : *t_luft* := Loesnluft[1][*nluft*]; *l_luft* := Loesnluft_x[*nluft*];

 $t_luft := 1.328000000$

 $l_luft := 22.11153650$

nspin := *findindeks*(Loesnspin_z) : *t_spin* := Loesnspin[1][*nspin*]; *l_spin* := Loesnspin_x[*nspin*];

 $t_spin := 0.9520000000$ $l_spin := 17.35194367$

En tennisbold der sendes afsted med 25m/s med en vinkel på 15° med *x*-aksen, vil altså flyve 22, 11 meter på 1, 328 sekunder. Tilføres der yderligere et topspin, der får bolden til at rotere med 20m/s vil den flyve 17, 354 meter på 0, 952 sekunder. Tilførslen af topspin får altså tennisbolden til at tilbagelægge en kortere distance på kortere tid. Derudover antyder figur 8.3 at topspinnet får bolden til at dykke hurtigere når den nærmer sig modspillerens banehalvdel. For at illustrere dette yderligere regnes flere tilfælde hvor starthastighed og vinkel ændres. Beregningerne foretages efter samme metode som ovenfor.

Det første eksempel der betragtes, er en sammenligning af banekurver for en tennisbold i luft med og uden spin, hvor startvinklen θ og skudlængden har samme værdier for de to tilfælde. Sæt $\theta = 6^{\circ}$ for begge tilfælde. For tennisbolden uden spin sættes $v_0 = 32m/s$. For tennisbolden med spin sættes $v_0 = 49$, 1m/s og ||w|| = 17m/s. Derved fås banekurver som plottet på figur 8.4.



Figur 8.4. Banekurve for tennisbold i luft (blå) med begyndelsesværdier: $v_0 = 32m/s$ og $\theta = 6^\circ$, samt banekurve for tennisbold med topspin (sort) med begyndelsesværdier $v_0 = 49, 1m/s, ||w|| = 17m/s, \theta = 6^\circ$.

Ved beregning af flyvetid og skudlængde for de to tilfælde fås følgende værdier

 $t_luft := 0.8880000000$ $l_luft := 20.42289024$ $t_spin := 0.5670000000$ $l_spin := 20.42375238$

Værdierne af *l_luft* og *l_spin* viser at tennisbolden i begge tilfælde flyver tilnærmelsesvis 20,42 meter. Der er dog markant forskel på flyvetiden for de to tilfælde, hvor tennisbolden uden spin bruger 0,888 sekunder og med spin bruger 0,567 på at tilbagelægge de 20,42 meter. Dette skyldes

den høje starthastighed, v_0 , der kan påføres tennisbolden med spin. Det fremgår ydermere af figur 8.4 at den vinkel tennisbolden med spin rammer modstanderens banehalvdel med, er skarpere end for tilfældet uden spin, hvilket gør bolden sværere at returnere for modstanderen.

I det næste eksempel holdes starthastigheden og startvinklen fast og startvinkelhastigheden varieres. Starthastigheden, v_0 , sættes til 30m/s og startvinklen, θ , sættes til 15° . Der plottes herefter banekurver for vinkelhastigheder $||w|| \in [0, 5, 10, 15, 20]m/s$. De approksimerede banekurver er plottet på figur 8.5



Figur 8.5. Banekurver for fast $v_0 = 30m/s$ og $\theta = 15^\circ$, samt varieret værdi af $||w|| \in [0, 5, 10, 15, 20]m/s$.

Af figur 8.5 ses det at, hvis en tennisspiller ønsker at sende bolden afsted fra en 15 graders vinkel med 30m/s, skal han tilføre bolden et topspin for at holde bolden indenfor banens lægnde. Det ses at jo højere topspin, jo kortere distance tilbagelægger bolden.

Referencer

Dette kapitel er primært skrevet på baggrund af [Gander og Hřebíček, 2004, s. 27-30]. Afsnit 8.1 er skrevet på baggrund af [Štěpánek, 1988, s. 140].

Afrunding 9

Dette projekt har beskæftiget sig med metoder til at approksimere løsninger til systemer af første ordens differentialligninger, på formen

$$\frac{dy}{dx} = f(x, y) \tag{9.1}$$

hvor *y* kan være en vektor. Det primære fokus har været rettet mod metoderne; Eulers metode, Forbedret Euler, Heuns metode og Runge-Kutta metoden af orden 4, RK4. Det er vist at Eulers metode er af orden 1, og Forbedret Euler samt Heuns metode er af orden 2. Metodernes orden blev illustreret ved numeriske eksperimenter, hvor følgende fejlforhold blev betragtet

$$\frac{E_{endepunkt}^{n}}{E_{endepunkt}^{2n}} \quad \text{og} \quad \frac{E_{max}^{n}}{E_{max}^{2n}} \tag{9.2}$$

For Eulers metode gik ovenstående forhold mod 2, dvs. at fejlen blev halveret ved en fordobling af antallet af skridt *n*. For Forbedret Euler og Heuns metode var dette tal 4, og for RK4 16. RK4 er givet vist den mest effektive metode af de fire. Det skal dog bemærkes at RK4, kræver en højere grad af differentiabilitet for funktionen f i (9.1).

