

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

REKENAFDELING

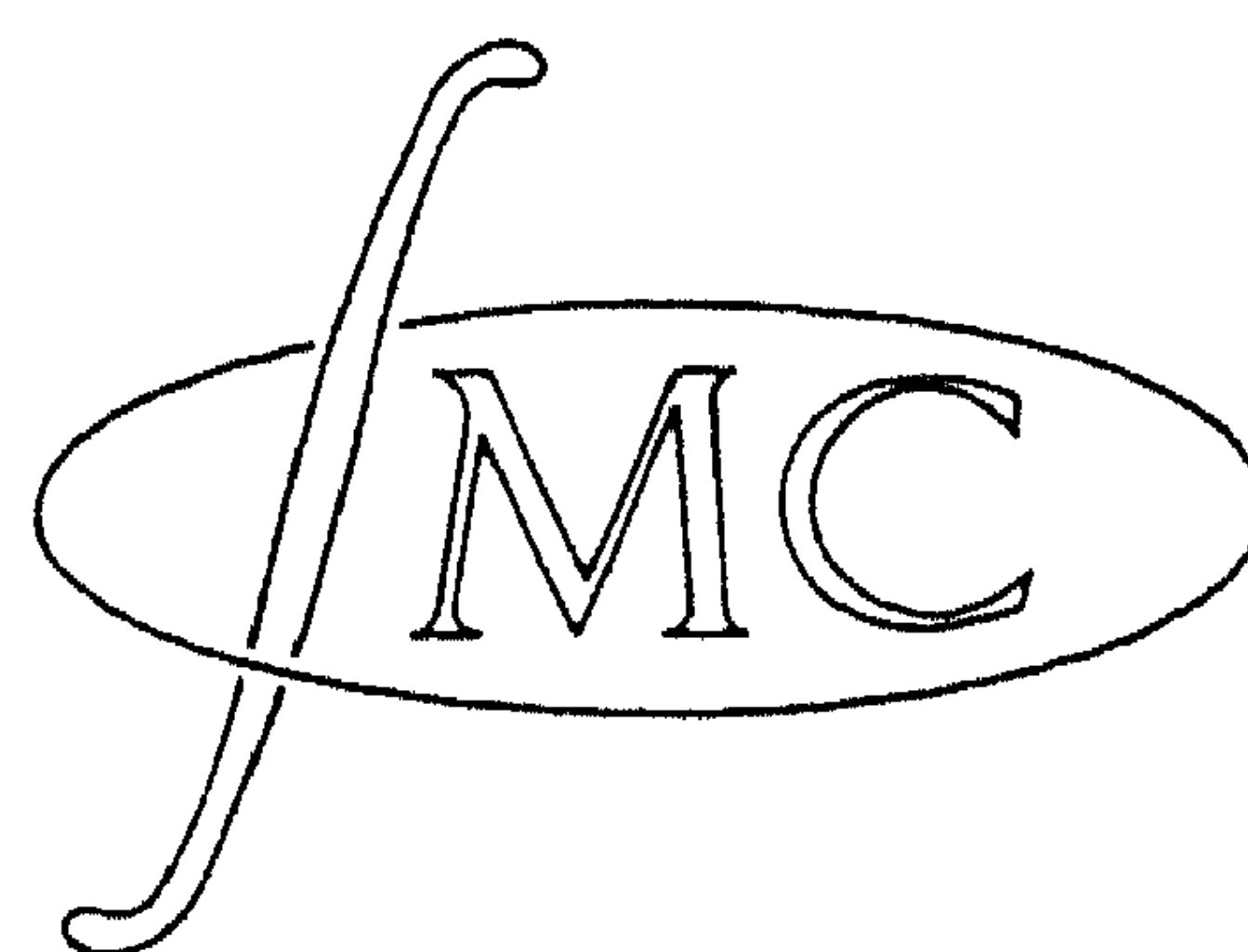
CURSUS: WETENSCHAPPELIJK REKENEN B

Onderdeel: "Numerieke Analyse"

door

Prof.dr. E. van Spiegel

Deel Ia



1964

Printed at the Mathematical Centre at Amsterdam, 49, 2nd Boerhaavestraat.
The Netherlands.

The Mathematical Centre, founded the 11th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications, and is sponsored by the Netherlands Government through the Netherlands Organization for Pure Scientific Research (Z.W.O.) and the Central National Council for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

NUMERIEKE ANALYSE

door

Prof.dr.E.van Spiegel

cursus 1964

Inleiding

In deze cursus zal een aantal numerieke methoden behandeld worden. Sommige hiervan zijn reeds behandeld in de cursus Wetenschappelijk Rekenaar A. De bedoeling van deze cursus is dan ook tweeledig:

- 1° Nieuwe numerieke processen te leren.
- 2° Van de reeds bekende processen de achtergronden te bespreken.

1. Integratieformules van Newton-Cotes

Voor de afleiding van de Newton-Cotes integratieformules gaan we uit van de interpolatieformule van Lagrange.

Van een functie $f(x)$ gedefinieerd op het interval (a,b) zijn in de punten $x_0, x_1, \dots, x_n$ de funktiewaarden $f(x_0), f(x_1), \dots, f(x_n)$ bekend.

We zoeken, als benaderingsfunctie voor $f(x)$, een polynoom $p(x)$ van de graad $\leq n$ waarvoor geldt

$$p(x_i) = f(x_i) \quad i = 0, 1, 2, \dots, n.$$

Daartoe schrijven we $p(x)$ in de vorm

$$p(x) = \sum_{k=0}^n \ell_k(x) \cdot f(x_k).$$

Als het ons gelukt om voor $\ell_k(x)$ een n -de graads polynoom te nemen waarvoor geldt

$$\ell_k(x_k) = 1 \tag{1.1}$$

$$\ell_k(x_j) = 0 \quad \text{als } j \neq k$$

dan is $p(x)$ juist het gezochte polynoom, dat in de $n+1$ basis-punten $x_0, x_1, \dots, x_n$ met $f(x)$ overeenstemt.

Opmerking

Met behulp van het zogenaamde Kronecker symbool

$$\delta_{ij} = \begin{cases} 0 & \text{als } i \neq j \\ 1 & \text{als } i = j \end{cases} \tag{1.2}$$

kunnen we de voorwaarden voor $\ell_k(x)$ kortweg als volgt weergeven

$$\ell_k(x_j) = \delta_{kj}.$$

Het n -de graads polynoom $\ell_k(x)$ dat aan (1.1) voldoet, is

$$\ell_k(x) = \frac{(x-x_0) \dots (x-x_{k-1})(x-x_{k+1}) \dots (x-x_n)}{(x_k-x_0) \dots (x_k-x_{k-1})(x_k-x_{k+1}) \dots (x_k-x_n)}. \quad (1.3)$$

Met behulp van de afkorting

$$\pi(x) = (x-x_0)(x-x_1) \dots (x-x_n) = \prod_{i=0}^n (x-x_i) \quad (1.4)$$

kunnen we in plaats van (1.3) ook schrijven

$$\ell_k(x) = \frac{\pi(x)}{(x-x_k) \pi'(x_k)}.$$

Het polynoom $\ell_k(x)$ heet het k -de Lagrange-polynoom.

De interpolatieformule van Lagrange luidt nu als volgt:

$$f(x) = \sum_{k=0}^n \ell_k(x) \cdot f(x_k) + E(x). \quad (1.5)$$

$E(x)$ heet de restterm.

We zullen nu voor $E(x)$ een uitdrukking afleiden. Omdat $x_0, x_1, \dots, x_n$ nulpunten van $E(x)$ zijn kunnen we schrijven

$$E(x) = \pi(x) \cdot K(x) \quad (1.6)$$

en moeten we een uitdrukking voor $K(x)$ afleiden. Hiertoe beschouwen we de functie $S(t)$ gedefinieerd door:

$$S(t) = f(t) - \sum_{k=0}^n \ell_k(t) \cdot f(x_k) - \pi(t) \cdot K(x). \quad (1.7)$$

Opmerking

$S(t)$ ontstaat uit (1.5) door hierin x te vervangen door t , behalve in $K(x)$. (idem (1.6))

$$S(t) = 0 \quad \text{als } t = x_k \quad k = 0, 1, 2, \dots, n \\ \text{en ook als } t = x,$$

dus $S(t)$ heeft $n+2$ nulpunten.

We veronderstellen nu dat $S(t)$ voldoende vaak differentieerbaar is. Dan volgt, door herhaald toepassen van de stelling van Rolle, uit

$$\begin{array}{ll} S(t) & \text{heeft } n+2 \text{ nulpunten,} \\ S'(t) & \text{heeft } n+1 \text{ nulpunten,} \\ S''(t) & \text{heeft } n \text{ nulpunten,} \\ \vdots & \\ S^{(n+1)}(t) & \text{heeft } 1 \text{ nulpunt.} \end{array}$$

Het nulpunt van $S^{(n+1)}(t)$, genaamd ξ , is gelegen tussen de kleinste en de grootste onder de getallen $x, x_0, x_1, \dots, x_n$. $(n+1)$ keer differentiëren van (1.7) geeft

$$S^{(n+1)}(t) = f^{(n+1)}(t) - (n+1)! K(x).$$

Hierin $t = \xi$ invullen levert ons de uitdrukking voor $K(x)$

$$K(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}$$

en de gezochte uitdrukking voor de restterm $E(x)$ luidt

$$E(x) = \frac{\pi(x) \cdot f^{(n+1)}(\xi)}{(n+1)!}.$$

Door integratie van de interpolatieformule van Lagrange (1.5) vinden we numerieke integratieformules.

$$\begin{aligned} \int_a^b f(x) dx &= \int_a^b \left[\sum_{k=0}^n \ell_k(x) \cdot f(x_k) + E(x) \right] dx \\ &= \sum_{k=0}^n \int_a^b \ell_k(x) dx \cdot f(x_k) + \int_a^b E(x) dx. \end{aligned}$$

Tot nu toe zijn geen speciale eisen gesteld aan de ligging van de punten $x_0, x_1, \dots, x_n$ in (a, b) .

Nemen we nu in het bijzonder aan dat $x_0 = a$, $x_n = b$ en dat de punten aequidistant liggen op (a, b) , dus $x_i - x_{i-1} = h$ dan krijgen we de Newton-Cotes integratieformules.

$$n = 1 \quad \int_{x_0}^{x_1} f(x) dx = \frac{h}{2} (f(x_0) + f(x_1)) - \frac{h^3}{12} f''(\xi) \quad \text{trapeziumregel}$$

$$n = 2 \quad \int_{x_0}^{x_2} f(x) dx = \frac{h}{3} (f(x_0) + 4f(x_1) + f(x_2)) - \frac{h^5}{90} f^{(4)}(\xi)$$

regel van Simpson

$$n = 3 \quad \int_{x_0}^{x_3} f(x) dx = \frac{3h}{8} (f(x_0) + 3f(x_1) + 3f(x_2) + f(x_3)) - \frac{3h^5}{80} f^{(4)}(\xi)$$

drie-achtste regel

Opmerking

In het huidige stadium, waarin men gebruik maakt van de moderne rekenautomaten hebben de hogere orde Newton-Cotes formules geen praktische betekenis. In het algemeen zal men de trapeziumregel of de regel van Simpson gebruiken.

2. De Romberg-Stiefel integratiemethode

Een andere methode om $\int_a^b f(x)dx$ te bepalen is de Romberg-Stiefel integratiemethode. Hierbij veronderstellen we dat $f(x)$ in ieder gewenst punt x te berekenen is. Voor het gemak beperken we ons tot het bepalen van $\int_0^1 f(x)dx$. Dit is geen wezenlijke beperking want door een eenvoudige transformatie is iedere integraal in deze vorm te brengen.

De eerste benadering T_1 krijgen we door toepassing van de trapeziumregel op het hele interval $(0,1)$

$$T_1 = \frac{1}{2} (f(0) + f(1)).$$

We halveren het interval en passen de trapeziumregel toe op beide deelintervallen $(0, \frac{1}{2})$ en $(\frac{1}{2}, 1)$. Zo vinden we als tweede benadering T_2

$$\begin{aligned} T_2 &= \frac{1}{4} (f(0) + 2f(\frac{1}{2}) + f(1)) \\ &= \frac{1}{2} (T_1 + f(\frac{1}{2})). \end{aligned}$$

Nogmaals halveren en de trapeziumregel op de vier deelintervallen toepassen levert ons als derde benadering T_3

$$\begin{aligned} T_3 &= \frac{1}{8} (f(0) + 2f(\frac{1}{4}) + 2f(\frac{1}{2}) + 2f(\frac{3}{4}) + f(1)) \\ &= \frac{1}{2} (T_2 + \frac{f(\frac{1}{4}) + f(\frac{3}{4})}{2}). \end{aligned}$$

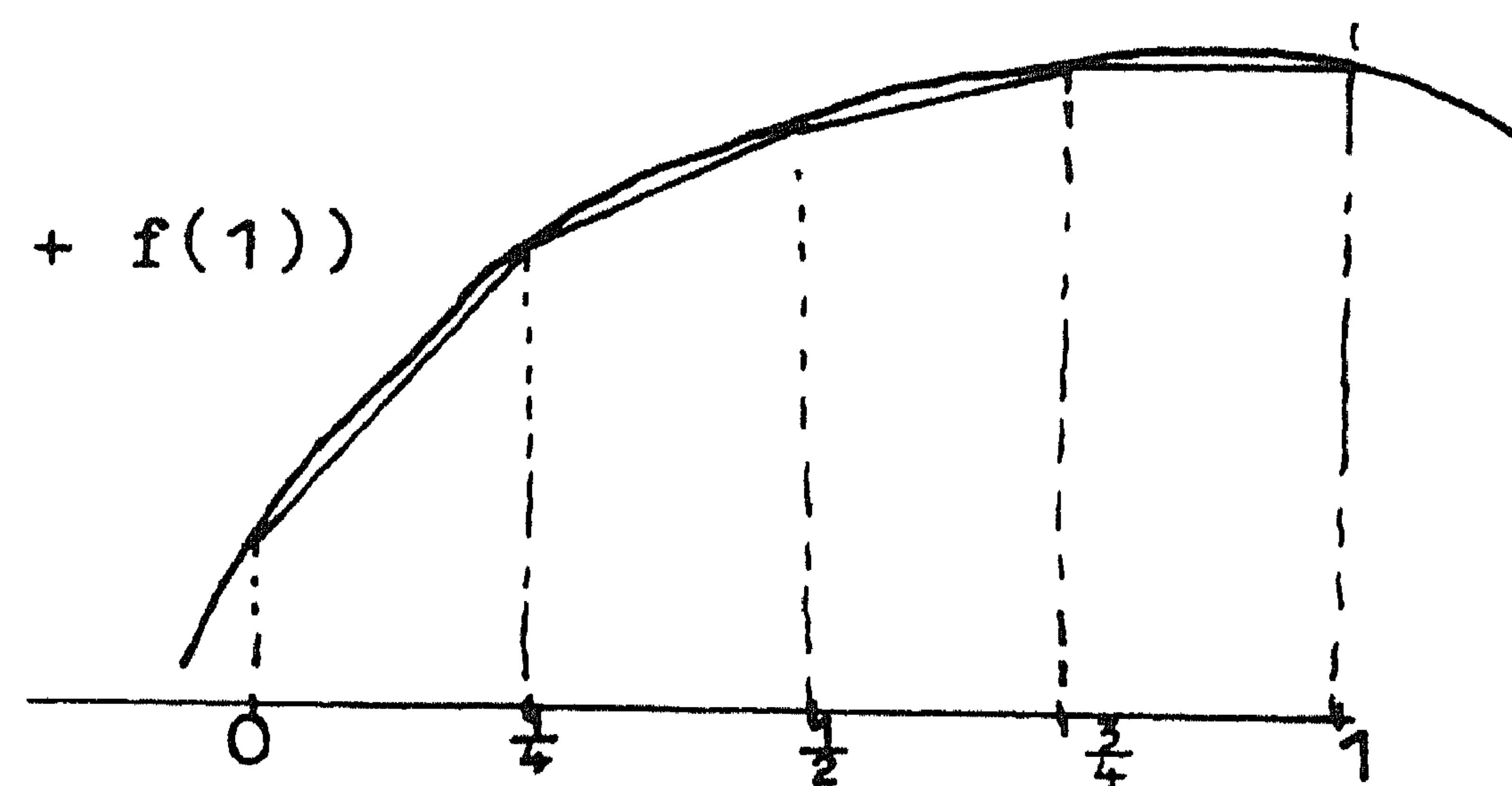
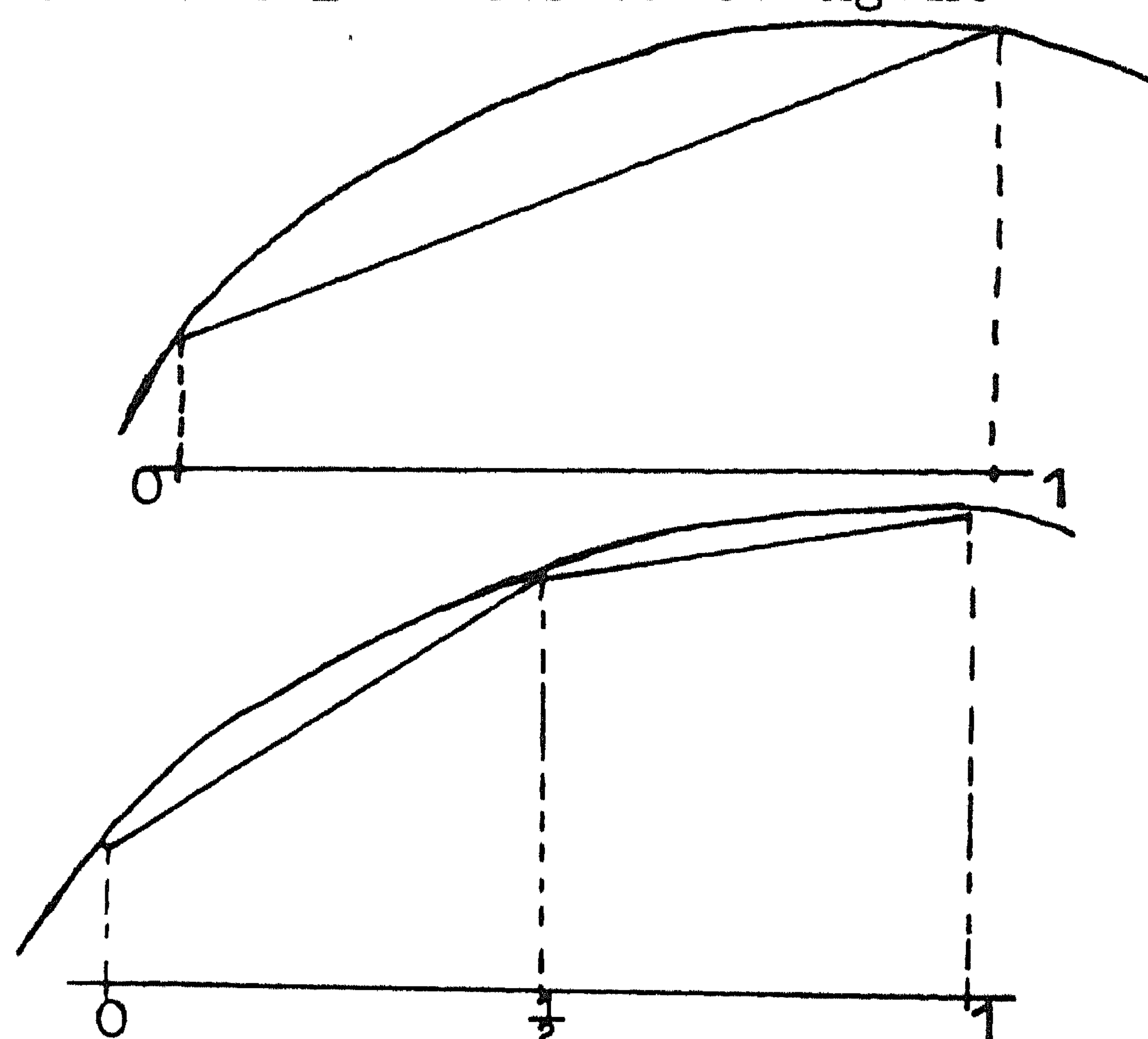
Op dezelfde manier vinden we

$$T_4 = \frac{1}{2} (T_3 + \frac{f(\frac{1}{8}) + f(\frac{3}{8}) + f(\frac{5}{8}) + f(\frac{7}{8})}{4}).$$

Algemeen blijkt

$$T_n = \frac{1}{2} (T_{n-1} + \text{gemiddelde van de nieuwe funktiewaarden}).$$

Duidelijk is dat $\lim_{n \rightarrow \infty} T_n = \int_0^1 f(x)dx$.



