Numerieke methoden voor stelsels gewone differentiaalvergelijkingen

Prof. Dr. Marnix Van Daele

Deel II

Lineaire Meerstapsmethoden

Hoofdstuk 4

Lineaire meerstapsmethoden

4.1 Definities

In paragraaf 2.1 hebben we de LMMn ingevoerd. Voor deze methoden wordt het rechterlid, $h\,\phi_f(y_{n+k}, y_{n+k-1}, \ldots, y_n, x_n; h)$, een lineaire combinatie van functiewaarden f geëvalueerd in de punten (x_{n+j}, y_{n+j}), $j = 0, 1, \ldots, k$. Door gebruik te maken van het kortschrift

$$f_{n+j} := f(x_{n+j}, y_{n+j}), \quad j = 0, 1, \ldots, k,$$

kunnen we de LMMn schrijven in de standaardvorm :

Definitie 4.1.1 *Een lineaire k-stapsmethode voor $y' = f(x, y)$ is van de vorm*

$$\sum_{j=0}^{k} \alpha_j\, y_{n+j} = h \sum_{j=0}^{k} \beta_j\, f_{n+j}\,, \tag{4.1}$$

waarbij $\alpha_k = 1$ *en* $|\alpha_0| + |\beta_0| \neq 0\,.$ □

De voorwaarde $\alpha_k = 1$ heft de willekeur op die ontstaat uit het feit dat we beide zijden van (4.1) kunnen vermenigvuldigen met dezelfde constante zonder de methode te wijzigen. De tweede voorwaarde sluit uit dat α_0 en β_0 beide nul worden; derhalve worden methoden van de vorm

$$y_{n+2} - y_{n+1} = h\, f_{n+1}$$

vermeden; die methode is essentieel een één-stapsmethode en is in de praktijk niet onderscheidbaar van

$$y_{n+1} - y_n = h\, f_n\,, \tag{4.2}$$

die bekend staat als *de regel van Euler*, de eenvoudigste van alle numerieke methoden.

Er bestaat een alternatieve notatie voor LMMn. We introduceerden reeds de eerste karakteristieke veelterm ρ geassocieerd aan de methode (2.2) als een veelterm van graad k met coëfficiënten α_j. Voor LMMn lijkt het natuurlijk een gelijkaardige veelterm te definiëren met coëfficiënten β_j :

Definitie 4.1.2 *De eerste en tweede karakteristieke veeltermen $\rho(\xi)$ en $\sigma(\xi)$ geassocieerd aan (4.1) zijn*

$$\rho(\xi) := \sum_{j=0}^{k} \alpha_j \, \xi^j, \quad \sigma(\xi) := \sum_{j=0}^{k} \beta_j \, \xi^j \tag{4.3}$$

met $\xi \in \mathbb{C}$. □

Opmerking 4.1.1

Soms is het voor berekeningen handig gebruik te maken van operatoren. Zo kan de LMM (4.1) geschreven worden als

$$\rho(E) \, y_n = h \, \sigma(E) \, f_n \,, \tag{4.4}$$

waarbij E de verschuivingsoperator voorstelt, die gedefinieerd is door $E \, f(x) := f(x+h)$. □

De voorwaarden die opgelegd worden aan α_k, α_0 en β_0 impliceren dat ρ een monische veelterm is van graad k en dat ρ en σ geen gemeenschappelijke factor ξ bezitten.

De methode (4.1) is expliciet als $\beta_k = 0$ en impliciet als $\beta_k \neq 0$, m.a.w. (4.4) is impliciet als σ van graad k is en expliciet als die graad kleiner is dan k.