Afslutningsvist blev et større eksempel gennemregnet. Her blev det illustreret hvorledes banekurven for en tennisbold påvirkes, når tennisbolden tilføres et topspin. Her viste beregningerne at givet en starthastighed v_0 vil tilførslen af et topspin forkorte både skudlængden og flyvetid. Det betyder at en tennisbold med topspin, kan sendes afsted med en større hastighed, i forhold til en bold uden spin, og stadig ramme jorden inden modstanderens baglinje.

Det vil ydermere være interessant at betragte det modsatte tilfælde, hvor topspinnet udskiftes med et bagspin. Et bagspin anvendes af golfspillere, der netop ønsker den modsatte effekt af topspinnet, således at golfbolden flyver længere ved en givet starthastighed, end den ville have gjort uden spin.


A.1 Definitioner

Følgende definition stammer fra [Fitzpatrick, 2006, s. 245]

DEFINITION A.1

Lad $S \subseteq \mathbb{R}$, $n \in \mathbb{N}$ og $f_n : S \to \mathbb{R}$. $\{f_n\}_{n \in \mathbb{N}}$ er **ligelig konvergent** med grænse $f : S \to \mathbb{R}$ hvis der for alle $\epsilon > 0$ eksisterer et $N \in \mathbb{N}$ således at

 $|f_n(x) - f(x)| < \epsilon$ for alle $n \ge N$ og $x \in S$

A.2 Partielle afledede

Følgende sætning og korollar er fra [Rudin, 1976, s. 235-236]

SÆTNING A.2

Lad *f* være defineret på en åben mængde $E \subset \mathbb{R}^2$. Antag at $D_1 f$, $D_{21} f$ og $D_2 f$ eksisterer for alle punkter i *E* og $D_{21} f$ er kontinuert i et givet punkt $(a, b) \in E$, så gælder det at $D_{12} f$ eksisterer i (a, b) og

$$(D_{12}f)(a,b) = (D_{21}f)(a,b)$$

KOROLLAR A.3 Anvend notation som i sætning A.2. Hvis $f \in C^2(E)$, så er

$$(D_{12}f) = (D_{21}f)$$

Taylors sætning

Følgende sætning stammer fra [Edwards og Penney, 2008, s. 748]

SÆTNING A.4 (Taylors sætning)

Antag at de n + 1 første afledede af funktionen f eksisterer på et interval I. Lad $a, b \in I$, så gælder det at

$$f(b) = f(a) + f'(a)(b-a) + \frac{f''(a)}{2!}(b-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(b-a)^n + \frac{f^{(n+1)}(z)}{(n+1)!}(b-a)^{n+1}$$

for et tal z mellem $a \circ g b$.

Definition A.5 og sætning A.6 er fra [Trench, 2013, s. 349-351].

DEFINITION A.5

Lad $r \ge 1$ og lad alle partielle afledede af orden $\le r - 1$ af funktionen f være differentiable i en omegn af $x_0 \in \mathbb{R}^n$. Den r'te afledede af f i x_0 , er defineret ved

$$d_{x_0}^{(r)}f = \sum_{i_1,i_2,\dots,i_r=1}^n \frac{\partial^r f(x_0)}{\partial x_{i_r} \partial x_{i_{r-1}} \cdots \partial x_{i_1}} dx_{i_1} dx_{i_2} \cdots dx_{i_r}$$

ydermere defineres

$$(d_{x_0}^{(r)})f = f(x_0) \tag{A.1}$$

SÆTNING A.6 (Taylors sætning for funktion af flere variable)

Lad $x_0, x \in \mathbb{R}$ og L være et linjesegment der forbinder disse to punkter. Antag at funktionen f og funktionens partielle afledede af orden $\leq k$ er differentiable i punkterne x_0, x og på L. Så gælder det at

$$f(x) = \sum_{r=0}^{k} \frac{1}{r!} (d_{x_0}^{(r)} f)(x - x_0) + \frac{1}{(k+1)!} (d_{\widetilde{x}}^{(k+1)} f)(x - x_0)$$

for et $\tilde{x} \neq x_0$, *x* på *L*.