De convergentie van de rij $T_1, T_2, T_3, \dots$ zal in het algemeen niet erg snel zijn. Om uit deze rij een nieuwe rij te maken die sneller convergeert redeneren we als volgt.

Bij T_1 maken we een fout $-\frac{1}{12} f''(\xi_1)$ (de zgn "truncation error")

Bij T_2 maken we een fout $-\frac{1}{8} \cdot \frac{1}{12} (f''(\xi_2) + f''(\xi_3)) = -\frac{1}{4} \cdot \frac{1}{12} f''(\xi'_1)$.

Veronderstellen we nu dat f'' niet veel varieert in $(0,1)$ d.w.z. $f''(\xi_1) \approx f''(\xi'_1)$, dan is $\frac{4T_2 - T_1}{3}$ een betere benadering voor de integraal omdat de fout ongeveer geelimineerd is.

We passen dit idee nu direct toe op de volgende manier. T_1 en T_2 zijn benaderingen voor $\int_0^1 f(x)dx$ die exact zijn voor een eerstegraads functie.

We nemen nu het gewogen gemiddelde van T_1 en T_2

$$S_1 = p T_1 + q T_2 \text{ met } p+q = 1 \quad (2.1)$$

en kiezen p en q zo dat S_1 exact de waarde van de integraal levert als $f(x)$ een tweedegraads polynoom is. Het is daarvoor voldoende te zorgen dat de integraal van $f(x) = x^2$ exact geleverd wordt. Voor deze functie geldt

$$\int_0^1 x^2 dx = \frac{1}{3} \quad T_1 = \frac{1}{2} \quad T_2 = \frac{3}{8}$$

zodat we voor p en q de vergelijkingen krijgen

$$\left. \begin{array}{l} p + q = 1 \\ \frac{1}{2}p + \frac{3}{8}q = \frac{1}{3} \end{array} \right\} \implies p = -\frac{1}{3} \quad q = \frac{4}{3}.$$

Invullen van deze p en q in (2.1) geeft

$$S_1 = \frac{4T_2 - T_1}{3}. \quad (2.2)$$

Deze S_1 blijkt ook exact de waarde van de integraal te geven als $f(x)$ een derdegraads polynoom is, hetgeen men gemakkelijk kan nagaan door $f(x) = x^3$ te nemen. S is dus veel beter dan T_1 en T_2 .

Opmerking

Daar S_1 de integraal is van de kwadratische functie door de drie basispunten, moet (2.2) overeenstemmen met de regel van Simpson

$$S_1 = \frac{1}{6} (f(0) + 4f(\frac{1}{2}) + f(1))$$

wat men kan verifiëren door (2.2) uit te schrijven.

Met T_2 en T_3 maken we op dezelfde wijze $S_2 = \frac{4T_3 - T_2}{3}$ die de waarde van de integraal geeft als we $f(x)$ door twee paraboolbogen vervangen.

$$S_3 = \frac{4T_4 - T_3}{3} \quad S_4 = \frac{4T_5 - T_4}{3} \quad \text{enz.}$$

We krijgen zo een rij $S_1, S_2, S_3, \dots$ die sneller convergeert naar de integraal dan de T -rij.

Nu passen we op de rij $S_1, S_2, S_3, \dots$ ook het procédé van het gewogen gemiddelde toe. De combinatie

$$C_1 = p S_1 + q S_2 \quad \text{met } p+q = 1 \quad (2.3)$$

is exact de integraal voor een derdegraads polynoom. We kiezen p en q zodanig, dat C_1 ook een vierdegraads polynoom exact integreert. Met $f(x) = x^4$ vinden we

$$\int_0^1 x^4 dx = \frac{1}{5} \quad T_1 = \frac{1}{2} \quad T_2 = \frac{9}{32} \quad T_3 = \frac{113}{512}.$$

Hieruit berekenen we $S_1 = \frac{5}{24}$ $S_2 = \frac{77}{384}$ en vinden dan de vergelijkingen voor p en q

$$\left. \begin{array}{l} \frac{5}{24}p + \frac{77}{384}q = \frac{1}{5} \\ p + q = 1 \end{array} \right\} \Rightarrow p = -\frac{1}{15} \quad q = \frac{16}{15}.$$

De gevonden benadering

$$C_1 = \frac{4^2 S_2 - S_1}{4^2 - 1} \quad (2.4)$$

blijkt eveneens beter te zijn dan we verwachten, nl. ook voor een vijfdegraads polynoom wordt de integraal door C_1 exact gegeven.

Analoog maken we

$$C_2 = \frac{4^2 S_3 - S_2}{4^2 - 1} \quad C_3 = \frac{4^2 S_4 - S_3}{4^2 - 1} \quad \text{enz.}$$

Uit de rij $C_1, C_2, C_3, \dots$ vormen we volgens hetzelfde procédé een nieuwe rij, waarvoor geldt

$$D_1 = \frac{4^3 C_2 - C_1}{4^3 - 1} \quad D_2 = \frac{4^3 C_3 - C_2}{4^3 - 1} \quad \text{enz.}$$

Zoals we verwachten blijkt D_1 (en dus ook D_2 enz.) ook voor een zevengraads polynoom de integraal exact te geven. Uit de D -rij kunnen we weer een rij maken

$$E_1 = \frac{4^4 D_2 - D_1}{4^4 - 1}$$

die ook een negendegraads polynoom exact integreert, enz.

De achtereenvolgende benaderingen schrijft men het meest overzichtelijk op in de volgende integratietabel

T_1				
T_2	S_1			
T_3	S_2	C_1		
T_4	S_3	C_2	D_1	
T_5	S_4	C_3	D_2	E_1

Het berekenen van een nieuwe T waarde in de eerste kolom impliceert het uitrekenen van nieuwe funktiewaarden. De overige kolommen vult men dan aan door de genoemde lineaire combinaties. Het is daarbij numeriek beter in plaats van met (2.2), (2.4) enz. de elementen te berekenen volgens

$$S_1 = T_2 + \frac{T_2 - T_1}{3}$$

$$C_1 = S_2 + \frac{S_2 - S_1}{15} \quad \text{enz.}$$

zodat men steeds de aanvullende correctie uitrekent.

Daar S_4 een betere benadering is dan T_5 , verkregen uit dezelfde funktiewaarden, C_3 weer beter is dan S_4 enz., zal men de tabel rijsgewijs invullen. De nauwkeurigheid van het resultaat wordt dan gegeven door de overeenstemming van twee elementen uit dezelfde rij. Als b.v. D_2 en E_1 in 5 decimalen overeenstemmen dan nemen we aan dat E_1 de integraal in 5 decimalen nauwkeurig geeft. Op deze wijze berekent men niet meer funktiewaarden dan strikt noodzakelijk is, wat speciaal bij functies, waarvan het bepalen van de funktiewaarde tijdrovend is, een groot voordeel is.

Opmerking

We hebben gezien dat de formule voor S_1 niets anders is dan de regel van Simpson (opmerking blz. NA-3). Op dezelfde manier kan men inzien dat (2.4) en de Newton-Cotes formule voor $n = 4$ dezelfde zijn.

Bij D_1 is dit echter anders. Om D_1 te kunnen berekenen zijn 9 funktiewaarden nodig. Maar D_1 is exact de integraal voor polynomen van de graad ≤ 7 . In het algemeen dus niet voor een achtstegraads polynoom. De N-C formule voor 9 punten geeft in het geval van een achtstegraads polynoom wel een exact antwoord. D_1 heeft dus niets meer te maken met de integratie van een interpolatiepolynoom. Het blijkt echter, dat toch de R-S methode de voorkeur verdient boven de N-C formule. Schrijven we nl. zowel de formule voor D_1 als de N-C formule als een gewogen combinatie van de 9 funktiewaarden, dan blijken de gewichten bij D_1 positief te zijn terwijl de N-C formule negatieve gewichten heeft!

Dit is ten eerste onnatuurlijk als men bedenkt dat de integraal per definitie de limiet van een som van funktiewaarden is. Ten tweede is het numeriek erg onaantrekkelijk daar het een extra verlies van cijfers veroorzaakt.

Voorbeeld Berekening van $\pi = 4 \int_0^1 \frac{dx}{1+x^2}$

funktietabel				gemiddelde	T
4	2			3	3
3.2				3.2	3.1
3.76471	2.56			3.16236	3.13118
3.93846	3.50685	2.87640	2.26549	3.14680	3.13899
3.98444	3.86415	3.64413	3.35738	3.14290	3.14094
3.03858	2.71618	2.40941	2.12890		

Integratietabel

T	S	C	D	E
3				
3.1	3.13333			
3.13118	3.14157	3.14212		
3.13899	3.14159	3.14159	3.14158	
3.14094	3.14159	3.14159	3.14159	3.14159

$$\pi = 3.141592\dots$$

3. Gauss integratie

In de tijd dat men alle berekeningen met de hand moest doen, werden bij voorkeur integratieformules gebruikt, gebaseerd op constante intervallen, vooral omdat de integrand meestal berekend werd met behulp van funktietabellen. Nu, door het gebruik van elektronische rekenmachines, we in het algemeen mogen aannemen dat de berekening van $f(x)$ in ieder willekeurig punt x gelijkwaardig is, kunnen we ook de beperking tot constante intervallengten laten vallen. De winst die we daarmee halen is, dat we minder funktiewaarden behoeven te berekenen, zoals blijkt uit het volgende

Voorbeeld

5 punts Newton-Cotes formule:

$$\int_{-1}^1 f(x)dx = \frac{1}{45} (7f(-1)+32f(-\frac{1}{2})+12f(0)+32f(\frac{1}{2})+7f(1)) - \frac{f^{(6)}(\xi)}{15120}.$$

3 punts Gauss-Legendre formule:

$$\int_{-1}^1 f(x)dx = \frac{1}{9} (5f(-\sqrt{\frac{3}{5}})+8f(0)+5f(\sqrt{\frac{3}{5}})) - \frac{f^{(6)}(\xi)}{15750}$$

Zoals de Lagrange interpolatieformule in feite de grondslag is voor de Newton-Cotes integratieformules, zo kunnen de Gauss integratieformules worden afgeleid uit de Hermite interpolatieformule.

Van een functie $f(x)$ gedefinieerd op (a,b) zijn in de punten $x_1, \dots, x_m$ zowel de funktiewaarden $f(x_1), \dots, f(x_m)$ als de afgeleide $f'(x_1), \dots, f'(x_m)$ bekend. We zoeken nu een polynoom $p(x)$ waarvoor geldt

$$\begin{aligned} p(x_i) &= f(x_i) \\ p'(x_i) &= f'(x_i) \end{aligned} \quad i = 1, 2, \dots, m.$$

De graad van $p(x)$ moet dan $2m-1$ zijn.

We pakken dit probleem op dezelfde wijze aan als bij Lagrange.

We stellen

$$p(x) = \sum_{k=1}^m g_k(x) \cdot f(x_k) + \sum_{k=1}^m h_k(x) \cdot f'(x_k) \quad (3.1)$$

en eisen dat $g_k(x)$ en $h_k(x)$ polynomen zijn van de graad $2m-1$, die voldoen aan de voorwaarden

$$\begin{aligned} g_i(x_j) &= \delta_{ij} & h_i(x_j) &= 0 \\ \text{en} & & & \\ g_i'(x_j) &= 0 & h_i'(x_j) &= \delta_{ij} \end{aligned} \quad (3.2)$$

Nu is $\ell_i(x)$ (zie (1.3)) een polynoom van de graad $m-1$ waarvoor geldt $\ell_i(x_j) = \delta_{ij}$. $\{\ell_i(x)\}^2$ heeft de graad $2m-2$ en voldoet aan $\{\ell_i(x_j)\}^2 = \delta_{ij}$, terwijl voor de afgeleide $(\{\ell_i(x)\}^2)' = 2 \cdot \ell_i(x) \cdot \ell_i'(x)$ geldt $(\{\ell_i(x_j)\}^2)' = 0$ als $i \neq j$. We nemen daarom

$$\begin{aligned} g_i(x) &= \lambda_i(x) \cdot \{\ell_i(x)\}^2 \\ h_i(x) &= \mu_i(x) \cdot \{\ell_i(x)\}^2 \end{aligned} \quad (3.3)$$

waarin $\lambda_i(x)$ en $\mu_i(x)$ lineaire funkties zijn die zo gekozen worden dat $g_i(x)$ en $h_i(x)$ voldoen aan (3.2).

Nu geldt

$$\begin{aligned} g_i(x_j) &= \lambda_i(x_j) \cdot \delta_{ij} & h_i(x_j) &= \mu_i(x_j) \cdot \delta_{ij} \\ g_i'(x_j) &= \lambda_i'(x_j) \cdot \delta_{ij} + 2\lambda_i(x_j) \cdot \ell_i'(x_j) \cdot \delta_{ij} \\ h_i'(x_j) &= \mu_i'(x_j) \cdot \delta_{ij} + 2\mu_i(x_j) \cdot \ell_i'(x_j) \cdot \delta_{ij} \end{aligned}$$

Hieruit halen we voor $\lambda_i(x)$ en $\mu_i(x)$ de volgende voorwaarden

$$\begin{aligned}\lambda_i(x_i) &= 1 & \mu_i(x_i) &= 0 \\ \lambda_i'(x_i) + 2\lambda_i(x_i) \cdot \ell_i'(x_i) &= 0 \\ \mu_i'(x_i) &= 1\end{aligned}$$

en we vinden tenslotte

$$\begin{aligned}\lambda_i(x) &= 1 - 2\ell_i'(x_i) \cdot (x - x_i) \\ \mu_i(x) &= x - x_i\end{aligned}\tag{3.4}$$

Als we (3.4) invullen in (3.3) en deze vervolgens in (3.1) dan krijgen we het interpolatie polynoom van Hermite.

$$p(x) = \sum_{k=1}^m [(1 - 2\ell_k'(x_k) \cdot (x - x_k)) \cdot \{\ell_k(x)\}^2 \cdot f(x_k) + (x - x_k) \cdot \{\ell_k(x)\}^2 \cdot f'(x_k)].\tag{3.5}$$

De bepaling van de restterm $E(x) = f(x) - p(x)$ gaat ook op dezelfde manier als bij Lagrange. Omdat x_k voor $k = 1, 2, \dots, m$ een tweevoudig nulpunt is, stellen we

$$E(x) = \{\pi(x)\}^2 \cdot K(x)$$

en beschouwen vervolgens

$$F(t) = f(t) - p(t) - \{\pi(t)\}^2 \cdot K(x).\tag{3.6}$$

Dan is $F(t) = 0$ voor $t = x_k$ $k = 1, 2, \dots, m$ en ook voor $t = x$, dus $F(t)$ heeft $m+1$ nulpunten. Volgens Rolle heeft $F'(t)$ dan m nulpunten, gelegen tussen de nulpunten van $F(t)$. Daar ook $F'(t) = 0$ voor $t = x_k$ $k = 1, 2, \dots, m$ heeft $F'(t)$ dus $2m$ nulpunten.

Hieruit volgt, op dezelfde manier als bij $S(t)$ (zie blz. NA-2), dat $F^{(2m)}(t)$ één nulpunt heeft, dat we weer ξ noemen.

We differentieren (3.6) $2m$ keer en vullen vervolgens $t = \xi$ in. Dan vinden we

$$K(x) = \frac{f^{(2m)}(\xi)}{(2m)!}.$$

Voor de restterm $E(x)$ krijgen we tenslotte de uitdrukking

$$E(x) = \frac{\{\pi(x)\}^2 \cdot f^{(2m)}(\xi)}{(2m)!}.\tag{3.7}$$

Als dus $f(x)$ een polynoom is waarvan de graad $\leq 2m-1$, dan is het m -punts interpolatiepolynoom van Hermite exact gelijk aan de functie.

Uit de Hermite interpolatieformule

$$f(x) = \sum_{k=1}^m g_k(x) \cdot f(x_k) + \sum_{k=1}^m h_k(x) \cdot f'(x_k) + E(x) \quad (3.8)$$

leiden we af de integratieformule

$$\begin{aligned} \int_a^b w(x) \cdot f(x) dx &= \sum_{k=1}^m \int_a^b w(x) \cdot g_k(x) dx \cdot f(x_k) + \\ &+ \sum_{k=1}^m \int_a^b w(x) \cdot h_k(x) dx \cdot f'(x_k) + \int_a^b w(x) \cdot E(x) dx. \end{aligned}$$

Met de afkortingen

$$G_k = \int_a^b w(x) \cdot g_k(x) dx = \int_a^b w(x) \cdot \{1 - 2\ell'_k(x_k)(x-x_k)\} \cdot \{\ell_k(x)\}^2 dx \quad (3.9)$$

$$H_k = \int_a^b w(x) \cdot h_k(x) dx = \int_a^b w(x) \cdot (x-x_k) \cdot \{\ell_k(x)\}^2 dx$$

wordt de Hermite integratieformule

$$\int_a^b w(x) \cdot f(x) dx = \sum_{k=1}^m G_k \cdot f(x_k) + \sum_{k=1}^m H_k \cdot f'(x_k) + E. \quad (3.10)$$

De gewichtsfunctie $w(x)$ zal bij een bepaald type integratieformule een vaste functie zijn, die mede de gewichtscoefficienten G_k en H_k bepaalt. We beperken ons tot gewichtsfuncties $w(x)$ die niet-negatief zijn in (a,b)

$$w(x) \geq 0 \quad a \leq x \leq b \quad (3.11)$$

De fout E uit de integratieformule (3.10) is

$$E = \int_a^b w(x) \cdot E(x) dx = \frac{1}{(2m)!} \int_a^b w(x) \cdot f^{(2m)}(\xi) \cdot \{\pi(x)\}^2 dx.$$

Deze brengen we in een eenvoudiger vorm door toepassing van de Eerste middelwaarde stelling van de integraalrekening.

Zij $f(x)$ een integreerbare functie waarvoor geldt $f(x) \geq 0$ voor $a \leq x \leq b$, zij $g(x)$ een continue functie op (a,b) , dan is er een η te vinden op (a,b) zo dat

$$\int_a^b f(x) \cdot g(x) dx = g(\eta) \int_a^b f(x) dx.$$

Bewijs. Uit $f(x) \geq 0$ volgt

$$\min_{a \leq x \leq b} g(x) \cdot \int_a^b f(x) dx \leq \int_a^b g(x) \cdot f(x) dx \leq \max_{a \leq x \leq b} g(x) \int_a^b f(x) dx$$

en uit de continuïteit van $g(x)$ volgt dan de stelling.

Daar $w(x)$ niet-negatief is verondersteld, is $w(x) \cdot \{\pi(x)\}^2 \geq 0$. Toepassing van de eerste middelwaarde stelling geeft dan

$$E = \frac{f^{(2m)}(\eta)}{(2m)!} \int_a^b w(x) \cdot \{\pi(x)\}^2 dx \quad (3.12)$$

voor zekere η in (a, b) .

Opmerking

We vergelijken dit resultaat met het corresponderende geval van de Lagrange integratieformule, gebaseerd op m punten.

$$\int_a^b w(x) \cdot f(x) dx = \sum_{k=1}^m W_k \cdot f(x_k) + E. \quad (3.13)$$

Hierin is

$$W_k = \int_a^b w(x) \cdot \ell_k(x) dx \quad (3.14)$$

en

$$E = \frac{1}{m!} \int_a^b w(x) \cdot f^{(m)}(\xi) \cdot \pi(x) dx. \quad (3.15)$$

Het afleiden van een soortgelijke vorm voor E als (3.12) is in dit geval veel moeilijker. Daar $\pi(x)$ niet tekenvast is op (a, b) kunnen we namelijk niet de middelwaarde stelling toepassen.