Als voldoende startwaarden beschikbaar gesteld zijn, kan de rij $\{y_n\}$ bij een expliciete methode rechtstreeks berekend worden. Voor een impliciete methode is het noodzakelijk bij elke stap een stelsel (van over het algemeen niet-lineaire) vergelijkingen op te lossen :

$$y_{n+k} = h \, \beta_k \, f(x_{n+k}, \, y_{n+k}) + g \,, \tag{4.5}$$

waarbij g een bekende functie is van vooraf berekende waarden van y_{n+j}, nl.

$$g = h \, [\beta_{k-1} \, f_{n+k-1} + \ldots + \beta_0 \, f_n] - \alpha_{k-1} \, y_{n+k-1} - \ldots - \alpha_0 \, y_n \,.$$

De meest eenvoudige methode is vastpunt-iteratie, d.w.z. we gebruiken het algoritme

$$y_{n+k}^{[\nu+1]} = h \, \beta_k f(x_{n+k}, \, y_{n+k}^{[\nu]}) + g \,, \tag{4.6}$$

waarbij $y_{n+k}^{[0]}$ een zekere aanvangsapproximatie is voor $y(x_{n+k})$. Het probleem hierbij is echter dat dit schema pas convergeert als $|h \, \beta_k \, L| < 1$, waarbij L de Lipschitz-constante is van f. Dit leidt soms tot onbruikbaar kleine stapwaarden h. In dat geval wordt overgestapt op pseudo-Newton-iteratie (zie Appendix A).

4.2 Enkele families

De algemene klasse (4.1) van LMMn bevat enkele heel bekende subklassen.

4.2.1 Benadering van $f(x,\, y)$ door een interpolatieveelterm

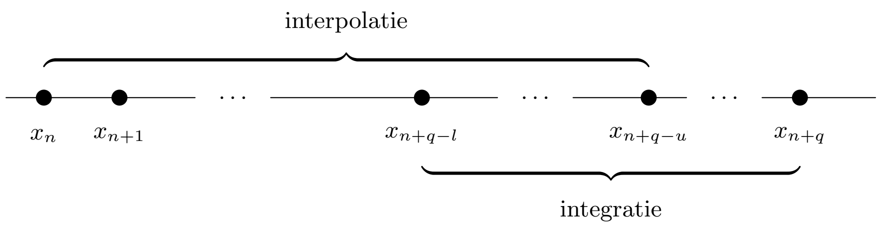
Integreren we het IVP (2.1) tussen x_{n+q-l} en x_{n+q} met $l > 0$, dan ontstaat de identiteit

$$y(x_{n+q}) - y(x_{n+q-l}) = \int\limits_{x_{n+q-l}}^{x_{n+q}} f(x,\, y)\, \mathrm{d}x\,. \tag{4.7}$$

We benaderen $f(x,\, y)$ door een veelterm (uitgedrukt in termen van achterwaartse differenties) van graad $q-u$ met $0 \le u \le q$ vanaf het punt $(x_{n+q-u},\, y(x_{n+q-u}))$, m.a.w.

$$f(x,\, y) \approx \sum_{i=0}^{q-u} (-1)^i \binom{-t}{i} \nabla^i f_{n+q-u}\,, \qquad t = \frac{x - x_{n+q-u}}{h}\,.$$

Het stapgetal van deze methode is dus $k = \max(q,\, l)$.



Figuur 4.1: Ligging van de verschillende punten voor het geval waarbij $q > l$.

De foutterm op interpolatieniveau bedraagt, rekening houdend met $y' = f(x,\, y(x))$,

$$(-1)^{q+1-u} \binom{-t}{q+1-u} h^{q+1-u}\, y^{(q+2-u)}(\xi(x))\,.$$

M.b.v. de interpolatieveelterm ontstaat uit (4.7) de methode

$$y_{n+q} - y_{n+q-l} = h \sum_{i=0}^{q-u} \gamma_i\, \nabla^i f_{n+q-u}\,, \tag{4.8}$$

waarbij

$$\gamma_i = (-1)^i \int\limits_{u-l}^{u} \binom{-t}{i}\, \mathrm{d}t\,. \tag{4.9}$$

De fout die hierbij gemaakt wordt, bedraagt dan (in de veronderstelling dat $y_{n+j} \equiv y(x_{n+j})$, $j = 0,\, 1,\, \ldots,\, q-1$) :

$$y(x_{n+q}) - y_{n+q} = h^{q+2-u} \int\limits_{u-l}^{u} (-1)^{q+1-u} \binom{-t}{q+1-u} y^{(q+2-u)}(\xi(x_{n+q-u} + t\,h))\; \mathrm{d}t\,. \tag{4.10}$$

De orde van de aldus bekomen methoden is bijgevolg $p = q + 1 - u$.