Algoritmer **B**

B.1 Forholdet mellem fejl i endepunkt

Algoritmer til at beregne forholdet mellem $E_{endepunkt}^{n}$ og $E_{endepunkt}^{2n}$ for

- Forbedret Euler algoritme 9
- Heuns metode algoritme 10
- RK4 algoritme 11

ALGORITME 9 (Forholdet mellem $E_{endepunkt}^{n}$ og $E_{endepunkt}^{2n}$ for Forbedret Euler).

```
fejlendepunkt_forbeuler := proc(f, y, xnul, ynul, a, b, m, ABC)
```

 $\begin{array}{l} Digits := ABC;\\ eksakt := y(b);\\ app := []; fejl := []; forhold := [1];\\ for i from 1 to m do\\ app := [op(app), forbed reteuler_1(f, xnul, ynul, a, b, 2^i, ABC)];\\ fejl := [op(fejl), abs(evalf(eksakt - app[i][2][2^i + 1]))];\\ end do;\\ for i from 2 to m do\\ forhold := \left[op(forhold), \frac{fejl[i-1]}{fejl[i]}\right];\\ end do;\\ for i from 2 to m do\\ printf("%4d% + 1.5f % + 1.5f % 1.5e\% 1.5f n", 2^i, app[i][2][2^i + 1], eksakt, fejl[i], forhold[i])\\ end do;\\ end proc;\\ \end{array}$

ALGORITME 10 (Forholdet mellem $E_{endepunkt}^n$ og $E_{endepunkt}^{2n}$ for Heuns metode). $fejlendepunkt_heun := proc(f, y, xnul, ynul, a, b, m, ABC)$ Digits := ABC; eksakt := y(b);app := []; fejl := []; forhold := [1]; for i from 1 to m do $app := [op(app), heun_1(f, xnul, ynul, a, b, 2^i, ABC)];$ $fejl := [op(fejl), abs(evalf(eksakt - app[i][2][2^i + 1]))];$ end do; for i from 2 to m do forhold := $\left[op(forhold), \frac{fejl[i-1]}{fejl[i]} \right];$ end do; for i from 2 to m do $printf("\%4d\% + 1.5f\% + 1.5f\% 1.5e\% 1.5fn", 2^{i}, app[i][2][2^{i} + 1], eksakt, fejl[i], forhold[i])$ end do; end proc;

ALGORITME 11 (Forholdet mellem $E_{endevunkt}^n$ og $E_{endevunkt}^{2n}$ for RK4). $fejlendepunkt_RK4 := proc(f, y, xnul, ynul, a, b, m, ABC)$ Digits := ABC;eksakt := y(b);app := []; fejl := []; forhold := [1];for i from 1 to m do $app := [op(app), RK4_1(f, xnul, ynul, a, b, 2^i, ABC)];$ $fejl := [op(fejl), abs(evalf(eksakt - app[i][2][2^i + 1]))];$ end do; for i from 2 to m do $forhold := \left[op(forhold), \frac{fejl[i-1]}{fejl[i]} \right];$ end do; for i from 2 to m do $printf("\%4d\% + 1.5f\% + 1.5f\% 1.5e\% 1.5fn", 2^{i}, app[i][2][2^{i} + 1], eksakt, fejl[i], forhold[i])$ end do; end proc;