Als een integratieformule exacte resultaten oplevert wanneer $f(x)$ een willekeurig polynoom is van de graad $\leq r$, doch niet meer exact is voor minstens één polynoom van de graad $r+1$, dan zeggen we dat de integratieformule een nauwkeurigheidsgraad r heeft.

Anders gezegd.

Een integratieformule heeft de nauwkeurigheidsgraad r als ze exacte resultaten geeft voor $1, x, x^2, \dots, x^r$, maar niet voor x^{r+1} .

Uit (3.12) volgt dat de nauwkeurigheidsgraad van de m-punts Hermite integratieformule precies $2m-1$ is.

Voor de m-punts Lagrange integratieformule volgt uit (3.15) dat de nauwkeurigheidsgraad minstens $m-1$ is.

We bewijzen nu dat de nauwkeurigheidsgraad bij Lagrange hoogstens $2m-1$ is.

Nemen we $f(x) = \{\pi(x)\}^2$ dan wordt (3.13)

$$\int_a^b w(x) \cdot \{\pi(x)\}^2 dx = \sum_{k=1}^m W_k \cdot \{\pi(x_k)\}^2 + E.$$

Daar $\pi(x_k) = 0$ en $\int_a^b w(x) \cdot \{\pi(x)\}^2 dx \neq 0$ moet $E \neq 0$. Dit betekent dat voor het polynoom $\{\pi(x)\}^2$, dat van de graad $2m$ is, de integratieformule niet exact is. Dus is de nauwkeurigheidsgraad $< 2m$. De nauwkeurigheidsgraad van de Lagrange formule hangt af van de keuze van de basispunten $x_1, x_2, \dots, x_m$. We zullen zien dat het mogelijk is $x_1, x_2, \dots, x_m$ zo te kiezen dat de nauwkeurigheidsgraad $2m-1$ is.

Met behulp van $\ell_k(x) = \frac{\pi(x)}{(x-x_k) \cdot \pi'(x_k)}$ kunnen we de formule voor H_k uit (3.9) omvormen tot

$$H_k = \frac{1}{\pi'(x_k)} \int_a^b w(x) \cdot \pi(x) \cdot \ell_k(x) dx. \quad (3.16)$$

Als we de basispunten $x_1, x_2, \dots, x_m$ zo kunnen kiezen dat

$$H_k = 0 \quad k = 1, 2, \dots, m \quad (3.17)$$

dan gaat de Hermite integratieformule over in een formule van het type (3.13) waarbij de nauwkeurigheidsgraad $2m-1$ behouden blijft.

Intermezzo

In de lineaire ruimte $C(a,b)$ van alle continue functies op het gesloten interval (a,b) kunnen we bij een tweetal functies $f(x)$ en $g(x)$ het inwendig produkt (f,g) als volgt definieren

$$(f,g) = \int_a^b f(x) \cdot g(x) dx. \quad (3.18)$$

Dit kan men beschouwen als een generalisatie van het inwendig produkt $(\underline{a}, \underline{b})$ van twee vectoren $\underline{a} = (a_1, a_2, \dots, a_n)$ en $\underline{b} = (b_1, b_2, \dots, b_n)$ uit R_n , dat als volgt gedefinieerd is

$$(\underline{a}, \underline{b}) = \sum_{i=1}^n a_i \cdot b_i.$$

We zeggen dat twee functies $f(x)$ en $g(x)$ loodrecht op elkaar staan (orthogonaal zijn) als $(f, g) = 0$, dus als $\int_a^b f(x) \cdot g(x) dx = 0$.

We krijgen een algemener inwendig produkt door een gewichtsfunctie $w(x)$ te introduceren,

$$(f, g) = \int_a^b w(x) \cdot f(x) \cdot g(x) dx \quad (3.19)$$

en we zeggen:

De functies $f(x)$ en $g(x)$ staan loodrecht op elkaar over het interval (a, b) met betrekking tot de gewichtsfunctie $w(x)$ als

$$\int_a^b w(x) \cdot f(x) \cdot g(x) dx = 0.$$

We kunnen nu (3.17) ook zo uitdrukken:

$\pi(x)$ staat loodrecht op de polynomen $\ell_k(x)$ $k=1, 2, \dots, m$ over het interval (a, b) met betrekking tot de gewichtsfunctie $w(x)$.

Omdat $\ell_k(x)$ de graad $m-1$ heeft, is een voldoende voorwaarde voor (3.17) dat $\pi(x)$ loodrecht staat op alle polynomen van de graad $\leq m-1$ over (a, b) met betrekking tot $w(x)$.

We bewijzen nu dat deze voorwaarde ook noodzakelijk is.

Veronderstel

$$1) H_k = 0 \quad k = 1, 2, \dots, m$$

$$2) \text{ nauwkeurigheidsgraad van de integratieformule is } 2m-1.$$

Zij $f(x)$ een polynoom van de graad $\leq 2m-1$, dat geschreven kan worden als

$$f(x) = \pi(x) \cdot p_{m-1}(x),$$

waarin $p_{m-1}(x)$ een willekeurig polynoom van de graad $\leq m-1$ is.

Omdat $\pi(x_k) = 0$, is ook $f(x_k) = 0$ $k = 1, 2, \dots, m$.

We vullen deze $f(x)$ in (3.10) in en vinden dan

$$\int_a^b w(x) \cdot f(x) dx = \int_a^b w(x) \cdot \pi(x) \cdot p_{m-1}(x) dx = 0$$

omdat $H_k = 0$, $E = 0$ en $f(x_k) = 0$; waaruit de noodzakelijkheid van de genoemde voorwaarde volgt.

We hebben dus de volgende stelling bewezen.

Stelling

Dan en alleen dan, indien het polynoom $\pi(x)$, van de graad m , loodrecht staat op alle polynomen van de graad $\leq m-1$ over (a,b) met betrekking tot $w(x)$, reduceert de Hermite integratieformule (3.10) tot de formule

$$\int_a^b w(x) \cdot f(x) dx = \sum_{k=1}^m G_k \cdot f(x_k) + E \quad (3.20)$$

waarin

$$\begin{aligned} G_k &= \int_a^b w(x) \{\ell_k(x)\}^2 dx - 2\ell_k'(x_k) \cdot H_k \\ &= \int_a^b w(x) \cdot \{\ell_k(x)\}^2 dx \quad (\text{zie (3.9)}) \end{aligned} \quad (3.21)$$

en

$$E = \frac{f^{(2m)}(\xi)}{(2m)!} \int_a^b w(x) \cdot \{\pi(x)\}^2 dx,$$

en de basispunten $x_1, x_2, \dots, x_m$ de nulpunten van $\pi(x)$ zijn.

Formule (3.20) wordt de Gauss integratieformule genoemd.

We vergelijken (3.20) met (3.13). Beide formules zijn exact, d.w.z. $E = 0$, voor iedere polynoom van de graad $\leq m-1$. Dan moeten de gewichtsfuncties G_k en W_k dezelfde zijn, en dus geldt de betrekking

$$\int_a^b w(x) \cdot \{\ell_k(x)\}^2 dx = \int_a^b w(x) \cdot \ell_k(x) dx \quad (3.22)$$

onder de voorwaarde $H_k = 0$.

Opmerking

Deze betrekking zullen we gebruiken als we de waarde van G_k willen berekenen. Het bepalen van W_k zal in het algemeen namelijk eenvoudiger zijn.

Opmerking

Uit (3.21) zien we dat alle gewichtscoefficienten G_k van de Gauss integratieformule positief zijn.

We geven enkele speciale gevallen van de Gauss integratieformule.

Gauss-Legendre

Het interval (a,b) is eindig. Zonder beperking mogen we daarvoor $(-1,1)$ nemen. De gewichtsfunctie nemen we constant, $w(x) \equiv 1$. Dan geldt voor $f(x)$ gedefinieerd op $(-1,1)$

$$\int_{-1}^1 f(x) dx = \sum_{k=1}^m G_k \cdot f(x_k) + E$$

als we $x_1, x_2, \dots, x_m$ zo kiezen, dat $\pi(x)$ loodrecht staat op alle polynomen van de graad $\leq m-1$ over $(-1,1)$.

Er bestaan polynomen die aan deze eis voldoen. Deze staan bekend als de polynomen van Legendre.

De differentiaalvergelijking

$$(1-x^2) \frac{d^2 y}{dx^2} - 2x \frac{dy}{dx} + n(n+1)y = 0 \quad (3.23)$$

heet de differentiaalvergelijking van Legendre.

Als n geheel en ≥ 0 , dan heeft de differentiaalvergelijking een n -de graads polynoom $P_n(x)$ als oplossing.

Dit kunnen we op de volgende manier inzien.

Stel $u(x) = (x^2-1)^n$ dan is $u'(x) = 2n x (x^2-1)^{n-1}$ dus $(x^2-1)u' - 2n x u = 0$.

Deze formule differentieren we $m+1$ keer, met de regel van Leibniz

$$(x^2-1)u^{(m+2)} + \binom{m+1}{1} \cdot 2x \cdot u^{(m+1)} + \binom{m+1}{2} \cdot 2 \cdot u^{(m)} - 2nxu^{(m+1)} - 2n\binom{m+1}{1}u^{(m)} = 0$$

$$(x^2-1)u^{(m+2)} + 2(m+1-n)xu^{(m+1)} + (m+1)(m-2n)u^{(m)} = 0.$$

We stellen $m = n$ en krijgen dan

$$(x^2 - 1)u^{(n+2)} + 2xu^{(n+1)} - n(n+1)u^{(n)} = 0$$

waaruit blijkt dat $u^{(n)}(x)$ oplossing is van (3.23).
Dan is ook

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n \quad (\text{Rodrigues}) \quad (3.24)$$

een oplossing van (3.23).

Het polynoom $P_n(x)$ wordt genoemd het n-de Legendre polynoom.

Het is een polynoom van de graad n ; de coëfficiënt van x^n is $\frac{(2n)!}{2^n (n!)^2}$; de factor $\frac{1}{2^n \cdot n!}$ in (3.24) is zo gekozen dat $P_n(1) = 1$.

Enkele voorbeelden

$$P_0(x) = 1$$

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

$$P_1(x) = x$$

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

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

$$P_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x)$$

$P_n(x)$ kunnen we berekenen met de volgende recursieformule

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

$P_n(x)$ heeft de eigenschap

$$\int_{-1}^1 P_n(x) \cdot x^s dx = 0 \quad s < n \quad (3.25)$$

en als $s = n$ geldt

$$\int_{-1}^1 P_n(x) \cdot x^n dx = \frac{(-1)^n}{2^n} \int_{-1}^1 (x^2 - 1)^n dx = \frac{2^{n+1} \cdot (n!)^2}{(2n+1)!} \quad (3.26)$$

Men kan dit bewijzen door middel van partiële integratie.

Uit (3.25) volgt dat $P_n(x)$ loodrecht staat op alle polynomen van de graad $< n$ over $(-1, 1)$.

In het bijzonder geldt de relatie

$$\int_{-1}^1 P_m(x) \cdot P_n(x) dx = 0 \quad m \neq n. \quad (3.27)$$

We berekenen nog $\int_{-1}^1 \{P_n(x)\}^2 dx$. Vanwege (3.25) behoeven we van één $P_n(x)$ in de integrand alleen de kopterm $\frac{(2n)!}{2^n(n!)^2} x^n$ te nemen. Samen met (3.26) vinden we dan

$$\int_{-1}^1 \{P_n(x)\}^2 dx = \frac{(2n)!}{2^n(n!)^2} \int_{-1}^1 P_n(x) \cdot x^n dx = \frac{2}{2n+1}. \quad (3.28)$$

(3.27) en (3.28) noemen we de orthogonaliteitsrelaties. Op grond van deze relaties noemen we de polynomen $P_n(x)$ een stel orthogonale polynomen.

Omdat het polynoom $P_m(x)$ loodrecht staat op alle polynomen van de graad $\leq m-1$ kunnen we nemen (zie blz NA-16)

$$\pi(x) = \frac{2^m \cdot (m!)^2}{(2m)!} P_m(x).$$

De punten $x_1, x_2, \dots, x_m$ zijn dan de nulpunten van $P_m(x)$.

Dit is echter alleen zinvol als al deze nulpunten verschillend zijn en tussen -1 en $+1$ liggen.

We tonen dit aan met de stelling van Rolle.

$(x^2-1)^m$ heeft m -voudig nulpunt in -1 en $+1$

$\frac{d}{dx} (x^2-1)^m$ heeft $(m-1)$ -voudig nulpunt in -1 en $+1$
en ook een nulpunt tussen -1 en $+1$

$\frac{d^2}{dx^2} (x^2-1)^m$ heeft $(m-2)$ -voudig nulpunt in -1 en $+1$
en twee verschillende nulpunten tussen -1 en $+1$

enz.

$\frac{d^m}{dx^m} (x^2-1)^m$ heeft geen nulpunt in -1 en $+1$
en m verschillende nulpunten tussen -1 en $+1$

Dus ook $P_m(x)$ heeft m verschillende nulpunten, en alle nulpunten liggen tussen -1 en $+1$.

Geval $m = 3$

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

$$\text{nulpunten zijn } x_1 = -\sqrt{\frac{3}{5}} \quad x_2 = 0 \quad x_3 = \sqrt{\frac{3}{5}}$$

$$G_k = \int_{-1}^1 \{\ell_k(x)\}^2 dx = \int_{-1}^1 \ell_k(x) dx \quad (\text{zie (3.22)}) \text{ geeft}$$

$$G_1 = \frac{5}{9} \quad G_2 = \frac{8}{9} \quad G_3 = \frac{5}{9}.$$

De 3-punts Gauss-Legendre integratieformule is

$$\int_{-1}^1 f(x) dx = \frac{1}{9} [5f(-\sqrt{\frac{3}{5}}) + 8f(0) + 5f(\sqrt{\frac{3}{5}})] + E.$$

E berekenen we met (3.12)

$$E = \frac{f^{(6)}(\xi)}{6!} \int_{-1}^1 \{\pi(x)\}^2 dx = \frac{f^{(6)}(\xi)}{6!} \cdot \left(\frac{2}{5}\right)^2 \int_{-1}^1 \{P_3(x)\}^2 dx = \frac{f^{(6)}(\xi)}{15750}.$$

Opmerking

Deze uitdrukking is numeriek onbruikbaar om de fout in de integratieformule te schatten. Om een indruk te krijgen van de fout gaan we analoog te werk als bij Newton-Cotes. We halveren het interval en passen op beide helft Gauss-Legendre toe. De fout wordt dan in het algemeen een factor 2^6 kleiner.

Na de uitvoerige afleiding van Gauss-Legendre kunnen we bij de andere formules volstaan met in het kort aangeven van de afleiding, daar die volkomen analoog verloopt als bij Gauss-Legendre.

Gauss-Laguerre

Het interval is half oneindig, we nemen $(0, \infty)$

$$w(x) = e^{-x}.$$

Voor $\pi(x)$ moeten we weer een stel orthogonale polynomen hebben, nu over $(0, \infty)$ met betrekking tot e^{-x} .

Dit zijn de polynomen van Laguerre.

$$L_n(x) = e^x \frac{d^n}{dx^n} (x^n e^{-x})$$

Deze voldoen aan de differentiaalvergelijking van Laguerre.

$$xy'' + (1-x)y' + ny = 0.$$

Voorbeelden

$$L_0(x) = 1$$

$$L_2(x) = 2 - 4x + x^2$$

$$L_1(x) = 1-x$$

$$L_3(x) = 6 - 18x + 9x^2 - x^3.$$

Recursieformule

$$L_{n+1}(x) = (1 + 2n - x) L_n(x) - n^2 \cdot L_{n-1}(x).$$

Orthogonaliteitsrelaties

$$\int_0^\infty e^{-x} L_m(x) \cdot L_n(x) dx = 0 \quad m \neq n$$

$$\int_0^\infty e^{-x} \{L_n(x)\}^2 dx = (n!)^2.$$

We nemen

$$\pi(x) = (-1)^m L_m(x)$$

waarvan alle nulpunten positief zijn.

Geval m = 2

2-punts Gauss-Laguerre formule

$$\int_0^\infty e^{-x} f(x) dx = \frac{1}{4} [(2 + \sqrt{2}) f(2 - \sqrt{2}) + (2 - \sqrt{2}) f(2 + \sqrt{2})] + \frac{f^{(4)}(\xi)}{6}.$$

Gauss-Hermite

Het interval $(-\infty, \infty)$, $w(x) = e^{-x^2}$.

De differentiaalvergelijking van Hermite

$$y'' - 2xy' + 2ny = 0$$

heeft als oplossingen de polynomen van Hermite

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} (e^{-x^2}).$$

Deze polynomen zijn orthogonaal over $(-\infty, \infty)$ met betrekking tot e^{-x^2}

Voorbeelden

$$H_0(x) = 1$$

$$H_2(x) = 4x^2 - 2$$

$$H_1(x) = 2x$$

$$H_3(x) = 8x^3 - 12x.$$

Recursieformule voor $H_n(x)$

$$H_{n+1}(x) = 2x H_n(x) - 2n H_{n-1}(x).$$

Orthogonaliteitsrelaties

$$\int_{-\infty}^{\infty} e^{-x^2} H_n(x) \cdot H_m(x) dx = \begin{cases} 0 & m \neq n \\ 2^n n! \sqrt{\pi} & m = n \end{cases}$$

We nemen in dit geval

$$\pi(x) = \frac{1}{2^m} H_m(x).$$

Geval $m = 3$

3-punts Gauss-Hermite formule

$$\int_{-\infty}^{\infty} e^{-x^2} f(x) dx = \frac{\sqrt{\pi}}{6} \left[f\left(-\sqrt{\frac{3}{2}}\right) + 4f(0) + f\left(\sqrt{\frac{3}{2}}\right) \right] + \frac{\sqrt{\pi} f^{(6)}(\xi)}{960}$$

Opmerking

Zowel van de gewichtscoefficiënten, als van de basispunten van de verschillende Gauss integratieformules bestaan uitgebreide tabellen.

Gauss-Chebyshev

Het interval is $(-1, 1)$, maar nu met de gewichtsfunctie $w(x) = \frac{1}{\sqrt{1-x^2}}$
 Dan is

$$\pi(x) = \frac{1}{2^{m-1}} T_m(x)$$

waarin $T_m(x)$ het m-de Chebyshev polynoom is gedefinieerd door

$$T_m(x) = \cos(m \arccos x).$$

Voorbeelden

$$T_0(x) = 1$$

$$T_2(x) = 2x^2 - 1$$

$$T_1(x) = x$$

$$T_3(x) = 4x^3 - 3x.$$

Recursieformule voor $T_n(x)$

$$T_{n+1}(x) = 2x T_n(x) - T_{n-1}(x)$$

Orthogonaliteitsrelaties

$$\int_{-1}^1 \frac{T_m(x) \cdot T_n(x)}{\sqrt{1-x^2}} dx = \begin{cases} 0 & m \neq n \\ \pi & m = n = 0 \\ \frac{\pi}{2} & m = n \neq 0. \end{cases}$$

De nulpunten van $T_n(x) = \cos(n \arccos x)$ berekenen we door $x = \cos \varphi$ te stellen. Dan is $T_n(x) = \cos n\varphi$, en de nulpunten zijn

$$\varphi_k = \frac{(2k-1)\pi}{2n} \quad \text{ofwel} \quad x_k = \cos \frac{(2k-1)\pi}{2n} \quad k = 1, 2, \dots, n$$

$$G_k = \int_{-1}^1 w(x) \ell_k(x) dx = \int_{-1}^1 \frac{T_m(x)}{(x-x_k) T'_m(x_k)} \frac{dx}{\sqrt{1-x^2}} =$$

$$= \frac{1}{T'_m(x_k)} \int_0^\pi \frac{\cos m\varphi}{\cos \varphi - \cos \varphi_k} d\varphi = \frac{1}{T'_m(x_k)} \cdot \frac{\pi \sin m\varphi_k}{\sin \varphi_k} = \frac{\pi}{m}$$

(Glauert)

dus alle gewichtscoefficiënten zijn hetzelfde.