Indien $l = 1$ bezit $\binom{-t}{q+1-u}$ een vast teken over het integratie-interval, zodat we steunend op de Middelwaardestelling vinden

$$y(x_{n+q}) - y_{n+q} = h^{q+2-u}\, \gamma_{q+1-u}\, \overline{y}^{(q+2-u)}(\eta)\,, \tag{4.11}$$

waarbij $\eta \in [\min(x_n,\, x_{n+q-l}),\, x_{n+q}]$. De $\overline{y}$-notatie in (4.11) wijst er op dat, indien y meerdimensionaal is, de componenten geëvalueerd kunnen worden in verschillende punten, d.w.z.

$$\overline{y}(\eta) = [\,{}^1y(\eta_1), {}^2y(\eta_2), \ldots, {}^my(\eta_m)]^T\,,$$

waarbij elke η_i behoort tot $[\min(x_n,\, x_{n+q-l}),\, x_{n+q}]$.

Als $l > 1$, kan een afschatting bekomen worden op een andere manier.

De methode (4.8) kan herschreven worden in de vorm (4.1) door de coëfficiënten te groeperen volgens de punten x_{n+j} :

$$\begin{aligned}
y_{n+q} - y_{n+q-l} &= h \sum_{i=0}^{q-u} \gamma_i\, \nabla^i f_{n+q-u} \\
&= h \sum_{i=0}^{q-u} \gamma_i \sum_{m=0}^{i} (-1)^m \binom{i}{m} f_{n+q-u-m} \\
&= h \sum_{m=0}^{q-u} \left((-1)^m \sum_{i=m}^{q-u} \gamma_i \binom{i}{m} \right) f_{n+q-u-m} \\
&= h \sum_{j=0}^{q-u} \left((-1)^{q-u-j} \sum_{i=q-u-j}^{q-u} \gamma_i \binom{i}{q-u-j} \right) f_{n+j} \\
&= h \sum_{j=0}^{q-u} \beta_j\, f_{n+j}
\end{aligned}$$

met $\beta_j = (-1)^{q-u-j} \displaystyle\sum_{i=q-u-j}^{q-u} \gamma_i \binom{i}{q-u-j}$.

In de praktijk wordt de formule slechts gebruikt voor $u = 0$ of $u = 1$. Dit geeft aanleiding tot impliciete, resp. expliciete methoden. De meest bekende methoden (zie Tabel 4.1) uit deze klasse zijn deze met $l = 1$ en $l = 2$.

	$l = 1$	$l = 2$
$u = 0$	Adams–Moulton	Milne–Simpson
$u = 1$	Adams–Bashforth	Nyström

Tabel 4.1: Overzicht van de benaming van de belangrijkste methoden.

Wegens constructie zijn alle methoden die zo worden bekomen steeds convergent. Ze zijn consistent gezien voor elke $q \geq 1$ de interpolatiefout voor de functies $f(x,\, y) \equiv c$ met c constant nul is. De nulstabiliteit volgt uit $\rho(\xi) = \xi^{q-l}\,(\xi^l - 1)$.