B.2 Forholdet mellem maksimal fejl

Algoritmer til at beregne forholdet mellem E_{max}^n og E_{max}^{2n} for

- Forbedret Euler algoritme 12
- Heuns metode algoritme 13
- RK4 algoritme 14

```
ALGORITME 12 (Forholdet mellem E_{max}^n og E_{max}^{2n} for Forbedret Euler).
fejlmax\_euler := proc(f, y, xnul, ynul, a, b, m, ABC)
                  Digits := ABC;
                  app := []; fejlmax := []; forhold := [1];
                  for i to m do
                     app := [op(app), forbedreteuler_1(f, xnul, ynul, a, b, 2i, ABC)]
                  end do;;
                  for n from 1 to m do
                    fejl := [];
                    ap := app[n];
                    k := 2\hat{n} + 1;
                    forl from 1 to k do
                       fejl := [op(fejl), abs(evalf(y(ap[1][l]) - ap[2][l]))]
                     end do;;
                     fejlmax := [op(fejlmax), max(fejl)]
                  end do;;
                  for i from 2 to m do
                    forhold := \left[ op(forhold), \frac{fejlmax[i-1]}{fejlmax[i]} \right]
                  end do;;
                  for i from 2 to m do
                printf("%4d\%1.5e\%1.5f", 2<sup>i</sup>, fejlmax[i], forhold[i])
                  end do;
      end proc;
```

```
ALGORITME 13 (Forholdet mellem E_{max}^n og E_{max}^{2n} for Heuns metode).
    fejlmax\_euler := proc(f, y, xnul, ynul, a, b, m, ABC)
                      Digits := ABC;
                      app := []; fejlmax := []; forhold := [1];
                      for i to m do
                        app := [op(app), heun_1(f, xnul, ynul, a, b, 2i, ABC)]
                      end do;;
                      for n from 1 to m do
                        fejl := [];
                        ap := app[n];
                        k := 2\hat{n} + 1;
                        forl from 1 to k do
                           fejl := [op(fejl), abs(evalf(y(ap[1][l]) - ap[2][l]))]
                        end do;;
                        fejlmax := [op(fejlmax), max(fejl)]
                      end do;;
                      for i from 2 to m do
                        forhold := \left[ op(forhold), \frac{fejlmax[i-1]}{fejlmax[i]} \right]
                      end do;;
                      for i from 2 to m do
                   printf("\%4d\%1.5e\%1.5f", 2^i, fejlmax[i], forhold[i])
                      end do;
          end proc;
```

```
ALGORITME 14 (Forholdet mellem E_{max}^n og E_{max}^{2n} for Eulers metode).
    fejlmax\_euler := proc(f, y, xnul, ynul, a, b, m, ABC)
                      Digits := ABC;
                      app := []; fejlmax := []; forhold := [1];
                      for i to m do
                         app := [op(app), RK4_1(f, xnul, ynul, a, b, 2i, ABC)]
                      end do;;
                      for n from 1 to m do
                         fejl := [];
                         ap := app[n];
                         k := 2\hat{n} + 1;
                        for l from 1 to k do
                           fejl := [op(fejl), abs(evalf(y(ap[1][l]) - ap[2][l]))]
                         end do;;
                         fejlmax := [op(fejlmax), max(fejl)]
                      end do;;
                      for i from 2 to m do
                        forhold := \left[ op(forhold), \frac{fejlmax[i-1]}{fejlmax[i]} \right]
                      end do;;
                      for i from 2 to m do
                    printf("%4d\%1.5e\%1.5f", 2^i, fejlmax[i], forhold[i])
                      end do;
          end proc;
```

- Bauer, W. og Westfall, G. D. (2011). *University Physics with Modern Physics*. Number ISBN: 978-0-07-285736-8. McGraw-Hill.
- Butcher, J. C. (1965). On the attainable order of runge-kutta methods. *Mathematics of Computation*, 19(91):408–417.
- Butcher, J. C. (1985). The nonexistence of ten-stage eight order explicit runge-kutta methods. BIT Numerical Mathematics, 25(3):521–540.
- Butcher, J. C. (2008). *Numerical Methods for Ordinary Differential Equations*. Number ISBN: 978-0-470-72335-7. John Wiley & Sons, Ltd.
- Edwards, C. H. og Penney, D. E. (2008). *Calculus Early Transcendentals*, volume 7. Pearson Education, Inc.
- Faber, T. E. (1995). *Fluid dynamics for physicists*. Number ISBN: 0-521-41943-3. Cambridge University Press.
- Fitzpatrick, P. (2006). *Advanced calculus*. Number ISBN:978-0-8218-4791-6. American Mathematical Soc.
- Gander, W. og Hřebíček, J. (2004). Solving Problems in Scientific Computing Using Maple and MATLAB. Number ISBN: 3-540-21127-6. Springer.
- Gautschi, W. (1997). *Numerical Analysis An Introduction*. Number ISBN:0-8176-3895-4. Birkhäuser Boston.
- Møller, J. S. (2012). Ordinary differential equations: An introduction for mathematicians : Lecture notes. Online. URL: http://aula.au.dk/main/document/document.php?cidReq=IMFDIFFE12.
- Petrovski, I. G. (1973). *Ordinary Differential Equations*. Number ISBN: 0-486-64683-1. Dover Publications, Inc.
- Rudin, W. (1976). Principles of mathematical analysis, volume 3. McGraw-Hill New York.
- Trench, W. F. (2013). Introduction to real analysis. Online: Free Edition 2.04. URL: http: //ramanujan.math.trinity.edu/wtrench/texts/TRENCH_REAL_ANALYSIS.PDF.
- Štěpánek, A. (1988). The aerodynamics of tennisballs -the topspin lob. *American Journal of Physics*, 56(2):138–142.