De m-punts Gauss-Chebyshev integratieformule is dus

$$\int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx = \frac{\pi}{m} \sum_{k=1}^m f\left(\cos \frac{(2k-1)\pi}{2m}\right) + \frac{2\pi}{2^{2m}(2m!)} f^{(2m)}(\xi).$$

Opmerking

Er bestaat verband tussen de Gauss integratie en de kleinste kwadraten approximatie methode.

Kleinste kwadraten approximatie

Van een continue functie $f(x)$ op (a,b) noemen we het polynoom $p(x)$ onder de polynomen van de graad $\leq n$, de beste benadering in de zin van de kleinste kwadraten als de coëfficiënten $c_0, c_1, \dots, c_n$ van $p(x)$ zo gekozen zijn, dat

$$\int_a^b w(x) \cdot \{f(x) - p(x)\}^2 dx \quad \text{minimaal is.}$$

Gauss integratie

$$\int_a^b w(x) \cdot f(x) dx = \sum_{k=1}^m G_k \cdot f(x_k) + \frac{f^{(2m)}(\xi)}{(2m)!} \int_a^b w(x) \cdot \{\pi(x)\}^2 dx. \quad (3.20)$$

Hierbij zijn de basispunten $x_1, x_2, \dots, x_m$ zo gekozen dat

$$\int_a^b w(x) \cdot \pi(x) \cdot x^s dx = 0 \quad s < m \quad (3.29)$$

(3.29) blijkt gelijkwaardig te zijn met

$$E = \int_a^b w(x) \cdot \{\pi(x)\}^2 dx \quad \text{is minimaal.} \quad (3.30)$$

Stel namelijk $\pi(x) = x^m + c_{m-1}x^{m-1} + \dots + c_1x + c_0$
dan volgt uit (3.30)

$$\frac{\partial E}{\partial c_s} = 2 \int_a^b \omega(x) \cdot \pi(x) \cdot \frac{\partial \pi(x)}{\partial c_s} dx = 0,$$

dus

$$\int_a^b \omega(x) \cdot \pi(x) \cdot x^s dx = 0, \quad s < m.$$

We kunnen daarom ook zeggen dat de Gauss formule (3.20) ontstaat uit de Lagrange formule (3.13) als we daarin $x_1, x_2, \dots, x_m$ zo kiezen dat het interpolatie-polynoom, onder de polynomen van de graad $\leq m-1$, de beste benadering van $f(x)$ is in de zin van de kleinste kwadraten, als we veronderstellen dat $f(x)$ een m -de graads polynoom is.

4. Gewone differentiaalvergelijkingen

We onderscheiden de numerieke methoden voor het oplossen van gewone differentiaalvergelijkingen in twee soorten.

Eénstaps methoden (one-step methods). Bij deze methoden kan y_{n+1} berekend worden als alleen y_n bekend is. Geen kennis is nodig van voorgaande waarden $y_{n-1}, y_{n-2}, \dots$. Een dergelijke methode heet zelfstartend.

Voorbeelden zijn: Euler, Runge-Kutta.

Meerstaps methoden (multistep methods). Deze methoden hebben een aparte startprocedure nodig om de waarden $y_n, y_{n-1}, y_{n-2}, \dots$ uit te rekenen die nodig zijn voor de berekening van y_{n+1} .

Voorbeelden zijn: Milne, Adams.

Eénstaps methoden.

Een voorbeeld van een éénstaps methode is de methode van Euler. Het is de meest eenvoudige van alle numerieke methoden voor het oplossen van een differentiaalvergelijking met gegeven beginvoorwaarde (we noemen dit een beginwaarde-probleem). Euler's methode wordt in de praktijk weinig gebruikt vanwege de geringe nauwkeurigheid. We zullen desondanks Euler's methode nogal uitgebreid beschouwen, omdat het daarbij mogelijk is een aantal kenmerken op eenvoudige wijze te demonstreren, die specifiek zijn óók voor de meer gecompliceerde éénstaps methoden.

We beschouwen de eerste orde differentiaalvergelijking

$$y' = f(x, y) \quad y(x_0) = y_0. \quad (4.1)$$

Onder numeriek oplossen van deze vergelijking verstaan we het vinden

van de rij funktiewaarden $y(x_1), y(x_2), \dots$ met $x_i - x_{i-1} = h$.
 We integreren daarvoor (4.1), dit geeft

$$\int_{x_0}^{x_1} y' dx = \int_{x_0}^{x_1} f(x, y) dx$$

ofwel

$$y_1 = y_0 + \int_{x_0}^{x_1} f(x, y) dx.$$

Daar de integraal in het algemeen niet op te lossen is, zoeken we een benadering hiervoor. De simpelste benadering krijgen we door de integrand te vervangen door de waarde in x_0 , dus $f(x_0, y_0)$. Dit geeft de formule

$$y_1 = y_0 + h \cdot f(x_0, y_0).$$

Algemeen

$$y_{n+1} = y_n + h \cdot f(x_n, y_n) \quad (\text{Euler}). \quad (4.2)$$

Meetkundige interpretatie van Euler's formule:

$y' = f(x, y)$ definieert een richtingsveld in het (x, y) -vlak. De oplossing $y(x)$ van de d.v. is een kromme door (x_0, y_0) , waarvan de richting

in ieder punt overeenkomt met het richtingsveld.

Met Euler's formule krijgen we een polygoon dat door (x_0, y_0) gaat en waarvan de richting van de lijnstukken overeenkomt met de veldrichting in het linkerhoekpunt.

Een andere manier om (4.2) af te leiden.

Neem in (4.1) als benadering voor y' de voorwaartse differentie

$$\frac{y_{n+1} - y_n}{h} = f(x_n, y_n)$$

y_{n+1} hieruit oplossen geeft direct formule (4.2).

Tenslotte kunnen we (4.2) ook met behulp van de reeks van Taylor afleiden. Als $y(x_n)$ bekend is willen we $y(x_{n+1}) = y(x_n + h)$ berekenen.

De reeksontwikkeling van $y(x)$ rond het punt x_n is

$$y(x_n + h) = y(x_n) + hy'(x_n) + \frac{1}{2}h^2 y''(x_n) + \dots$$

We krijgen Euler's formule door van deze reeks alleen het lineaire stuk te nemen. We zien nu meteen welke afbrek-fout (truncation error) we daarbij maken.

Voorbeeld

$$y' = x - y^2$$

$$y(0) = 0$$

x_n	y_n	$f(x_n, y_n)$
0	0	0
0.1	0	0.10000
0.2	0.01000	0.19990
0.3	0.02999	0.29910
0.4	0.05990	

De differentievergelijking in het numerieke proces, die we als benadering voor de differentiaalvergelijking nemen, is een recursieformule voor de rij y_n . Is $f(x, y)$ voldoende simpel dan is het mogelijk de differentievergelijking exact op te lossen. We nemen als voorbeeld Euler's formule voor de differentiaalvergelijking

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

Formule (4.2) wordt dan

$$y_{n+1} = y_n + h y_n = (1+h) y_n.$$

En dus, omdat $y(0) = 1$ is de oplossing van de differentievergelijking

$$y_n = (1 + h)^n.$$

Nu is $x = nh$. Vullen we daarom $h = \frac{x}{n}$ in dan krijgen we

$$y_n = \left(1 + \frac{x}{n}\right)^n$$

en dit gaat naar e^x als $h \rightarrow 0$. We zien hieruit dat de oplossing van de differentiaalvergelijking willekeurig goed benaderd kan worden met Euler's methode, door de stapgrootte h maar voldoende klein te nemen.

Bij het oplossen van een differentiaalvergelijking met een numerieke methode maken we twee soorten fouten.

1. Discretiseringsfout of afbreekfout. Deze fout ontstaat doordat we de differentiaalvergelijking vervangen door een differentievergelijking. De oplossing y_n van de differentievergelijking is, ook bij

exact rekenen, slechts een benadering van de oplossing $y(x_n)$ van de differentiaalvergelijking in het punt $x_n = x_0 + nh$. Het ⁿverschil

$$e_n = y_n - y(x_n) \quad (4.3)$$

zullen we de discretiseringsfout noemen.

2. Afrondingsfout. We werken met getallen van eindige precisie. Na iedere bewerking wordt het resultaat afgerond. Daarom vinden we ook niet de exacte oplossing y_n van de differentievergelijking maar een benadering hiervan die we met $\tilde{y}_n$ zullen aanduiden. Het verschil

$$r_n = \tilde{y}_n - y_n \quad (4.4)$$

noemen we de afrondingsfout.
De totale fout is dan

$$\tilde{y}_n - y(x_n) = e_n + r_n. \quad (4.5)$$

Omdat zowel $y(x_n)$ als y_n oplossingen van een vergelijking zijn is de fout e_n een wel-gedefinieerde grootte, waar we dikwijls analytisch wel iets over kunnen zeggen. De afrondingsfouten daarentegen zijn niet vastgelegd door de oplossingsmethode, maar hangen samen met het aantal cijfers waarin gerekend wordt, de volgorde waarin de bewerkingen worden uitgevoerd, enz. .

Uit (4.5) volgt dat we het meest efficiënt rekenen als beide fouten vergelijkbaar van grootte zijn. Het is echter wel zinvol te proberen r_n zo klein te houden, dat de fout in de oplossing essentieel gelijk is aan e_n . Dat geeft ons de mogelijkheid om de nauwkeurigheid van het antwoord te bepalen.

We zullen nu enkele typen van foutenschattingen bekijken.

Een algoritme voor het oplossen van een numeriek probleem kan dikwijls beschouwd worden als een proces dat van een parameter, zeg p , afhangt. Deze parameter is bijvoorbeeld het aantal iteraties $p = n$ van een ite-

ratieproces, zoals Newton's $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$; of bij een differen-

tiaalvergelijking het aantal stappen $p = (b-a)/h$ dat nodig is om de differentiaalvergelijking te integreren van a naar b .

We veronderstellen nu dat de uitkomst van het algoritme convergeert naar de exacte oplossing van het probleem als $p \rightarrow \infty$. Verstaan we onder $e(p)$ de fout van het proces, dan kunnen we over de bruikbaarheid van de methode alleen iets zeggen als het gedrag van $e(p)$ voor $p \rightarrow \infty$ bekend is.

Er zijn verschillende mogelijkheden.

1. Het proces is convergent. In dit geval weten we alleen dat $e(p) \rightarrow 0$ als $p \rightarrow \infty$.
2. We weten iets over de convergentiesnelheid van het proces als er een bekende functie $\varphi(p)$ is met $\varphi(p) \rightarrow 0$ als $p \rightarrow \infty$ zodanig dat

$$e(p) = O(\varphi(p)) \quad p \rightarrow \infty.$$

In het geval van Simpson integratie bijv. geldt $e(h) = O(h^5)$.
De convergentiesnelheid is van de orde h^5 .

3. Een echte schatting krijgen we als we weten dat

$$|e(p)| \leq |\varphi(p)| \quad \text{voor} \quad p \geq p_0.$$

4. We hebben een asymptotische formule voor $e(p)$ als we weten dat

$$\frac{e(p)}{\varphi(p)} \rightarrow 1 \quad \text{als} \quad p \rightarrow \infty.$$

Van deze vier mogelijkheden is 3. de enige die een concrete uitspraak doet over de werkelijke grootte van de fout. Vaak zal echter deze fout-schatting veel te grof zijn. In de praktijk blijkt men dikwijls meer te hebben aan de orde van grootte van de fout, of aan het asymptotisch gedrag.

Het is soms mogelijk de convergentie van een numeriek proces te bewijzen zonder te veronderstellen dat het probleem een oplossing heeft. In dat geval kunnen we zeggen dat we de existentie van de oplossing van het probleem hebben bewezen met behulp van de convergentie van het proces.

Opmerking

Niet iedere differentiaalvergelijking heeft een oplossing. Ook is het niet zeker dat er slechts één oplossing bestaat. Zowel voor de existentie als voor de eenduidigheid van de oplossing zal de differentiaalvergelijking aan bepaalde voorwaarden moeten voldoen.

Existentiethorema voor de differentiaalvergelijking $y' = f(x, y)$
 $y(a) = \eta$. Als $f(x, y)$ voldoet aan de voorwaarden:

- a) $f(x, y)$ is gedefinieerd en continu in de strook $a \leq x \leq b$,
 $-\infty < y < \infty$.
- b) Er is een constante L zodat voor iedere x uit (a, b) en ieder tweetal y_1 en y_2 geldt

$$|f(x, y_1) - f(x, y_2)| \leq L |y_1 - y_2| \quad (4.6)$$

(Lipschitz)

dan bestaat er precies één functie $y(x)$, die oplossing is van de differentiaalvergelijking, d.w.z.

- 1) $y(x)$ is continu en differentieerbaar in (a, b) ,
- 2) $y(x)$ voldoet aan de vergelijking $y'(x) = f(x, y(x))$,
- 3) $y(a) = \eta$.

Een opmerking over voorwaarde b).

Het is de z.g. Lipschitz-voorwaarde die zegt dat het differentie-quotiënt van $f(x,y)$ naar y begrensd is voor alle x . Aan de Lipschitz-voorwaarde is zeker voldaan als $\frac{\partial f}{\partial y}$ bestaat en begrensd is. Dit laatste geval zullen we meestal tegenkomen.

Het bewijs van het existentiëtheorema kan gegeven worden door te laten zien dat onder de voorwaarden a) en b) de Euler methode convergent is. De limiet is dan de (eenduidig bepaalde) oplossing van de differentiaalvergelijking.

We zullen nu een schets van dit bewijs geven.

Allereerst maken we een rij benaderingen $y_p(x)$ van $y(x)$ met de methode van Euler.

We definiëren

$$h_p = \frac{b-a}{2^p}, \quad p = 0, 1, 2, \dots$$

en geven de oplossing van Euler's formule met stapgrootte h_p aan met $y_p(x)$. Dan is $y_p(x)$ een polygoon met hoekpunten in de punten $x_m = a + mh_p$, dus als $x_m \leq x \leq x_{m+1}$ geldt

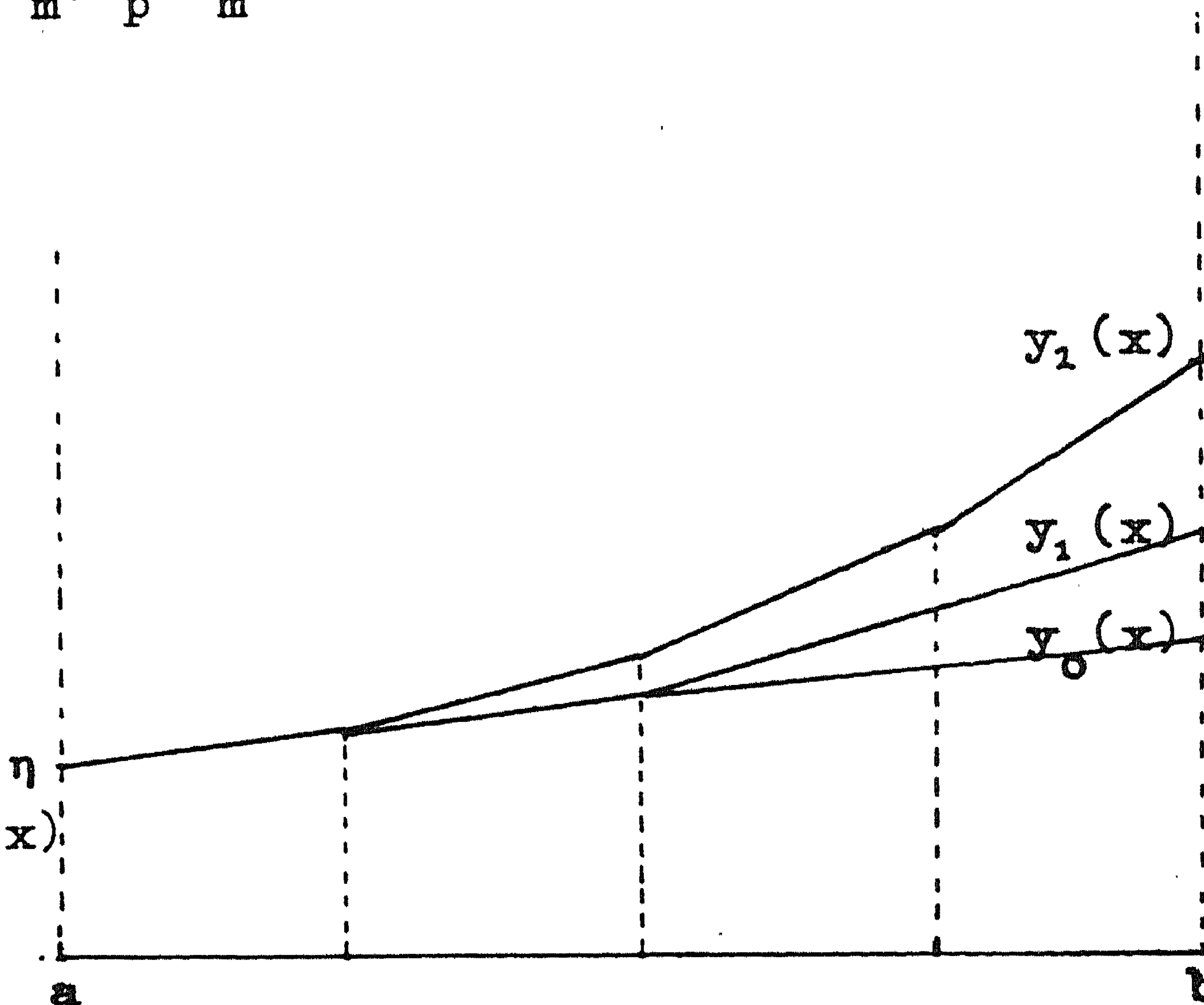
$$y_p(x) = y_p(x_m) + (x-x_m) \cdot f(x_m, y_p(x_m)) \quad (4.7)$$

Men kan nu laten zien dat de rij functies $y_p(x)$ voor $p \rightarrow \infty$

uniform convergent is in het interval (a,b) . Uit de uniforme convergentie volgt, dat de limietfunctie $y(x)$ continu is, omdat $y_p(x)$ continu is voor alle p .

Dat $y(a) = \eta$ is duidelijk, want $y_p(a) = \eta$ voor alle p .

Nu moet nog bewezen worden dat $y(x)$ differentieerbaar is en voldoet aan de differentiaalvergelijking. Voor willekeurige p geldt



$$y_p(x) - y_p(a) = y_p(x) - y_p(x_m) + \sum_{i=0}^{m-1} [y_p(x_{i+1}) - y_p(x_i)]$$

waarvoor we met behulp van (4.7) kunnen schrijven

$$y_p(x) - y_p(a) = (x-x_m) \cdot f(x_m, y_p(x_m)) + \sum_{i=0}^{m-1} h_p \cdot f(x_i, y_p(x_i)) \quad (4.8)$$

Met behulp van de uniforme convergentie van de rij $y_p(x)$ en de continuïteit van $f(x,y)$ kan men dan bewijzen dat het rechterlid van (4.8) voor $p \rightarrow \infty$ convergeert naar $\int_a^x f(x,y(x))dx$, zodat we tenslotte voor de limietfunctie $y(x)$ vinden

$$y(x) = y(a) + \int_a^x f(x,y(x))dx,$$

waaruit direct volgt dat $y(x)$ differentieerbaar is en voldoet aan de differentiaalvergelijking.

Met de Lipschitz-voorwaarde kan men aantonen dat $y(x)$ de enige oplossing van de differentiaalvergelijking is.

We zullen voor de discretiseringsfout e_n bij Euler een schatting afleiden. We hebben daarvoor het volgende lemma nodig.

Lemma

Als de rij getallen ξ_n voldoen aan de recurrente relatie

$$|\xi_{n+1}| \leq A |\xi_n| + B \quad (4.9)$$

waarbij A en B niet-negatieve constanten zijn, onafhankelijk van n , dan geldt

$$|\xi_n| \leq A^n |\xi_0| + \begin{cases} \frac{A^n - 1}{A - 1} \cdot B & A \neq 1 \\ nB & A = 1 \end{cases} \quad (4.10)$$

Bewijs. Met volledige inductie. Voor $n = 1$ is (4.10) gelijk aan (4.9) en dus waar.