De waarden γ_i kunnen berekend worden door invoering van de genererende functie $G(z) = \sum_{i=0}^{\infty} \gamma_i\, z^i$:

$$G(z) = \sum_{i=0}^{\infty} \left(\int_{u-l}^{u} (-1)^i \binom{-t}{i} \, \mathrm{d}t \right) z^i = \int_{u-l}^{u} \left(\sum_{i=0}^{\infty} \binom{-t}{i} (-z)^i \right) \, \mathrm{d}t = \int_{u-l}^{u} (1-z)^{-t} \, \mathrm{d}t \,,$$

waaruit

$$G(z) = \frac{(1-z)^l - 1}{(1-z)^u \, \ln(1-z)} \,. \tag{4.12}$$

Door $G(z)$ in een reeks te ontwikkelen rond $z = 0$ vinden we op deze manier de volgende tabel :

	γ_0	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	$\cdots$
Adams–Bashforth	1	$\frac{1}{2}$	$\frac{5}{12}$	$\frac{3}{8}$	$\frac{251}{720}$	$\frac{95}{288}$	$\frac{19087}{60480}$	
Adams–Moulton	1	$-\frac{1}{2}$	$-\frac{1}{12}$	$-\frac{1}{24}$	$-\frac{19}{720}$	$-\frac{3}{160}$	$-\frac{863}{60480}$	
Nyström	2	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{29}{90}$	$\frac{14}{45}$	$\frac{1139}{3780}$	
Milne–Simpson	2	-2	$\frac{1}{3}$	0	$-\frac{1}{90}$	$-\frac{1}{90}$	$\frac{37}{3780}$	

De nullen die optreden in de bovenstaande tabel zorgen voor eigenaardige effecten. Dit komt tot uiting wanneer we het rechterlid in termen van de coëfficiënten β_j herschrijven :

- Adams–Bashforth: $p = q$ en $k = q$

$$\begin{aligned}
q = 1 \quad &: \quad y_{n+1} - y_n = h\, f_n \\
q = 2 \quad &: \quad y_{n+1} - y_n = \frac{h}{2} (3\, f_n - f_{n-1}) \\
q = 3 \quad &: \quad y_{n+1} - y_n = \frac{h}{12} (23\, f_n - 16\, f_{n-1} + 5\, f_{n-2}) \\
q = 4 \quad &: \quad y_{n+1} - y_n = \frac{h}{24} (55\, f_n - 59\, f_{n-1} + 37\, f_{n-2} - 9\, f_{n-3}) \\
&\;\;\vdots
\end{aligned}$$

- Adams–Moulton: $p = q + 1$ en $k = \max(1,\, q)$

$$\begin{aligned}
q = 0 \quad &: \quad y_{n+1} - y_n = h\, f_{n+1} \\
q = 1 \quad &: \quad y_{n+1} - y_n = \frac{h}{2} (f_{n+1} + f_n) \\
q = 2 \quad &: \quad y_{n+1} - y_n = \frac{h}{12} (5\, f_{n+1} + 8\, f_n - f_{n-1})
\end{aligned}$$

$$
\begin{aligned}
q = 3 \quad &: \quad y_{n+1} - y_n = \frac{h}{24}\,(9\,f_{n+1} + 19\,f_n - 5\,f_{n-1} + f_{n-2}) \\
q = 4 \quad &: \quad y_{n+1} - y_n = \frac{h}{720}\,(251\,f_{n+1} + 646\,f_n - 264\,f_{n-1} + 106\,f_{n-2} - 19\,f_{n-3}) \\
&\;\vdots
\end{aligned}
$$

- Nyström : $p = q$ en $k = \max(2,\, q)$

$$
\begin{aligned}
q = 1 \quad &: \quad y_{n+1} - y_{n-1} = 2\,h\,f_n \qquad\qquad \text{opgelet : } p = 2 \\
q = 2 \quad &: \quad y_{n+1} - y_{n-1} = 2\,h\,f_n \\
q = 3 \quad &: \quad y_{n+1} - y_{n-1} = \frac{h}{3}\,(7\,f_n - 2\,f_{n-1} + f_{n-2}) \\
q = 4 \quad &: \quad y_{n+1} - y_{n-1} = \frac{h}{3}\,(8\,f_n - 5\,f_{n-1} + 4\,f_{n-2} - f_{n-3}) \\
&\;\vdots
\end{aligned}
$$