Veronderstel dat (4.10) geldt voor $n = N$, dan vinden we in het geval $A \neq 1$

$$\begin{aligned} |\xi_{N+1}| &\leq A \cdot \left\{ A^N \cdot |\xi_0| + \frac{A^N - 1}{A - 1} \cdot B \right\} + B \\ &= A^{N+1} \cdot |\xi_0| + \frac{A^{N+1} - 1}{A - 1} \cdot B \end{aligned}$$

en dit is (4.10) voor $n = N + 1$.
Het geval $A = 1$ bewijst men analoog.

In de praktijk komt het vaak voor dat A een grootte is van de vorm $1+\delta$, waarbij δ een klein positief getal is. Dan is

$$A = 1+\delta < e^\delta, \quad \delta > 0$$

en dus

$$|\xi_n| \leq e^{n\delta} |\xi_0| + \frac{e^{n\delta} - 1}{\delta} \cdot B \quad (4.11)$$

We bewijzen nu de volgende stelling.

Stelling

Veronderstel dat $f(x,y)$ voldoet aan de Lipschitz-voorwaarde, en dat $y(x)$ tweemaal continu differentieerbaar is. Zij

$$N(x) = \frac{1}{2} \max_{a \leq t \leq x} |y''(t)|$$

dan geldt voor de fout e_n in Euler's methode

$$|e_n| \leq h \cdot N(x_n) \cdot E_L(x_n - a) \quad (4.12)$$

waarbij

$$E_L(x) = \begin{cases} \frac{e^{Lx} - 1}{L} & L > 0 \\ x & L = 0 \end{cases} \quad \text{(Lipschitzfunctie)} \quad (4.13)$$

Bewijs.

$$y_{m+1} = y_m + h \cdot f(x_m, y_m) \quad \text{(Euler's formule)}$$

$$y(x_{m+1}) = y(x_m + h) = y(x_m) + h \cdot y'(x_m) + \frac{h^2}{2} y''(\xi) \quad \text{(Taylor's formule)}$$

$$\text{waarbij } x_m < \xi < x_{m+1}$$

Dan is

$$e_{m+1} = e_m + h [f(x_m, y_m) - f(x_m, y(x_m))] - \frac{1}{2} h^2 y''(\xi).$$

Dus

$$|e_{m+1}| \leq |e_m| + h |f(x_m, y_m) - f(x_m, y(x_m))| + \frac{1}{2} h^2 |y''(\xi)|.$$

We gebruiken de voorwaarde van Lipschitz

$$|e_{m+1}| \leq (1+hL) |e_m| + h^2 N(x_{m+1}).$$

Nu geldt ook, wegens de definitie van $N(x)$, voor $k = 1, 2, \dots, n-1$

$$|e_{k+1}| \leq (1+hL) |e_k| + h^2 N(x_n).$$

Stellen we $A = (1+hL)$ en $B = N(x_n)$ dan zijn A en B onafhankelijk van k . We passen het lemma van blz. NA-30 toe en vinden dan

$$|e_k| \leq e^{khL} \cdot |e_0| + \frac{e^{khL} - 1}{hL} \cdot h^2 N(x_n).$$

Nu is $e_0 = 0$ en $x_n = a + nh$, nemen we verder $k = n$ dan krijgen we tenslotte voor e_n de schatting

$$|e_n| \leq \frac{e^{(x_n-a)L} - 1}{L} \cdot h \cdot N(x_n)$$

die gelijkwaardig is met (4.12).

Daar $y''(t) = f_x(t, y) + f_y(t, y) \cdot y'(t) = f_x(t, y) + f_y(t, y) \cdot f(t, y)$ vinden we voor $N(x)$ de volgende schatting

$$2N(x) \leq \max \{|f_x + f_y \cdot f|; \quad a \leq t \leq x, \quad -\infty < y < \infty\}.$$

Stelling

Laat $N(x)$ gedefinieerd zijn zoals in stelling (NA-31) en $y(x)$ en $f(x, y)$ voldoen aan de in deze stelling gestelde voorwaarden. Veronderstel dat de rij y_n oplossing is van

$$y_{n+1} = y_n + h \cdot f(x_n, y_n) + c \cdot \theta_n \cdot h^2 \quad (4.14)$$

$$y_0 = \eta,$$

waarin $c \geq 0$ een constante is en $|\theta_n| \leq 1$.

Dan geldt als schatting voor $e_n = y_n - y(x_n)$ de volgende ongelijkheid

$$|e_n| \leq h \cdot \{N(x_n) + c\} \cdot E_L(x_n - a). \quad (4.15)$$

Het bewijs van (4.15) is geheel analoog aan het bewijs van (4.12).

Deze stelling zegt dat als de waarden y_n niet precies de oplossing van Euler's formule zijn, dan convergeert deze rij toch ook naar de exacte oplossing van de differentiaalvergelijking voor $h \rightarrow 0$, als de afwijking per stap niet meer bedraagt dan ch^2 .

We beschouwen de ongelijkheid (4.12) voor een vast punt $x = x_n$. Dan geeft deze een bovengrens in de vorm $K(x).h$ aan van de fout die we maken in het berekenen van $y(x)$. Dat de werkelijke fout in een concreet geval veel kleiner kan zijn dan deze bovengrens zien we aan de volgende voorbeelden.

Voorbeeld

$y' = y$ $y(0) = 1$ opgelost met Euler's methode, met $h = 2^{-6}$.
In dit geval is de Lipschitz constante $L = 1$, $N(x) = \frac{1}{2}e^x$ en dus wordt het rechterlid van (4.12)

$$K(x).h = \frac{1}{2}h.e^x(e^x - 1).$$

x_n	1	2	3	4	5
y_n	2.69735	7.27567	19.62499	52.93537	142.7850
e_n	-0.02093	-0.11339	-0.46055	-1.66278	-5.6282
$K(x_n).h$	0.03649	0.36882	2.99487	22.86218	170.9223

Voorbeeld

$y' = -y$ $y(0) = 1$ opgelost met Euler's methode.
 $h = 2^{-6}$, $L = 1$, $N(x) = \frac{1}{2}$, $K(x).h = \frac{1}{2}h(e^x - 1)$.

x_n	1	2	3	4	5
y_n	0.364987	0.133215	0.048622	0.017746	0.006477
e_n	-0.002892	-0.002120	-0.001165	-0.000570	-0.000261
$K(x_n).h$	0.013424	0.049914	0.149106	0.418735	1.151666

We zien dat de foutenschatting (4.12) in deze twee eenvoudige voorbeelden volkomen waardeloos is, en we mogen niet verwachten dat het in een meer gecompliceerd geval beter zal zijn.

We zullen nu een resultaat afleiden dat niet aangeeft hoe groot de fout in Euler's methode mogelijkerwijs zou kunnen zijn (4.12), maar hoe groot de fout werkelijk is, althans bij benadering.

Uit de ongelijkheid (4.12) kunnen we in ieder geval besluiten dat de getallen $h^{-1}|e_n|$ begrensd zijn als $h \rightarrow 0$, als we voor x_n een vast punt x nemen.

In onderstaande tabel worden de waarden y_n , e_n en $h^{-1} \cdot e_n$ van de differentiaalvergelijking $y' = -y$ $y(0) = 1$, in het punt $x = 1$, gegeven, berekend met verschillende stapgrootten $h = 2^{-k}$ $k = 1, 2, 3, \dots$

k	y_n	e_n	$h^{-1} e_n$
1	0.250000	-0.117879	-0.235758
2	0.316406	-0.051473	-0.205893
3	0.343609	-0.024270	-0.194164
4	0.356074	-0.011805	-0.188885
5	0.362055	-0.005824	-0.186373
6	0.364987	-0.002892	-0.185147
7	0.366438	-0.001441	-0.184540
8	0.367160	-0.000719	-0.184239
.	.		
.	.		
.	.		
∞	0.367879		

We zien in deze tabel dat de rij getallen $h^{-1} \cdot e_n$ niet alleen begrensd is, maar ook een limiet heeft voor $n \rightarrow \infty$. We zullen bewijzen dat dit algemeen het geval is bij Euler's methode. In het bovenstaande speciale geval zal blijken dat

$$\lim_{n \rightarrow \infty} h^{-1} \cdot e_n = -\frac{1}{2}e^{-1} = -0.183940.$$

We veronderstellen dat $f(x, y)$ voldoet aan de voorwaarden a) en b) (NA-28). Verder veronderstellen we dat $f(x, y)$ continue en begrensde eerste en tweede partiële afgeleiden heeft. Dan bestaat de derde afgeleide van $y(x)$ en is eveneens continu en begrensd, en we mogen schrijven

$$y(x_{n+1}) = y(x_n) + hf(x_n, y(x_n)) + \frac{h^2}{2} \cdot y''(x_n) + \frac{h^3}{6} \cdot y'''(\xi) \quad (4.17)$$

waarbij $x_n < \xi < x_{n+1}$.

(4.17) aftrekken van (4.2) geeft (daar $y_n = y(x_n) + e_n$)

$$e_{n+1} = e_n + h \cdot [f(x_n, y(x_n) + e_n) - f(x_n, y(x_n))] - \frac{1}{2}h^2 y''(x_n) - \frac{1}{6}h^3 y'''(\xi). \quad (4.18)$$

De uitdrukking tussen vierkante haken kunnen we met Taylor's formule als volgt schrijven

$$f(x_n, y(x_n) + e_n) - f(x_n, y(x_n)) = e_n \cdot f_y(x_n, y(x_n)) + \frac{1}{2}e_n^2 \cdot f_{yy}(x_n, y^*) \quad (4.19)$$

waarbij y^* gelegen is tussen y_n en $y(x_n)$.

We vullen (4.19) in (4.18) in, en delen deze vervolgens door h . Stellen we dan nog

$$\bar{e}_n = h^{-1} e_n$$

dan krijgen we de relatie

$$\bar{e}_{n+1} = \bar{e}_n + h \cdot [f_y(x_n, y(x_n)) \cdot \bar{e}_n - \frac{1}{2}y''(x_n)] + h^2 \cdot r_n \quad (4.20)$$

waarbij

$$r_n = -\frac{1}{6}y'''(\xi) + \frac{1}{2}f_{yy}(x_n, y^*) \cdot \bar{e}_n^2.$$

Nu is $|\bar{e}_n| \leq N(x_n) \cdot E_L(x_n - a)$, dus begrensd bij vaste x_n . Hieruit en uit de begrensdheid van de eerste en tweede afgeleiden van $f(x, y)$ volgt gemakkelijk dat ook r_n begrensd is, dus

$$|r_n| \leq c_1.$$

Als we de funktie $g(x)$ op de volgende manier definieren

$$g(x) = f_y(x, y(x)),$$

dan kunnen we (4.20) opvatten als Euler's formule toegepast op de

differentiaalvergelijking voor $e(x)$

$$e'(x) = g(x) \cdot e(x) - \frac{1}{2} y''(x) \quad (4.21)$$

met een additionele fout per stap die kleiner is dan $c_1 h^2$.

Omdat $\bar{e}_0 = 0$ nemen we als beginwaarde bij (4.21) $e(a) = 0$.

Als we nu op (4.21) met benaderingsformule (4.20) stelling (NA-32) toepassen, dan krijgen we het beloofde resultaat voor e_n . Er geldt namelijk de volgende

Stelling

Laat $f(x,y)$ voldoen aan de voorwaarden a) en b) van blz NA-28 en verder continue en begrensde eerste en tweede partiële afgeleiden hebben. Dan geldt voor de fout e_n in de methode van Euler voor het oplossen van de differentiaalvergelijking $y' = f(x,y)$ met $y(a) = \eta$

$$e_n = h \cdot e(x_n) + O(h^2) \quad (4.22)$$

waarbij $e(x)$ de oplossing is van de differentiaalvergelijking

$$e'(x) = f_y(x, y(x)) \cdot e(x) - \frac{1}{2} y''(x) \quad (4.23)$$

die voldoet aan de beginwaarde $e(a) = 0$.

Meer precies geformuleerd: voor e_n geldt

$$\left| \frac{1}{h} \cdot e_n - e(x_n) \right| \leq h \cdot c_3 \cdot E_L(x_n - a). \quad (4.24)$$

Hierin kunnen we volgens de uitgebreide stelling (NA-32) nemen

$$c_3 = \frac{1}{2} \max_{a \leq x \leq b} |e''(x)| + c_1,$$

terwijl L de Lipschitz constante is voor het rechterlid van (4.23) waarvoor we kunnen nemen

$$L = \max_{a \leq x \leq b} |f_y(x, y(x))|.$$

We passen het verkregen resultaat toe op dezelfde voorbeelden.

In het eerste voorbeeld $y' = y$ $y(0) = 1$ hebben we de oplossing $y(x) = e^x$.
 Dan is $g(x) = 1$ en $y''(x) = e^x$ zodat de differentiaalvergelijking voor $e(x)$ wordt

$$e'(x) = e(x) - \frac{1}{2}e^x \quad e(0) = 0.$$

De oplossing hiervan is $e(x) = -\frac{1}{2}xe^x$. Voor e_n vinden we dus

$$e_n = -\frac{1}{2}x_n e^{x_n} \cdot h + O(h^2).$$

In het tweede voorbeeld $y' = -y$ $y(0) = 1$ is de exacte oplossing $y(x) = e^{-x}$, $g(x) = -1$, $y''(x) = e^{-x}$. Dan wordt de differentiaalvergelijking voor $e(x)$

$$e'(x) = -e(x) - \frac{1}{2}e^{-x} \quad e(0) = 0$$

en de oplossing hiervan is $e(x) = -\frac{1}{2}xe^{-x}$. In dit geval vinden we voor de afbreekfout

$$e_n = -\frac{1}{2}x_n e^{-x_n} \cdot h + O(h^2).$$

Stellen we hierin $x_n = 1$, dan vinden we (daar $h = \frac{1}{n}$)

$$\lim_{h \rightarrow 0} h^{-1} \cdot e_n = -\frac{1}{2}e^{-1}$$

hetgeen overeenstemt met blz. NA-34.

Ter illustratie geven we nog onderstaande tabel voor het geval $y' = -y$ $y(0) = 1$, $h = 2^{-6}$

x_n	1	2	3	4	5
e_n	-0.002892	-0.002120	-0.001165	-0.000570	-0.000261
$h \cdot e(x_n)$	-0.002874	-0.002114	-0.001167	-0.000572	-0.000263

Ook zonder $e(x)$ expliciet te kennen kunnen we (4.22) gebruiken om een reële schatting van de fout te krijgen, en daarmee een nauwkeuriger antwoord.

Daar een benaderende oplossing afhankelijk is van de stapgrootte noteren we zo'n oplossing nu als $y(x, h)$. Dan geldt

$$y(x, h) = y(x) + h \cdot e(x) + O(h^2). \quad (4.25)$$

De fout $h \cdot e(x) + O(h^2)$ zal in de meeste gevallen essentieel gelijk zijn aan $h \cdot e(x)$, speciaal bij kleine h . Veronderstel nu dat we twee benaderende oplossingen hebben bepaald voor respectievelijk de stapgrootte h en qh . Voor deze laatste geldt dan

$$y(x, qh) = y(x) + qh \cdot e(x) + O(h^2).$$

Uit deze relaties kunnen we de grootte van de fout bepalen

$$h \cdot e(x) = \frac{y(x, h) - y(x, qh)}{1 - q} + O(h^2). \quad (4.26)$$

Vullen we deze betrekking in (4.25) in dan vinden we

$$y(x) = y(x, h) - \frac{y(x, h) - y(x, qh)}{1 - q} + O(h^2). \quad (4.27)$$

(4.27) geeft een nauwkeuriger resultaat dan (4.25) in die zin dat de fout van hogere orde in h is. Dit procédé, om uit twee numerieke oplossingen een beter benadering af te leiden, is afkomstig van L.F. Richardson, en werd door hem genoemd: Deferred approach to the limit. We zouden het ook extrapolatie naar $h = 0$ kunnen noemen. Het is toepasbaar op elk numeriek proces waarvan bekend is dat de fout essentieel gelijk is aan $c \cdot h^n$.

We demonstreren de h -extrapolatie nog in onderstaande tabel voor $y' = -y$ $y(0) = 1$ in het punt $x = 1$, met $q = \frac{1}{2}$.

k	$y(1, 2^{-k})$	$y(\text{geëxtrapoleerd})$
4	0.356074	
5	0.362055	0.368036
6	0.364987	0.367919
7	0.366438	0.367889
8	0.367160	0.367882
.	.	
.	.	
.	.	
∞	0.367879	

We behandelen nu nog enkele meer effectieve one-step methoden in die zin dat de nauwkeurigheid evenredig is met h^p met $p > 1$ waarbij h de stapgrootte is.

Een éénstaps methode voor het oplossen van het beginwaarde probleem (4.1) is een algoritme dat uit de waarden x_n , y_n en h de waarde y_{n+1} als benadering van de oplossing in het punt $x_{n+1} = x_n + h$ berekent. In formule kunnen we dit als volgt weergeven

$$y_{n+1} = y_n + h \cdot \Phi(x_n, y_n; h). \quad (4.28)$$

In het geval van Euler's methode hadden we $\Phi(x_n, y_n; h) = f(x_n, y_n)$ dus onafhankelijk van h .

$h \cdot \Phi(x_n, y_n; h)$ is de aangroeiing van de benaderende oplossing. De aangroeiing van de exacte oplossing bepalen we op de volgende manier.

Zij (x, y) een willekeurig punt, $a \leq x \leq b$ en $-\infty < y < \infty$.

Zij $z(t)$ de oplossing van de differentiaalvergelijking

$$z' = f(t, z),$$

die gaat door het punt (x, y) , dus $z(x) = y$. Voor willekeurige h waarvoor $a \leq x+h \leq b$ definiëren we dan

$$\Delta(x, y; h) = \begin{cases} \frac{z(x+h) - y}{h} & h \neq 0 \\ f(x, y) & h = 0 \end{cases} \quad (4.29)$$

dan is $h \cdot \Delta(x, y; h)$ de aangroeiing van x naar $x+h$ van de exacte oplossing van het probleem.

We veronderstellen nu dat $f(x, y)$ willekeurig vaak differentieerbaar is. Dan is ook de oplossing $y(x)$ van (4.1) willekeurig vaak differentieerbaar, en de afgeleiden van $y(x)$ zijn volkomen bepaald door $f(x, y)$ en zijn partiële afgeleiden. Ter vereenvoudiging schrijven we

$$y^{(n)} = f^{(n-1)}(x, y)$$

waarbij we met het rechterlid bedoelen de totale afgeleide naar x van de funktie $f(x, y(x))$. Met deze notatie wordt

$$z(x+h) - z(x) = h \cdot f(x, y) + \frac{h^2}{2!} f'(x, y) + \dots + \frac{h^p}{p!} f^{(p-1)}(x, y) + O(h^{p+1})$$

en we vinden voor de aangroeiing van de exacte oplossing

$$\Delta(x,y;h) = f(x,y) + \frac{h}{2!}f'(x,y) + \dots + \frac{h^{p-1}}{p!}f^{(p-1)}(x,y) + \mathcal{O}(h^p) \quad (4.30)$$

We willen $\Phi(x,y;h)$ zo kiezen, dat deze zo goed mogelijk overeenstemt met $\Delta(x,y;h)$. Een voor de hand liggende keuze is dan

$$\Phi(x,y;h) = f(x,y) + \frac{h}{2!}f'(x,y) + \dots + \frac{h^{p-1}}{p!}f^{(p-1)}(x,y). \quad (4.31)$$

Vullen we deze $\Phi(x,y;h)$ in (4.28) in dan krijgen we een methode die we de Taylor methode zullen noemen. We zeggen dat de orde van het proces p is, als de fout in $\Phi(x,y;h)$ $\mathcal{O}(h^p)$ is.

De methode van Euler is een bijzonder geval van de Taylor methode, nl voor $p = 1$.

Numeriek biedt deze methode weinig perspectief omdat we behalve $f(x,y)$ ook de afgeleiden $f'(x,y) \dots f^{(p-1)}(x,y)$ nodig hebben. Taylor's methode is dan ook in het algemeen geen praktische methode.

We proberen nu een formule voor $\Phi(x,y;h)$ te vinden zodat

$$\Phi(x,y;h) - \Delta(x,y;h) = \mathcal{O}(h^p) \quad p > 1$$

en waarbij geen afgeleiden van $f(x,y)$ nodig zijn. In plaats van de afgeleiden zullen we dan meer funktiewaarden moeten berekenen. We stellen daarom

$$\Phi(x,y;h) = a_1 f(x,y) + a_2 f(x+p_1 h, y+p_2 h) \quad (4.32)$$

waarin de constanten a_1 , a_2 , p_1 en p_2 nog bepaald moeten worden. Met behulp van de Taylor ontwikkeling voor $f(x,y)$ kunnen we voor (4.32) ook schrijven

$$\Phi(x,y;h) = (a_1 + a_2)f(x,y) + a_2[p_1 f_x + p_2 f_y \cdot f]h + \mathcal{O}(h^2). \quad (4.33)$$

Voor $p = 2$ wordt (4.30)

$$\Delta(x,y;h) = f(x,y) + \frac{1}{2}h[f_x + f_y \cdot f] + \mathcal{O}(h^2). \quad (4.34)$$

We kiezen de constanten a_1 , a_2 , p_1 en p_2 nu zo dat $\Phi(x,y;h)$ en $\Delta(x,y;h)$ zo goed mogelijk overeenkomen. Vergelijken van de overeenkomstige termen van (4.33) en (4.34) geeft de betrekkingen

$$a_1 + a_2 = 1 \quad a_2 p_1 = \frac{1}{2} \quad a_2 p_2 = \frac{1}{2}. \quad (4.35)$$

Als oplossingen hiervan vinden we

$$a_1 = 1 - \alpha \quad a_2 = \alpha \quad p_1 = p_2 = \frac{1}{2\alpha}$$

waarin de parameter $\alpha \neq 0$ willekeurig gekozen kan worden. Deze waarden van de constanten ingevuld in (4.32) geeft

$$\Phi(x,y;h) = (1-\alpha)f(x,y) + \alpha f\left(x + \frac{h}{2\alpha}, y + \frac{h}{2\alpha}f(x,y)\right). \quad (4.36)$$

De oplossingsmethode (4.28) met deze aangroeiingsfunctie wordt wel de vereenvoudigde Runge-Kutta methode genoemd. De aangroeiingsfunctie $\Phi(x,y;h)$ wordt bepaald als het gewogen gemiddelde van twee funktiewaarden.

Voor $\alpha = \frac{1}{2}$ krijgen we de verbeterde Euler methode ook wel bekend als de methode van Heun

$$y_{n+1} = y_n + \frac{h}{2}(f(x_n, y_n) + f(x_n + h, y_n + hf(x_n, y_n))).$$

Voor $\alpha = 1$ krijgen we de formule

$$y_{n+1} = y_n + hf\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}hf(x_n, y_n)\right)$$

die wel de verbeterde polygoonmethode of ook de gemodificeerde Euler methode genoemd wordt.

De klassieke Runge-Kutta methode kan op analoge wijze als hierboven verkregen worden. $\Phi(x,y;h)$ is dan het gewogen gemiddelde van vier funktiewaarden, waarbij het argument van iedere volgende funktiewaarde afhangt van de zojuist berekende.

We geven alleen de formule

$$\Phi(x,y;h) = \frac{1}{6} (k_1 + 2k_2 + 2k_3 + k_4) \quad (4.37)$$

met

$$\begin{aligned}
 k_1 &= f(x, y) \\
 k_2 &= f(x + \frac{1}{2}h, y + \frac{1}{2}hk_1) \\
 k_3 &= f(x + \frac{1}{2}h, y + \frac{1}{2}hk_2) \\
 k_4 &= f(x + h, y + hk_3)
 \end{aligned}
 \tag{4.38}$$

De orde van deze Runge-Kutta methode is 4. Dit kan op dezelfde manier bewezen worden als bij de tweede orde Runge-Kutta. We bewijzen het niet maar kunnen wel plausibel maken dat het zo is. Nemen we nl. $f(x, y) = f(x)$ dus onafhankelijk van y , dan wordt $k_2 = k_3$ en de Runge-Kutta formule gaat over in de Simpson integratieformule.

Nu is de fout in de Simpson formule $O(h^5)$, zodat het in dit speciale geval klopt, dat de orde van Runge-Kutta gelijk aan 4 is.

Meerstaps methoden

De meestaps methoden voor het oplossen van het beginwaarde-probleem

$$y' = f(x, y) \quad y(a) = \eta \tag{4.39}$$

zijn gebaseerd op interpolatieformules voor de functie $f(x, y)$.

Zij $z(x)$ een functie gedefinieerd op (a, b) en daar voldoende vaak differentieerbaar. Kiezen we in (a, b) de punten $x_0, x_1, \dots, x_n$, dan is er precies één polynoom $p(x)$ van de graad $\leq n$ waarvoor geldt

$$p(x_k) = z(x_k) \quad k = 0, 1, 2, \dots, n.$$

De Lagrange formule voor dit polynoom luidt

$$p(x) = \sum_{k=0}^n L_k(x) \cdot z(x_k).$$

De afwijking van het polynoom in een willekeurig punt x t.o.v. de funktiewaarde wordt gegeven door

$$E(x) = z(x) - p(x) = \frac{\pi(x) \cdot z^{(n+1)}(\xi)}{(n+1)!} \tag{4.40}$$

(zie ook blz NA-1 ev).

We zullen de afgeleide van $p(x)$ gebruiken als benadering voor $z'(x)$. De grootte van de fout in de basispunten x_k kunnen we met behulp van de volgende stelling schatten.

Stelling

Als $z(x)$ een continue $(n+1)$ -de afgeleide heeft in (a,b) dan is bij ieder punt x_k $k=0,1,2,\dots,n$, een ξ_k te vinden zodanig dat

$$z'(x_k) - p'(x_k) = \frac{\pi'(x_k) \cdot z^{(n+1)}(\xi_k)}{(n+1)!} . \quad (4.41)$$

Bewijs: We veronderstellen $x_0 < x_1 < x_2 < \dots < x_n$. Voor iedere keuze van de parameter λ heeft

$$Z(x) = z(x) - p(x) - \lambda \pi(x)$$

$n+1$ nulpunten, nl. de basispunten x_k . Volgens Rolle heeft $Z'(x)$ dan n nulpunten, gelegen tussen ieder tweetal basispunten. Bij een willekeurige k kiezen we $\lambda = \lambda_k$ zodanig dat ook $Z'(x_k) = 0$. Dit is het geval als

$$\lambda_k = \frac{z'(x_k) - p'(x_k)}{\pi'(x_k)} . \quad (4.42)$$

Bij deze keuze van λ heeft $Z'(x)$ dan $n+1$ verschillende nulpunten in (a,b) . Volgens Rolle heeft $Z^{(n+1)}(x)$ dan minstens één nulpunt, zeg ξ_k , in (a,b) . Dan is dus

$$Z^{(n+1)}(\xi_k) = z^{(n+1)}(\xi_k) - \lambda_k \cdot (n+1)! = 0. \quad (4.43)$$

Eliminatie van λ_k uit (4.42) en (4.43) geeft (4.41).

We nemen in het volgende aan dat de basispunten aequidistant gelegen zijn, dus $x_k - x_{k-1} = h$. Dan is $x_k = x_0 + k.h$. Verder schrijven we $z_k = z(x_k)$. De achterwaartse differenties van $z(x)$ zijn gedefinieerd door

$$\nabla z_k = z_k - z_{k-1} \quad \nabla^n z_k = \nabla(\nabla^{n-1} z_k). \quad (4.44)$$

Stellen we $\nabla^0 z_k = z_k$, dan is met volledige inductie eenvoudig te bewijzen dat

$$\nabla^n z_k = \sum_{m=0}^n (-1)^m \binom{n}{m} z_{k-m} \quad n = 0, 1, 2, \dots \quad (4.45)$$

Formule (4.45) geeft de achterwaartse differenties uitgedrukt in de funktiewaarden. Omgekeerd is het ook mogelijk de funktiewaarden uit te drukken in de differenties. Er geldt namelijk

$$z_{k-n} = \sum_{m=0}^n (-1)^m \binom{n}{m} \nabla^m z_k \quad n = 0, 1, 2, \dots \quad (4.46)$$

De (achterwaartse) interpolatieformule van Newton geeft het polynoom $p(x)$, van de graad $\leq n$, dat samenvalt met $z(x)$ in de punten $x_k, x_{k-1}, \dots, x_{k-n}$ uitgedrukt in de achterwaartse differenties:

$$p(x) = \sum_{m=0}^n (-1)^m \binom{-s}{m} \nabla^m z_k. \quad (4.47)$$

Hierin is $s = \frac{x - x_k}{h}$.

We veronderstellen in het vervolg dat $f(x, y)$ voldoet aan de voorwaarden a) en b) uit het existentie-theorema, blz NA-28.

Door integratie vinden we dat de exacte oplossing $y(x)$ van (4.39) voldoet aan de formule

$$y(x+h) - y(x) = \int_x^{x+h} f(t, y(t)) dt \quad (4.48)$$

voor ieder tweetal punten x en $x+h$ in het interval (a, b) . De meerstaps methoden die behandeld zullen worden, verkrijgen we door in (4.48) de onbekende funktie $f(x, y(x))$ te vervangen door het interpolatiepolynoom $p(x)$ dat in de punten x_i de waarde $f_i = f(x_i, y_i)$ heeft, waarbij y_i reeds berekende funktiewaarden zijn, of juist de te berekenen funktiewaarde is.

We nemen als interpolatiepunten $x_n, x_{n-1}, \dots, x_{n-k}$. Het interpolatiepolynoom dat $f(x, y(x))$ vervangt is dan

$$p(x) = \sum_{m=0}^k (-1)^m \binom{-s}{m} \nabla^m f_n.$$

Afhankelijk van de ligging van x en $x+h$ ten opzichte van de basispunten, kunnen we de meerstaps methoden onderscheiden in verschillende klassen.

Methode	x	x+h
Adams-Bashforth	x_n	x_{n+1}
Adams-Moulton	x_{n-1}	x_n
Nyström	x_{n-1}	x_{n+1}
Milne-Simpson	x_{n-2}	x_n

Adams-Bashforth

Bij deze methode wordt formule (4.48)

$$y_{n+1} - y_n = \int_{x_n}^{x_{n+1}} p(x) dx = h \sum_{m=0}^k \gamma_m \cdot \nabla^m f_n \quad (4.49)$$

waarbij de coëfficiënten

$$\gamma_m = (-1)^m \frac{1}{h} \int_{x_n}^{x_{n+1}} \binom{-s}{m} dx = (-1)^m \int_0^1 \binom{-s}{m} ds \quad (4.50)$$

onafhankelijk zijn van de funktie $f(x,y)$. Als nu de waarden $y_n, y_{n-1}, \dots, y_{n-k}$, en dus ook $f_n = f(x_n, y_n), f_{n-1}, \dots, f_{n-k}$ bekend zijn, kunnen we met (4.49) de waarde y_{n+1} berekenen. Vervolgens wordt n met 1 verhoogd en dezelfde methode wordt gebruikt om y_{n+2} te berekenen, etc. Omdat y_{n+1} direct berekend kan worden als $y_n, y_{n-1}, \dots$ bekend zijn noemen we de methode ook wel een expliciete methode. Deze methode heeft, zoals iedere multistep methode, een aparte start-procedure nodig om de eerste $k+1$ funktiewaarden te bepalen. Meestal wordt daarvoor gebruik gemaakt van de Taylor-reeksontwikkeling rond het beginpunt $x = a$, soms ook wel van een one-step methode zoals Runge-Kutta.

Om de coëfficiënten γ_m te vinden (die ook direct uit (4.50) berekend kunnen worden) maken wij gebruik van de genererende funktie van γ_m .

Definitie. De genererende funktie van een rij getallen $a_0, a_1, a_2, \dots$ is de funktie $G(t)$ die in een omgeving van $t = 0$ in een machtreeks ontwikkeld kan worden, waarbij de coëfficiënten juist de getallen a_n zijn. Dus

$$G(t) = \sum_{n=0}^{\infty} a_n t^n.$$

We bepalen de genererende funktie van γ_m .

$$\begin{aligned} G(t) &= \sum_{m=0}^{\infty} \gamma_m t^m = \sum_{m=0}^{\infty} (-t)^m \int_0^1 \binom{-s}{m} ds \\ &= \int_0^1 \sum_{m=0}^{\infty} \binom{-s}{m} (-t)^m ds = \int_0^1 (1-t)^{-s} ds \end{aligned}$$

zodat we na berekening van de laatste integraal vinden

$$G(t) = - \frac{t}{(1-t) \cdot \log(1-t)} . \quad (4.51)$$

Dit kunnen we ook als volgt schrijven

$$- \frac{\log(1-t)}{t} \cdot G(t) = \frac{1}{1-t} .$$

Vullen we hierin de reeksontwikkeling van de verschillende funkties in dan krijgen we

$$(1 + \frac{t}{2} + \frac{t^2}{3} + \dots)(\gamma_0 + \gamma_1 t + \gamma_2 t^2 + \dots) = 1 + t + t^2 + \dots$$

Gelijkstelling van de coëfficiënten van de overeenkomstige machten van t geeft de recurrente betrekking

$$\gamma_m + \frac{1}{2}\gamma_{m-1} + \frac{1}{3}\gamma_{m-2} + \dots + \frac{1}{m+1}\gamma_0 = 1 \quad m = 0, 1, 2, \dots \quad (4.52)$$

waarmee γ_m berekend kan worden. Voor de eerste waarden van m krijgen we dan

m	0	1	2	3	4	5	6
γ_m	1	$\frac{1}{2}$	$\frac{5}{12}$	$\frac{3}{8}$	$\frac{251}{720}$	$\frac{95}{288}$	$\frac{19087}{60480}$

Als de berekeningen met de hand worden uitgevoerd, zullen de differenties expliciet bepaald worden. Dit geeft een mogelijke controle op de reken-nauwkeurigheid, nl als $\nabla^k f_n$ groot wordt zal in het algemeen de nauwkeurigheid onvoldoende zijn.

Bij gebruik van een rekenmachine vervangen we de differenties door de funktiewaarden, met behulp van (4.45). Dan wordt (4.49)

$$y_{n+1} - y_n = h \sum_{i=0}^k \beta_{ki} f_{n-i} \quad (4.53)$$

waarin

$$\beta_{ki} = (-1)^i \sum_{m=i}^k \binom{m}{i} \gamma_m \quad i = 0, 1, 2, \dots, k; \quad k = 0, 1, 2, \dots$$

i	0	1	2	3	4	5
β_{0i}	1					
$2\beta_{1i}$	3	-1				
$12\beta_{2i}$	23	-16	5			
$24\beta_{3i}$	55	-59	37	-9		
$720\beta_{4i}$	1901	-2774	2616	-1274	251	
$1440\beta_{5i}$	4227	-7673	9482	-6798	2627	-425

De sterk toenemende grootte van de getallen β_{ki} alsmede de wisselende tekens, maken de Adams-Bashforth methode numeriek onaantrekkelijk.

Adams-Moulton

Hierbij wordt (4.48)

$$y_n - y_{n-1} = \int_{x_{n-1}}^{x_n} p(x) dx = h \sum_{m=0}^k \gamma_m^* \nabla^m f_n \quad (4.54)$$

waarin

$$\gamma_m^* = (-1)^m \frac{1}{h} \int_{x_{n-1}}^{x_n} \binom{-s}{m} dx = (-1)^m \int_{-1}^0 \binom{-s}{m} ds.$$

Voor de genererende funktie $G^*(t)$ van deze coëfficiënten vinden we

$$G^*(t) = - \frac{t}{\log(1-t)}$$

waaruit de recurrente betrekking volgt

$$\gamma_m^* + \frac{1}{2}\gamma_{m-1}^* + \dots + \frac{1}{m+1}\gamma_0^* = \begin{cases} 1 & m = 0 \\ 0 & m = 1, 2, \dots \end{cases}$$

m	0	1	2	3	4	5	6
γ_m^*	1	$-\frac{1}{2}$	$-\frac{1}{12}$	$-\frac{1}{24}$	$-\frac{19}{720}$	$-\frac{3}{160}$	$-\frac{863}{60480}$

We merken op dat de volgende relatie geldt

$$\frac{1}{1-t} \cdot G^*(t) = G(t)$$

waaruit met behulp van de reeksontwikkelingen volgt

$$\gamma_0^* + \gamma_1^* + \dots + \gamma_m^* = \gamma_m \quad (4.55)$$

In tegenstelling tot Adams-Bashforth is y_n in (4.54) in het algemeen niet expliciet in de overige funktiewaarden uit te drukken. We noemen deze methode daarom ook wel een impliciete methode. Veronderstellen we dat $y_{n-1}, y_{n-2}, \dots, y_{n-k}$ bekend zijn dan moeten we y_n oplossen uit de niet-triviale vergelijking (4.54). We doen dit met de volgende iteratie-formule

$$y_n^{(v)} = y_{n-1} + h \sum_{m=0}^k \gamma_m^* \nabla^m f_n^{(v-1)} \quad (4.56)$$

De iteratie verloopt dan als volgt.

Stel dat we als beginschatting $y_n^{(0)}$ hebben, dan berekenen we $f_n^{(0)} = f(x_n, y_n^{(0)})$ en daarmee

$$\nabla f_n^{(0)} = f_n^{(0)} - f_{n-1}, \quad \nabla^2 f_n^{(0)} = \nabla f_n^{(0)} - \nabla f_{n-1} \quad \text{etc.}$$

Deze differenties vullen we in het rechterlid van (4.56) in en berekenen daarmee een nieuwe schatting $y_n^{(1)}$. Met $y_n^{(1)}$ berekenen we $f_n^{(1)}$, en daarmee de differenties $\nabla^m f_n^{(1)}$. Met (4.56) vinden we dan de volgende schatting $y_n^{(2)}$ etc.

Het bewijs dat de zo verkregen rij $y_n^{(0)}, y_n^{(1)}, y_n^{(2)}, \dots$ onder zekere omstandigheden convergeert zal later gegeven worden.

We veronderstellen voorlopig dat het proces inderdaad convergeert.

We kunnen het rekenwerk per iteratiestap nog op de volgende wijze vereenvoudigen. We beschouwen

$$y_n^{(v+1)} - y_n^{(v)} = h \sum_{m=0}^k \gamma_m^* \{ \nabla^m f_n^{(v)} - \nabla^m f_n^{(v-1)} \}.$$

De differenties $\nabla^m f_n^{(v)}$ en $\nabla^m f_n^{(v-1)}$ verschillen alleen in de waarde $f_n^{(v)}$ en $f_n^{(v-1)}$. Daaruit volgt gemakkelijk dat

$$\nabla^m f_n^{(v)} - \nabla^m f_n^{(v-1)} = f_n^{(v)} - f_n^{(v-1)}.$$

Vullen we dit in bovenstaande formule in, dan krijgen we

$$\begin{aligned} y_n^{(v+1)} - y_n^{(v)} &= h \sum_{m=0}^k \gamma_m^* \{ f_n^{(v)} - f_n^{(v-1)} \} \\ &= h \cdot \{ f_n^{(v)} - f_n^{(v-1)} \} \cdot \sum_{m=0}^k \gamma_m^* = h \cdot \{ f_n^{(v)} - f_n^{(v-1)} \} \cdot \gamma_k \end{aligned}$$

en dus kunnen we de iteratieformule (4.56) vervangen door

$$y_n^{(v+1)} = y_n^{(v)} + h \cdot \gamma_k \cdot \{ f_n^{(v)} - f_n^{(v-1)} \}, \quad v = 1, 2, \dots \quad (4.57)$$

Om de mogelijke introductie van afrondingsfouten te elimineren vullen we de laatst gevonden $y_n^{(v)}$ uit (4.57) nog eenmaal in (4.56) in en itereren daarmee nog eenmaal.