- Milne-Simpson : $p = q + 1$ en $k = \max(2,\, q)$

$$
\begin{aligned}
q = 0 \quad &: \quad y_{n+1} - y_{n-1} = 2\,h\,f_{n+1} \\
q = 1 \quad &: \quad y_{n+1} - y_{n-1} = 2\,h\,f_n \\
q = 2 \quad &: \quad y_{n+1} - y_{n-1} = \frac{h}{3}\,(f_{n+1} + 4\,f_n + f_{n-1}) \qquad\qquad \text{opgelet : } p = 4 \\
q = 3 \quad &: \quad y_{n+1} - y_{n-1} = \frac{h}{3}\,(f_{n+1} + 4\,f_n + f_{n-1}) \\
q = 4 \quad &: \quad y_{n+1} - y_{n-1} = \frac{h}{90}\,(29\,f_{n+1} + 124\,f_n + 24\,f_{n-1} + 4\,f_{n-2} - f_{n-3}) \\
&\;\vdots
\end{aligned}
$$

Tussen deze uitdrukkingen bevinden zich enkele heel belangrijke methoden. De volgende tabel geeft een overzicht.

naam	orde	uitdrukking
de Euler-regel	1	$y_{n+1} - y_n = h\,f_n$
de achterwaartse Euler-regel	1	$y_{n+1} - y_n = h\,f_{n+1}$
de trapezium-regel	2	$y_{n+1} - y_n = \frac{h}{2}\,(f_{n+1} + f_n)$
de expliciete middelpunt-regel	2	$y_{n+1} - y_{n-1} = 2\,h\,f_n$
de Simpson-regel	4	$y_{n+1} - y_{n-1} = \frac{h}{3}\,(f_{n+1} + 4\,f_n + f_{n-1})$

l	δ_0	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	$\ldots$
0	0	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{7}$	
1	0	1	$-\frac{1}{2}$	$-\frac{1}{6}$	$-\frac{1}{12}$	$-\frac{1}{20}$	$-\frac{1}{30}$	$-\frac{1}{42}$	
2	0	1	$-\frac{3}{2}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{1}{30}$	$\frac{1}{60}$	$\frac{1}{105}$	

Tabel 4.2: De δ_i coëfficiënten voor verschillende waarden van l.

4.2.2 Benadering van $y(x)$ door een interpolatieveelterm

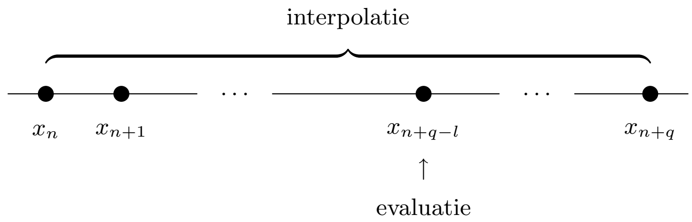
We benaderen $y(x)$ in het linkerlid van (2.1) door de interpolatieveelterm door de punten x_{n+j}, $j = 0, 1, \ldots, q$:

$$y(x) = y(x_{n+q} + t\,h) \approx \tilde{y}(x_{n+q} + t\,h) = \sum_{i=0}^{q} (-1)^i \binom{-t}{i} \nabla^i y_{n+q}\,, \qquad t = \frac{x - x_{n+q}}{h}\,.$$

De interpolatiefout is

$$y(x_{n+q} + t\,h) - \tilde{y}(x_{n+q} + t\,h) = (-1)^{q+1} \binom{-t}{q+1} h^{q+1} y^{(q+1)}(\xi(x_{n+q} + t\,h))\,.$$

We leiden de veelterm af naar t en gebruiken de waarde in $t = -l$ (of $x = x_{n+q-l}$) waarbij $l \geq 0$ als benadering voor $h\,y'(x_{n+q-l})$. We noemen deze benadering $h\,f_{n+q-l}$.