Omdat het iteratieproces bij iedere volgende stap terugkomt, loont het de moeite de beginschatting $y_n^{(0)}$ zo nauwkeurig mogelijk te bepalen. Daarmee maken we het aantal benodigde iteraties zo klein mogelijk. De formule waarmee we $y_n^{(0)}$ bepalen noemen we de predictor-formule. Als zodanig kunnen we de expliciete formule van Adams-Bashforth gebruiken. De Adams-Moulton formule noemen we dan de corrector-formule.

Voorbeeld

$$y' = x - y^2 \quad y(0) = 0.$$

Als we deze differentiaalvergelijking willen oplossen met Adams-Moulton, met $k = 2$ en $h = 0,1$ dan krijgen we de volgende formules.

$$\text{Predictor:} \quad y_n = y_{n-1} + h(f_{n-1} + \frac{1}{2} \nabla f_{n-1} + \frac{5}{12} \nabla^2 f_{n-1}) \quad \text{Adams-Bashforth.}$$

$$\text{Corrector:} \quad y_n = y_{n-1} + h(f_n - \frac{1}{2} \nabla f_n - \frac{1}{12} \nabla^2 f_n) \quad \text{Adams-Moulton.}$$

De drie benodigde startwaarden worden uit de reeksontwikkeling van de exacte oplossing verkregen.

$$y(x) = \frac{1}{2} x^2 - \frac{1}{20} x^5 + \frac{1}{160} x^8 \dots$$

Als we in (4.54) de differenties weer uitdrukken in de funktiewaarden dan krijgen we

$$y_n - y_{n-1} = h \sum_{i=0}^k \beta_{ki}^* f_{n-i} \quad (4.58)$$

waarin

$$\beta_{ki}^* = (-1)^i \sum_{m=i}^k \binom{m}{i} \gamma_m^* \quad i = 0, 1, 2, \dots, k; \quad k = 0, 1, 2, \dots$$

i	0	1	2	3	4	5
β_{0i}^*	1					
$2\beta_{1i}^*$	1	1				
$12\beta_{2i}^*$	5	8	-1			
$24\beta_{3i}^*$	9	19	-5	1		
$720\beta_{4i}^*$	251	646	-264	106	-19	
$1440\beta_{5i}^*$	475	1427	-798	482	-173	27

We vergelijken deze tabel met die van blz NA-47, en zien dat de coëfficiënten β_{ki}^* kleiner zijn dan β_{ki} . Dit betekent dat we bij het gebruik van (4.58) minder last zullen hebben van afrondingsfouten dan bij (4.53).

Ook de afbreekfout bij Adams-Moulton is kleiner dan bij Adams-Bashforth. Dit is intuïtief reeds duidelijk want bij Adams-Moulton gebruiken we het interpolatie-polynoom tussen de basispunten, terwijl bij Adams-Bashforth het polynoom buiten de basispunten gebruikt wordt (extrapolatie). In het speciale geval $k = 1$ staat de Adams-Moulton methode bekend als de trapeziumregel.

Nyström

Hier krijgen we voor (4.48)

$$y_{n+1} - y_{n-1} = h \sum_{m=0}^k \kappa_m \cdot \nabla^m f_n \quad (4.59)$$

waarin

$$\kappa_m = (-1)^m \int_{-1}^1 \binom{-s}{m} ds.$$

De genererende funktie van deze coëfficiënten is

$$K(t) = - \frac{t}{\log(1-t)} \cdot \frac{2-t}{1-t}.$$

Hiermee vinden we de recurrente betrekking

$$\kappa_m + \frac{1}{2} \kappa_{m-1} + \dots + \frac{1}{m+1} \kappa_0 = \begin{cases} 2 & m = 0 \\ 0 & m = 1, 2, \dots \end{cases}$$

m	0	1	2	3	4	5	6
κ_m	2	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{29}{90}$	$\frac{14}{45}$	$\frac{1139}{3780}$

Dat $\kappa_1 = 0$ betekent dat de formule (4.59) voor $k = 0$ en $k = 1$ dezelfde is, en dus ook dezelfde nauwkeurigheid heeft.

Nyström is ook een expliciete methode, die hoofdzakelijk gebruikt wordt als predictor bij de Milne-Simpson methode.

Milne-Simpson

In dit geval wordt (4.48)

$$y_n - y_{n-2} = h \sum_{m=0}^k \kappa_m^* \nabla^m f_n \quad (4.60)$$

waarin

$$\kappa_m^* = (-1)^m \int_{-2}^0 \binom{-s}{m} ds.$$

Voor de genererende funktie van κ_m^* vinden we

$$K^*(t) = - \frac{t}{\log(1-t)} \cdot (2-t)$$

waaruit de recurrente betrekking volgt

$$\kappa_m^* + \frac{1}{2}\kappa_{m-1}^* + \dots + \frac{1}{m+1}\kappa_0^* = \begin{cases} 2 & m = 0 \\ -1 & m = 1 \\ 0 & m = 2, 3, \dots \end{cases}$$

m	0	1	2	3	4	5	6
κ_m^*	2	-2	$\frac{1}{3}$	0	$-\frac{1}{90}$	$-\frac{1}{90}$	$-\frac{37}{3780}$

Uit de relatie

$$\frac{1}{1-t} \cdot K^*(t) = K(t)$$

volgt met behulp van de reeksontwikkelingen

$$\kappa_0^* + \kappa_1^* + \dots + \kappa_m^* = \kappa_m.$$

Wegens de speciale numerieke waarden van de coëfficiënten κ_m^* schrijven we (4.60) liever in de vorm

$$y_n - y_{n-2} = h \left[2f_{n-1} + \frac{1}{3}\nabla^2 f_n - \frac{1}{90}(\nabla^4 f_n + \nabla^5 f_n) + \dots \right] \quad (4.61)$$

We merken op dat $k = 3$ tot dezelfde formule aanleiding geeft als $k = 2$, omdat $\kappa_3^* = 0$. Dit betekent dat de formule met $k = 2$ een hogere nauwkeurigheid heeft dan verwacht mocht worden. Dit geval $k = 2$ staat bekend als de methode van Milne. Formule (4.61) uitgeschreven in funktiewaarden wordt dan

$$y_n = y_{n-2} + \frac{1}{3}h(f_n + 4f_{n-1} + f_{n-2}). \quad (4.62)$$

Als f onafhankelijk is van y , is dit de bekende Simpson integratie formule.

De Milne-Simpson methode is een impliciete methode. Als predictor gebruikt men liefst een formule met een afbreekfout van dezelfde orde van grootte als de corrector. Zo'n formule is hier bijvoorbeeld Nyström. Milne zelf had voorgesteld als predictor

$$y_n = y_{n-4} + \frac{1}{3}h(8f_{n-1} - 4f_{n-2} + 8f_{n-3}).$$

Deze ontstaat door integratie van het tweedegraads interpolatiepolynoom op de punten x_{n-1} , x_{n-2} en x_{n-3} over het interval (x_{n-4}, x_n) .

We leiden tenslotte nog een uitdrukking af voor de lokale afbreekfout bij de verschillende behandelde meerstaps methoden. Onder lokale afbreekfout verstaan we de fout die per integratie-stap gemaakt wordt. We gaan daarbij uit van de afbreekfout in de interpolatieformule.

Bij Adams-Bashforth krijgen we de interpolatieformule

$$y'(x) = \sum_{m=0}^k (-1)^m \binom{-s}{m} \nabla^m y'(x_n) + (-1)^{k+1} \binom{-s}{k+1} h^{k+1} y^{(k+2)}(\xi)$$

waaruit door integratie volgt

$$y(x_{n+1}) - y(x_n) = h \sum_{m=0}^k \gamma_m \nabla^m y'(x_n) + R_k^{AB}.$$

Hierin is de restterm

$$R_k^{AB} = (-1)^{k+1} h^{k+1} \int_{x_n}^{x_{n+1}} \binom{-s}{k+1} y^{(k+2)}(\xi) dx.$$

Daar in het interval (x_n, x_{n+1}) de funktie $\binom{-s}{k+1}$ tekenvast en $y^{(k+2)}(\xi)$

continu is mogen we de middelwaardestelling (NA-11) toepassen. Doen we dit dan krijgen we

$$R_k^{AB} = (-1)^{k+1} h^{k+1} y^{(k+2)}(\eta) \int_{x_n}^{x_{n+1}} \left(\frac{-s}{k+1}\right) dx.$$

Wegens de definitie van γ_m (zie blz NA-45) mogen we dit ook schrijven als

$$R_k^{AB} = h^{k+2} y^{(k+2)}(\eta) \gamma_{k+1} \quad (4.63)$$

Op geheel analoge wijze vinden we bij Adams-Moulton

$$y(x_n) - y(x_{n-1}) = h \sum_{m=0}^k \gamma_m^* \nabla^m y'(x_n) + R_k^{AM}$$

waarin voor de restterm geldt

$$R_k^{AM} = h^{k+2} y^{(k+2)}(\eta) \gamma_{k+1}^* \quad (4.64)$$

Uit de formules (4.63) en (4.64) zien we dat de afbreekfout in beide gevallen $O(h^{k+2})$ is. De hogere nauwkeurigheid van Adams-Moulton zit voornamelijk in de factor γ_{k+1}^* tegenover γ_{k+1} bij Adams-Bashforth. Om bij Nyström voor de restterm R_K^{NY} gedefinieerd door

$$y(x_{n+1}) - y(x_{n-1}) = h \sum_{m=0}^k \kappa_m \nabla^m y'(x_n) + R_K^{NY} \quad (4.65)$$

en evenzo bij Milne-Simpson voor R_K^{MS} gedefinieerd door

$$y(x_n) - y(x_{n-2}) = h \sum_{m=0}^k \kappa_m^* \nabla^m y'(x_n) + R_K^{MS} \quad (4.66)$$

een soortgelijke formule af te leiden als (4.63) en (4.64), kunnen we geen gebruik maken van de middelwaardestelling, omdat $\left(\frac{-s}{k+1}\right)$ in deze gevallen niet tekenvast is in het integratie-interval. De afleiding in het algemene geval is daardoor veel gecompliceerder. We beperken ons hier tot twee speciale gevallen.

Voor Nyström nemen we het geval $k = 0$ (dus ook $k = 1$).
 Voor het gemak nemen we aan dat $x_{n-1} = -h$, $x_n = 0$, $x_{n+1} = h$. Vullen we
 vervolgens $k = 0$ in (4.65) in dan krijgen we

$$R_0^{NY} = y(h) - y(-h) - 2h y'(0).$$

Met de Taylor-ontwikkeling voor $y(x)$ vinden we

$$y(h) = y(0) + hy'(0) + \frac{h^2}{2} y''(0) + \frac{h^3}{6} y'''(\xi_1) \quad 0 < \xi_1 < h,$$

$$y(-h) = y(0) - hy'(0) + \frac{h^2}{2} y''(0) - \frac{h^3}{6} y'''(\xi_2) \quad -h < \xi_2 < 0,$$

en dus

$$R_0^{NY} = \frac{h^3}{6} (y'''(\xi_1) + y'''(\xi_2)).$$

Omdat $y'''(x)$ continu is en dus in het interval (ξ_1, ξ_2) alle waarden tussen $y'''(\xi_1)$ en $y'''(\xi_2)$ aanneemt is er een punt ξ in (ξ_1, ξ_2) waarvoor geldt

$$y'''(\xi) = \frac{1}{2} (y'''(\xi_1) + y'''(\xi_2)).$$

Vullen we dit in de vorige formule in dan krijgen we

$$R_0^{NY} = \frac{1}{3} h^3 y'''(\xi) \quad x_{n-1} < \xi < x_{n+1}. \quad (4.67)$$

Bij Milne-Simpson bepalen we de restterm in het speciale geval $k = 2$
 (dus ook $k = 3$), de methode van Milne.

We kiezen nu $x_{n-2} = -h$, $x_{n-1} = 0$, $x_{n+1} = h$. Vervolgens vullen we $k = 2$
 in (4.66) in en krijgen dan

$$R_2^{MS} = y(h) - y(-h) - h[2y'(0) + \frac{1}{3} \nabla^2 y'(h)].$$

Het is gemakkelijk te verifiëren dat de volgende formule geldt

$$R_2^{MS} = \int_0^h [y'(t) + y'(-t) - 2y'(0) - \frac{t^2}{h^2} \nabla^2 y'(h)] dt.$$

We introduceren nu de functie

$$F(x) = \int_0^x \left[y'(t) + y'(-t) - 2y'(0) - \frac{t^2}{h^2} \nabla^2 y'(h) \right] dt \\ - \lambda \int_0^x \left[\binom{1 - \frac{t}{h}}{4} + \binom{1 + \frac{t}{h}}{4} \right] dt.$$

Dan geldt $F(0) = 0$, $F(-x) = -F(x)$. We kiezen λ zo dat ook $F(h) = F(-h) = 0$. Dit is mogelijk omdat

$$\int_0^h \left[\binom{1 - \frac{t}{h}}{4} + \binom{1 + \frac{t}{h}}{4} \right] dt = h \int_{-2}^0 \binom{-s}{4} ds = h \kappa_4^* = -\frac{1}{90} h \neq 0.$$

Met deze keuze van λ krijgen we

$$F(h) = R_2^{MS} + \frac{1}{90} \lambda h = 0.$$

We bepalen nu nog de waarde van λ . Omdat $F(\pm h) = F(0) = 0$ is er een ζ tussen 0 en h zodanig dat $F'(\pm \zeta) = 0$ (Rolle). Verder is

$$F'(x) = y'(x) + y'(-x) - 2y'(0) - \frac{x^2}{h^2} \nabla^2 y'(h) - \lambda \left[\binom{1 - \frac{x}{h}}{4} + \binom{1 + \frac{x}{h}}{4} \right]$$

waaruit men, door invullen, kan vinden dat $F'(0) = F'(\pm h) = 0$. Dus $F'(x)$ heeft vijf verschillende nulpunten in $(-h, h)$. Met Rolle vinden we dan op de bekende manier dat de vijfde afgeleide $F^{(5)}(x)$ minstens één nulpunt, zeg η , heeft. Daar

$$F^{(5)}(x) = y^{(5)}(x) + y^{(5)}(-x) - 2\lambda h^{-4}$$

vinden we voor λ de uitdrukking

$$\lambda = \frac{1}{2} h^4 [y^{(5)}(\eta) + y^{(5)}(-\eta)].$$

Uit de continuïteit van $y^{(5)}(x)$ volgt op dezelfde manier als hierboven dat er een punt ξ bestaat zodanig dat

$$\lambda = h^4 y^{(5)}(\xi)$$

zodat we tenslotte de formule krijgen

$$R_2^{MS} = -\frac{1}{90} h^5 y^{(5)}(\xi) \quad x_{n-2} < \xi < x_n. \quad (4.68)$$

We geven nu nog het op blz NA-49 beloofde bewijs van de convergentie van het iteratieproces bij Adams-Moulton en Milne-Simpson.

De algemene vorm van een corrector-formule is

$$y_n = \alpha h y'_n + \Phi_{n-1}. \quad (4.69)$$

Hierin is $y'_n = f(x_n, y_n)$, terwijl Φ_{n-1} een funktie is die alleen afhangt van de reeds bekende waarden $y_{n-1}, y_{n-2}, \dots$. De iteratieformule is

$$y_n^{(\nu+1)} = \alpha h f(x_n, y_n^{(\nu)}) + \Phi_{n-1}.$$

Als we veronderstellen dat y_n de gezochte oplossing van (4.69) is dan is

$$y_n - y_n^{(\nu+1)} = \alpha h [f(x_n, y_n) - f(x_n, y_n^{(\nu)})].$$

Omdat we geëist hebben dat $f(x, y)$ aan de Lipschitzvoorwaarde voldoet, d.w.z.

$$|f(x_n, y_n) - f(x_n, y_n^{(\nu)})| \leq L |y_n - y_n^{(\nu)}|$$

geldt

$$|y_n - y_n^{(\nu+1)}| \leq |\alpha| \cdot h \cdot L \cdot |y_n - y_n^{(\nu)}|.$$

Hieruit volgt dat de rij $y_n^{(0)}, y_n^{(1)}, y_n^{(2)}, \dots$ convergeert met limiet y_n als

$$|\alpha| \cdot h \cdot L < 1.$$

Het iteratieproces convergeert dus altijd als h voldoende klein is, nl. als $h < \frac{1}{|\alpha|L}$.

Om in de praktijk het iteratieproces, bij het bereiken van de gewenste nauwkeurigheid, te beëindigen hebben we een schatting voor $|y_n - y_n^{(v)}|$ nodig. We vinden deze als volgt.

$$|y_n - y_n^{(v)}| \leq |y_n - y_n^{(v+1)}| + |y_n^{(v+1)} - y_n^{(v)}|$$

$$|y_n - y_n^{(v)}| \leq |\alpha| \cdot h \cdot L \cdot |y_n - y_n^{(v)}| + |\alpha| \cdot h \cdot |f(x_n, y_n^{(v)}) - f(x_n, y_n^{(v-1)})|.$$

Dus

$$|y_n - y_n^{(v)}| \leq \frac{|\alpha| h}{1 - |\alpha| h L} \cdot |f(x_n, y_n^{(v)}) - f(x_n, y_n^{(v-1)})|.$$

Het rechterlid van deze ongelijkheid kan gemakkelijk berekend worden, vooral als we, zoals dikwijls gebeurt, veronderstellen dat h zo klein is dat we $1 - |\alpha| h L \approx 1$ mogen nemen.

5. Stabiliteit

In de vorige paragraaf hebben we formules afgeleid, die ons in staat stelden een beginwaarde-probleem stap voor stap op te lossen. Bij het gebruik van deze formules maken we per stap een fout, die we lokale afbreekfout genoemd hebben, en die dikwijls van de vorm $\alpha_n h^n f^{(n)}(\xi)$ is. Nu hebben we bij een willekeurige integratiestap niet alleen met de lokale afbreekfout te maken maar ook met de voortplanting van de reeds eerder gemaakte fouten. Om iets te kunnen zeggen over de nauwkeurigheid van de methode is het daarom belangrijk te weten wat het effect bij verdere integratiestappen is van een eenmaal geïntroduceerde fout. Als de fouten de neiging vertonen om sterk aan te groeien spreken we van instabiliteit.

We geven hier geen definitie van het begrip stabiliteit, maar we zullen in een aantal gevallen aangeven wat we onder stabiliteit of instabiliteit zullen verstaan. Verder zullen we naast de stabiliteit ook de convergentie van de methode onderzoeken.

We veronderstellen dat het beginwaarde-probleem bestaat uit een m-de orde lineaire differentiaalvergelijking

$$p_0 y^{(m)} + p_1 y^{(m-1)} + \dots + p_m y = q \quad (5.1)$$

met de beginvoorwaarden

$$y^{(i)}(0) = \eta_i, \quad i = 0, 1, 2, \dots, m-1. \quad (5.2)$$

De coëfficiënten van (5.1) mogen functies van x zijn.
De algemene oplossing van (5.1) is

$$y(x) = \sum_{s=1}^m \lambda_s y_s(x) + y_0(x) \quad (5.3)$$

waarbij $y_1(x), y_2(x), \dots, y_m(x)$ m onafhankelijke oplossingen zijn van de homogene vergelijking, (5.1) met $q = 0$, en $y_0(x)$ een particuliere oplossing is van de inhomogene vergelijking. Bij een speciale keuze van de coëfficiënten λ_s zal $y(x)$ ook aan de beginvoorwaarden (5.2) voldoen.