Figuur 4.2: Ligging van de verschillende punten voor het geval waarbij $q > l$.

Zo vinden we

$$h\,f_{n+q-l} = \sum_{i=0}^{q} \delta_i\, \nabla^i y_{n+q}\,, \tag{4.13}$$

waarbij δ_i gegeven is door

$$\begin{aligned} \delta_i &= (-1)^i \frac{\mathrm{d}}{\mathrm{d}t} \binom{-t}{i} \Big|_{t=-l} \\ &= \frac{1}{i!} \frac{\mathrm{d}}{\mathrm{d}t} t\,(t+1)\,\ldots\,(t+i-1) \Big|_{t=-l}\,. \end{aligned}$$

De methode (4.13) die aldus ontstaat heeft $k = \max(q,\, l)$ stappen en is van de orde $p = q$: dit laatste volgt uit

$$
\begin{aligned}
&\sum_{i=0}^{q} \delta_i\, \nabla^i y(x_n + q\,h) - h\, f(x_{n+q-l},\, y(x_{n+q-l})) \\
&= \sum_{i=0}^{q} \delta_i\, \nabla^i y_{n+q} - h\, y'(x_{n+q-l}) \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \tilde{y}(x_{n+q} + t\,h)\Big|_{t=-l} - h\, y'(x_{n+q-l}) \\
&= \frac{\mathrm{d}}{\mathrm{d}t} [\tilde{y}(x_{n+q} + t\,h) - y(x_{n+q} + t\,h)]\Big|_{t=-l} \\
&= -(-1)^{q+1}\, h^{q+1}\, \frac{\mathrm{d}}{\mathrm{d}t} \left(\binom{-t}{q+1} y^{(q+1)}(\xi(x_{n+q} + t\,h)) \right) \Big|_{t=-l} \\
&= -\delta_{q+1}\, h^{q+1}\, y^{(q+1)}(\xi(x_{n+q-l})) + \mathcal{O}(h^{q+2})\,.
\end{aligned}
$$

Werken we in de methode (4.13) de achterwaartse differenties weg, dan ontstaat

$$
\begin{aligned}
\sum_{i=0}^{q} \delta_i\, \nabla^i y_{n+q} &= \sum_{i=0}^{q} \delta_i \sum_{m=0}^{i} (-1)^m \binom{i}{m} y_{n+q-m} \\
&= \sum_{m=0}^{q} (-1)^m \sum_{i=m}^{q} \delta_i \binom{i}{m} y_{n+q-m} \\
&= \sum_{j=0}^{q} (-1)^{q-j} \sum_{i=q-j}^{q} \delta_i \binom{i}{q-j} y_{n+j} \\
&= \sum_{j=0}^{q} \tilde{\alpha}_j\, y_{n+j}
\end{aligned}
$$

waarbij $\tilde{\alpha}_j = (-1)^{q-j} \sum_{i=q-j}^{q} \delta_i \binom{i}{q-j}$. Voeren we tenslotte nog een herschaling uit (zodat $\alpha_q = 1$), dan ontstaat

$$
\sum_{i=0}^{q} \alpha_i\, y_{n+i} = h\, \beta_{q-l}\, f_{n+q-l}\,, \tag{4.14}
$$

waarbij $\beta_{q-l} = 1/\tilde{\alpha}_q = \left(\sum_{i=0}^{q} \delta_i \right)^{-1}$ en $\alpha_i = \beta_{q-l}\, \tilde{\alpha}_i$. De fout geassocieerd aan deze methode is dan

$$
-\beta_{q-l}\, \delta_{q+1}\, h^{q+1}\, y^{(q+1)}(\xi(x_{n+q-l})) + \mathcal{O}(h^{q+2})\,.
$$

Net zoals bij de methodes uit de vorige paragraaf is consistentie reeds vooraf ingebakken omdat $f(x,\, y) = c$ (met c constant) exact wordt benaderd voor $q \geq 1$. De nulstabiliteit daarentegen is niet gewaarborgd.