We veronderstellen nu dat de numerieke methode waarmee we (5.1) willen oplossen bestaat uit de differentievergelijking

$$a_0 y_n + a_1 y_{n-1} + \dots + a_k y_{n-k} = b_n \quad n = k, k+1, \dots \quad (5.4)$$

Verder nemen we aan dat, eventueel met een startprocedure, de waarden y_i $i = 0, 1, \dots, k-1$ zeer nauwkeurig, desnoods exact, bekend zijn.
De algemene oplossing van deze k-de orde lineaire differentievergelijking is

$$y_n = \sum_{s=1}^k \Lambda_s Y_n^{(s)} + Y_n^{(0)}. \quad (5.5)$$

Hierin zijn $Y_n^{(1)}, Y_n^{(2)}, \dots, Y_n^{(k)}$ k onafhankelijke oplossingen van de homogene vergelijking, (5.4) met $b_n = 0$ voor alle n , en $Y_n^{(0)}$ een particuliere oplossing van de inhomogene vergelijking. Om aan de startwaarden te voldoen moeten voor de coëfficiënten Λ_s speciale waarden gekozen worden.

Bij iedere oplossing (5.5) van de differentievergelijking hoort een stuksgewijs lineaire funktie $y(x,h)$ zodat $y(x_n,h) = y_n$. We noemen $y(x,h)$ een benaderende oplossing van de differentiaalvergelijking. De bij $y_n^{(s)}$ behorende funktie schrijven we als $Y_s(x,h)$. Dan is $Y_s(x_n,h) = y_n^{(s)}$ en geldt de relatie

$$y(x,h) = \sum_{s=1}^k \Lambda_s Y_s(x,h) + Y_0(x,h) \quad (5.6)$$

We noemen nu een numerieke methode convergent als voor $h \rightarrow 0$ de oplossing $y(x,h)$ van de differentievergelijking nadert tot de exacte oplossing $y(x)$ van de differentiaalvergelijking.

We noemen een numerieke methode stabiel als bij vaste h het asymptotisch gedrag na een groot aantal stappen, van de oplossing van de differentievergelijking overeenkomt met het asymptotisch gedrag van de oplossing van de differentiaalvergelijking.

In het geval $k = m$ mogen we verwachten dat bij iedere $Y_s(x,h)$ uit (5.6) een $y_s(x)$ uit (5.3) behoort waarvoor geldt dat $Y_s(x,h) \rightarrow y_s(x)$ als $h \rightarrow 0$. In het geval $k > m$, dat dikwijls zal voorkomen bij predictor-corrector methoden zullen er m oplossingen $Y_s(x,h) \rightarrow y_s(x)$ als $h \rightarrow 0$; de overige $k-m$ funkties $Y_s(x,h)$ hebben geen bijbehorende oplossing van de differentiaalvergelijking. We onderscheiden nu de volgende soorten instabiliteiten.

Inherente instabiliteit

Deze is geen gevolg van de gekozen numerieke methode, maar is inherent aan het probleem. Veronderstel namelijk dat de oplossing (5.3) een dominante term bevat, zeg $y_1(x)$ en dat deze onderdrukt is, omdat door de beginvoorwaarden $\lambda_1 = 0$ is. De corresponderende dominante term $Y_1(x,h)$ van (5.6) zal o.a. door afrondings- en afbreekfouten met een (mogelijk kleine) coëfficiënt $\Lambda_1 \neq 0$ voorkomen. Eenmaal geïntroduceerd zal deze term spoedig een overheersende invloed hebben op de oplossing $y(x,h)$. Het probleem zelf is hier instabiel, omdat een kleine verandering van de beginvoorwaarden een grote verandering in de oplossing geeft. Het is een analoog effect als bij een ill-conditioned stelsel lineaire vergelijkingen.

Partiele instabiliteit

We beschouwen het geval $k = m$, en we veronderstellen dat iedere $Y_s(x,h) \rightarrow y_s(x)$ als $h \rightarrow 0$. Dan convergeert de methode. Dit behoeft echter niet te betekenen dat het asymptotisch gedrag van $Y_s(x,h)$ en

$y_s(x)$ hetzelfde zijn. Als met name $Y_1(x, h)$ bij vaste h , zich geheel anders gedraagt dan $y_1(x)$ voor grote waarden van x , dan is de methode instabiel. In het algemeen bestaat er in deze gevallen een grenswaarde H_0 voor h , zodat we instabiliteit hebben als $h > H_0$ en stabiliteit als $h < H_0$. We noemen dit verschijnsel daarom partiele instabiliteit. Met deze definitie impliceert partiele instabiliteit dus convergentie.

Sterke instabiliteit

We beschouwen nu het geval $k > m$, en veronderstellen dat bij iedere $y_s(x)$ een $Y_s(x, h)$ bestaat zodat $Y_s(x, h) \rightarrow y_s(x)$ als $h \rightarrow 0$. Dan hebben de overige oplossingen $Y_s(x, h)$ geen overeenkomstige oplossing van de differentiaalvergelijking. Deze $Y_s(x, h)$ noemen we parasitaire oplossingen van de differentie-vergelijking. Het zijn functies die vreemd zijn aan het beginwaardeprobleem, maar door de speciale keuze van de numerieke methode geïntroduceerd zijn. Als de parasitaire $Y_s(x, h) \rightarrow 0$ voor $h \rightarrow 0$ dan hebben we convergentie.

Als bij vaste h het asymptotisch gedrag van de parasitaire oplossingen geen invloed heeft op dat van de echte oplossing, dan is de methode stabiel. Het komt echter dikwijls voor dat de parasitaire oplossingen niet verdwijnen als $h \rightarrow 0$ en daarbij voor vaste h sterker groeien dan een willekeurige oplossing van de differentiaalvergelijking als $x \rightarrow \infty$. Dit verschijnsel noemen we sterke instabiliteit; het impliceert dus geen convergentie en geen stabiliteit. Als we wel convergentie hebben maar geen stabiel asymptotisch gedrag noemen we de methode zwak instabiel.

Enkele voorbeelden

Inherente instabiliteit

De vergelijking

$$y' = y - x$$

heeft de algemene oplossing

$$y(x) = \lambda e^x + x + 1$$

waarbij λ een willekeurige constante is, die bepaald wordt door de beginvoorwaarde. Kiezen we speciaal $y(0) = 1$ dan is $\lambda = 0$ en de oplossing wordt dan $y = x + 1$, een functie die in vergelijking met e^x slechts langzaam groeit. Wordt daarom de term λe^x door de onnauwkeurige rekenwijze

geïntroduceerd dan zal deze spoedig de werkelijke oplossing verdringen. Dezelfde moeilijkheden krijgen we bij de tweede-orde differentiaalvergelijking

$$y'' = f(x) \cdot y$$

als $f(x) > 0$. Speciaal als $f(x) \rightarrow k^2$ voor $x \rightarrow \infty$ zal het asymptotisch gedrag van $y(x)$ gelijk zijn aan dat van $\lambda_1 e^{kx} + \lambda_2 e^{-kx}$. Daarvan zal de eerste term, zelfs als deze onderdrukt is door de beginvoorwaarden, uiteindelijk domineren bij iedere numerieke methode.

Partiele instabiliteit

De partiele instabiliteit komt vaak voor bij éénstapsmethoden zoals Runge-Kutta, of de Taylor methode. Het kan de aanleiding zijn om in plaats van een one-step methode een multistep methode te kiezen. We beschouwen de vergelijking

$$y' = -10y + 9 - 10x, \quad y(0) = 1.$$

De algemen oplossing hiervan is

$$y = \lambda e^{-10x} - x + 1$$

en de beginvoorwaarde is zo dat $\lambda = 0$ genomen moet worden. Als oplossingsmethode nemen we Taylor die we na de tweede afgeleide afbreken, dus

$$y_{n+1} = y_n + hy'_n + \frac{h^2}{2} y''_n.$$

Invullen geeft de recurrente betrekking

$$y_{n+1} = (1 - 10h + 50h^2)y_n + h(9 - 10x_n) + 50h^2(x_n - 1).$$

De hierbij behorende homogene differentievergelijking

$$y_{n+1} = (1 - 10h + 50h^2)y_n$$

heeft als oplossing

$$y_n^{(1)} = t^n \quad \text{waarbij } t = 1 - 10h + 50h^2.$$

Voor vaste x geldt dan dat $Y_q(x, h) = (1 - 10h + 50h^2)^{\frac{x}{h}}$ nadert tot $y_1(x) = e^{-10x}$ als $h \rightarrow 0$, en dus convergeert de methode.

Als echter $h > 0.2$ dan is $t > 1$ en dan zal $Y_q(x, h) \rightarrow \infty$ als $x \rightarrow \infty$, terwijl de overeenkomstige $y_1(x) \rightarrow 0$. De asymptotische gedragingen zijn voor $h > 0.2$ dus geheel verschillend. We zien dat de grenswaarde $H_0 = 0.2$. Voor een kleinere waarde $h = 0.15$ krijgen we $t = \frac{5}{8}$ en dus $Y_n^{(1)} = \left(\frac{5}{8}\right)^n$. Dus geldt dat evenals $y_1(x)$ in dit geval ook $Y_q(x, h)$, hoewel langzamer, naar nul nadert als $x \rightarrow \infty$.

Passen we op dit voorbeeld de 4^e orde Runge-Kutta formule toe dan krijgen we de formule

$$y_{n+1} = (1 - 10h + 50h^2 - \frac{500}{3}h^3 + \frac{1250}{3}h^4)y_n + \dots$$

De term tussen haakjes is groter dan 1 als $h > 0.27$. Dus is voor stabiele integratie vereist dat $h < 0.27$.

Door toepassing van een andere methode kunnen we de partiële instabiliteit vermijden. Gebruiken we bijvoorbeeld de trapeziumregel

$$y_{n+1} = y_n + \frac{h}{2}(y'_{n+1} + y'_n)$$

dan krijgen we in ons voorbeeld de differentieformule

$$(1 + 5h)y_{n+1} = (1 - 5h)y_n + h(9 - 5(x_n + x_{n+1})).$$

De bijbehorende homogene vergelijking is dan

$$y_{n+1} = \frac{1 - 5h}{1 + 5h} y_n$$

en dus hebben we onvoorwaardelijke stabiliteit, omdat $\frac{1 - 5h}{1 + 5h} < 1$ voor alle h .

Sterke instabiliteit

Predictor-corrector methoden kunnen last hebben van de meer katastrofale sterke instabiliteit, door de introductie van parasitaire oplossingen. We nemen als voorbeeld de differentievergelijking

$$11y_{n+1} = -27y_n + 27y_{n-1} + 11y_{n-2} + 3h(y'_{n+1} + 9y'_n + 9y'_{n-1} + y'_{n-2}) \quad (5.7)$$

Als de oplossing van de differentiaalvergelijking $y(x)$ een polynoom is van de graad ≤ 6 dan geeft (5.7) deze oplossing exact. De formule is dus zeer nauwkeurig in die zin dat de afbreekfout zeer klein is, namelijk $O(h^7)$.

Passen we (5.7) toe op de vergelijking $y' = \lambda y$ dan krijgen we de recurrente relatie

$$(11 - 3\lambda h)y_{n+1} + (27 - 27\lambda h)y_n - (27 + 27\lambda h)y_{n-1} - (11 + 3\lambda h)y_{n-2} = 0. \quad (5.8)$$

Intermezzo.

De methode van oplossen van een lineaire differentievergelijking

$$a_0 y_n + a_1 y_{n-1} + \dots + a_k y_{n-k} = b_n, \quad n = k, k+1, \dots \quad (5.9)$$

vertoont veel overeenkomst met de oplossingsmethode voor een lineaire differentiaalvergelijking.

Men kan eenvoudig bewijzen, op precies dezelfde manier als bij de differentiaalvergelijking dat de oplossingen van de homogene differentievergelijking, (5.9) met $b_n = 0$ voor alle n , een lineaire ruimte vor-

men bestaande uit alle combinaties van k onafhankelijke oplossingen $y_n^{(s)}$ en dat de algemene oplossing van de inhomogene vergelijking bestaat uit één particuliere oplossing van de inhomogene vergelijking $y_n^{(0)}$ plus de algemene oplossing van de homogene vergelijking. In formule

$$y_n = y_n^{(0)} + \sum_{s=1}^k \Lambda_s y_n^{(s)} \quad (5.10)$$

Opmerking: van dit resultaat wordt op blz (NA-59) ook gebruik gemaakt. Als de getallen $a_0, a_1, \dots, a_k$ onafhankelijk zijn van n dan hebben we een lineaire differentievergelijking met constante coëfficiënten. In dat geval kunnen we de onafhankelijke oplossingen $y_n^{(s)}$ eenvoudig vinden door te stellen $y_n^{(s)} = p^n$, (bij de differentiaalvergelijking stellen we in het analoge geval $y(x) = e^{px}$). Vullen we dit in (5.9) in dan vinden we dat p moet voldoen aan de karakteristieke vergelijking

$$a_0 p^k + a_1 p^{k-1} + \dots + a_k = 0 \quad (5.11)$$

Als de wortels p_s van deze vergelijking onderling verschillend zijn dan vinden we zo alle k onafhankelijke oplossingen $y_n^{(s)} = p_s^n$, $s = 1, 2, \dots, k$. Is p_1 tweevoudige wortel van (5.11) dan blijkt (door invullen) dat ook np_1^n een oplossing van (5.9) is. Is p_1 een drievoudige wortel dan zijn $n(n-1)p_1^n$ en np_1^n ook oplossingen, etc. Ook dit is analoog aan de differentiaalvergelijking waar bij een tweevoudige wortel p van de karakteristieke vergelijking twee oplossingen e^{px} en xe^{px} behoren, etc.

Formule (5.8) is een homogene lineaire differentievergelijking met constante coëfficiënten. De algemene oplossing daarvan is dan ook

$$y_n = A_1 p_1^n + A_2 p_2^n + A_3 p_3^n$$

waarbij p_1 , p_2 en p_3 de wortels zijn van de karakteristieke vergelijking

$$(11 - 3\lambda h)p^3 + (27 - 27\lambda h)p^2 - (27 + 27\lambda h)p - (11 + 3\lambda h) = 0.$$

Men kan uitrekenen dat deze vergelijking voor iedere waarde van λh een wortel $p_1 < -1.5$ heeft. De oplossing $y_n^{(1)} = p_1^n$ is dan ook een parasitaire oplossing met een snelgroeïende oscillerende waarde. We hebben in dit geval dus sterke instabiliteit. We merken nog op dat het verkleinen van h geen verbetering oplevert voor wat betreft de stabiliteit. De parasitaire oplossingen bestaan en een ervan groeit aan voor alle h . We zien dat (5.7) ondanks de kleine fout die we per stap maken erg on aantrekkelijk is juist vanwege de sterke instabiliteit van de methode.

We beschouwen nu nog algemeen de multistep methode waarvan de formule luidt

$$\alpha_0 y_n + \alpha_1 y_{n-1} + \dots + \alpha_k y_{n-k} = h(\beta_0 y_n' + \dots + \beta_k y_{n-k}'). \quad (5.12)$$

We veronderstellen dat $\alpha_0 \neq 0$, en $|\alpha_s| + |\beta_s| > 0$ voor $s = 0, 1, \dots, k$. We noemen k de orde van de methode; het is het aantal voorgaande punten waarmee het volgende punt berekend wordt. De graad p van (5.12) is de maximale graad van het polynoom waarvoor aan (5.12) exact voldaan is. De afbreekfout heeft dan de gedaante $K h^{p+1} f^{(p+1)}(\xi)$. Bij de differentieformule (5.12) definieren we twee polynomen

$$\begin{aligned} \rho(z) &= \alpha_0 z^k + \dots + \alpha_k \\ \sigma(z) &= \beta_0 z^k + \dots + \beta_k \end{aligned}$$

We veronderstellen nu dat $\rho(z)$ en $\sigma(z)$ geen gemeenschappelijke factor hebben en dat $p \geq 1$, en geven enkele resultaten welke afkomstig zijn van G. Dahlquist.

Stelling

De multistep methode (5.12) is dan en alleen dan convergent als alle nulpunten van het polynoom $\rho(z)$ op of binnen de eenheidscirkel liggen en de wortels die op de eenheidscirkel liggen enkelvoudig zijn.

Nu bevat (5.12) $2k+1$ onafhankelijke constanten α_s en β_s . We kunnen deze constanten zo kiezen dat de graad van (5.12) gelijk is aan $2k$. Er geldt echter de volgende stelling.

Stelling

De maximale graad p waarvoor aan de noodzakelijke voorwaarden voor convergentie voldaan is, is $p = k + 2$ voor even k , en $p = k + 1$ voor oneven k . Als $p = k + 2$ bij even k moet tevens $\alpha_s = -\alpha_{n-s}$ en

$\beta_s = \beta_{n-s}$. In dat geval liggen alle nulpunten van $\rho(z)$ op de eenheidscirkel.

De formule met de kleinste afbreekfout is dus niet de best mogelijke formule wat betreft de stabiliteit.

In het geval van een convergente methode is het nog wel mogelijk dat het polynoom $\rho(z) - \lambda h \sigma(z)$ dat behoort bij de differentiaalvergelijking $y' = \lambda y$ nulpunten buiten de eenheidscirkel heeft voor positieve waarden van h . De hieruit voortvloeiende instabiliteit is minder erg dan de sterke instabiliteit en wordt door Dahlquist zwakke instabiliteit genoemd.

Zo is bijvoorbeeld (5.7) sterk instabiel want $k = 3$ en $p = 6$.

De Milne-Simpson formule

$$y_{n+1} = y_{n-1} + \frac{h}{3}(y'_{n+1} + 4y'_n + y'_{n-1})$$

heeft $k = 2$ en $p = 4$. De polynomen zijn $\rho(z) = z^2 - 1$ en $\sigma(z) = \frac{1}{3}(z^2 + 4z + 1)$. De methode is wel convergent, de nulpunten van $\rho(z)$ zijn ± 1 . De nulpunten van $\rho(z) - \lambda \sigma(z)$ zijn

$$p_1 = e^{\lambda h}(1 + O(h^5)) \quad \text{en} \quad p_2 = -e^{-\frac{1}{3}\lambda h}(1 + O(h^3)).$$

p_1 correspondeert met de exacte oplossing van $y' = \lambda y$ terwijl p_2 parasitair is. Voor $\lambda > 0$ doet deze parasitaire oplossing niet terzake, maar geeft instabiliteit voor negatieve λ . Dus Milne-Simpson is zwak instabiel.

In het geval van een niet-lineaire differentiaalvergelijking $y' = f(x, y)$ is het veel moeilijker om na te gaan of de gebruikte methode stabiel is. We kunnen alleen iets zeggen als we de oplossing lokaal bekijken. In een klein gebied vervangen we de differentiaalvergelijking door de lineaire benadering

$$y' = f(x_n, y_n) + (x - x_0) \frac{\partial f}{\partial x} + (y - y_0) \frac{\partial f}{\partial y}.$$

De hierbij behorende homogene vergelijking is

$$y' = \lambda y \quad \text{met} \quad \lambda = \frac{\partial f}{\partial y}.$$

Een schatting van de waarde van $\frac{\partial f}{\partial y}$ kan ons iets zeggen over de lokale stabiliteit van de methode. Omdat $\frac{\partial f}{\partial y}$ in het algemeen niet constant is, is het mogelijk dat een methode in het ene gebied stabiel is, en in een ander instabiel. Verder is het niet duidelijk hoe de niet lineaire termen die we hebben weggelaten het gedrag van de oplossing zullen beïnvloeden.

Opmerking

Bij het kiezen van een rekenmethode is het niet nodig dat we ons beperken tot de formules die onder alle omstandigheden stabiel zijn. Alleen sterke instabiliteit moeten we altijd vermijden. Verder zullen alleen instabiliteiten die aanleiding geven tot grote relatieve fouten in de oplossing gevaarlijk zijn. Dit zal speciaal het geval zijn als de werkelijke oplossing een afnemende functie is.

Litteratuur:

- | | |
|-----------------|---|
| G.Dahlquist; | Proceedings of Symposia in Applied Mathematics
Volume XV. [1963]. |
| L.Fox; | Numerical Solution of Ordinary and Partial Differential Equations. Pergamon Press [1962]. |
| P.Henrici; | Discrete Variable Methods in Ordinary Differential Equations. John Wiley & Sons, Inc. [1962]. |
| F.B.Hildebrand; | Introduction to Numerical Analysis.
Mc Graw-Hill Book Company, Inc. [1956]. |
| E.Stiefel; | Einführung in die Numerische Mathematik.
B.G.Teubner Verlagsgesellschaft. [1961]. |