Als $1 \leq l \leq q$ ontstaan expliciete methoden. Het meest gebruikelijk is echter de keuze $l = 0$, die aanleiding geeft tot impliciete methoden, de k-staps BDF-methoden (*Backward Difference*

Formulas). In dat geval geldt $\delta_i = i^{-1}$, $i = 1, 2, \ldots$. Men heeft in het bijzonder aangetoond dat de BDF-methoden slechts nulstabiel zijn voor $1 \le q \le 6$. De coëfficiënten van deze BDF-methoden worden gegeven in Tabel 4.3.

q	α_6	α_5	α_4	α_3	α_2	α_1	α_0	β_q	$-\beta_q\,\delta_{q+1}$
1						1	-1	1	$-\frac{1}{2}$
2					1	$-\frac{4}{3}$	$\frac{1}{3}$	$\frac{2}{3}$	$-\frac{2}{9}$
3				1	$-\frac{18}{11}$	$\frac{9}{11}$	$-\frac{2}{11}$	$\frac{6}{11}$	$-\frac{3}{22}$
4			1	$-\frac{48}{25}$	$\frac{36}{25}$	$-\frac{16}{25}$	$\frac{3}{25}$	$\frac{12}{25}$	$-\frac{12}{125}$
5		1	$-\frac{300}{137}$	$\frac{300}{137}$	$-\frac{200}{137}$	$\frac{75}{137}$	$-\frac{12}{137}$	$\frac{60}{137}$	$-\frac{10}{137}$
6	1	$-\frac{360}{147}$	$\frac{450}{147}$	$-\frac{400}{147}$	$\frac{225}{147}$	$-\frac{72}{147}$	$\frac{10}{147}$	$\frac{60}{147}$	$-\frac{20}{343}$

Tabel 4.3: Coëfficiënten van de BDF-methode.

4.3 De differentieoperator $\mathcal{L}$ geassocieerd aan een LMM

In paragraaf 2.4 werd de orde van een numerieke methode ingevoerd. Voor LMM kunnen we deze definitie herformuleren aan de hand van de differentieoperator $\mathcal{L}$.

Definitie 4.3.1 *De* lineaire differentieoperator $\mathcal{L}$ *geassocieerd aan de LMM (4.1) is gedefinieerd door*

$$\mathcal{L}[z(x);\, h] := \sum_{j=0}^{k} [\alpha_j\, z(x + j\,h) - h\,\beta_j\, z'(x + j\,h)]\,, \tag{4.15}$$

waarbij $z(x) \in C^1[a, b]$ *een willekeurige functie is.* □

We kiezen de functie $z(x)$ zo dat ze voldoende afleidbaar is, ontwikkelen $z(x+j\,h)$ en $z'(x+j\,h)$ rond x, en groeperen termen gelijke machten van h. Zo vinden we dat

$$\mathcal{L}[z(x);\, h] = C_0\, z(x) + C_1\, h\, z^{(1)}(x) + \ldots + C_q\, h^q\, z^{(q)}(x) + \ldots \tag{4.16}$$

waarbij

$$\begin{aligned} C_0 &= \sum_{j=0}^{k} \alpha_j \equiv \rho(1) \\ C_1 &= \sum_{j=0}^{k} (j\,\alpha_j - \beta_j) \equiv \rho'(1) - \sigma(1) \\ C_q &= \sum_{j=0}^{k} \left[\frac{1}{q!} j^q\,\alpha_j - \frac{1}{(q-1)!}\, j^{q-1}\,\beta_j\right], \quad q = 2, 3, \ldots\,. \end{aligned} \tag{4.17}$$

Definitie 4.3.2 *De LMM (4.1) en zijn geassocieerde differentieoperator $\mathcal{L}$ zijn van de orde p en bezitten foutcontstante C_{p+1} a.s.a. $C_0 = C_1 = \ldots = C_p = 0$ en $C_{p+1} \neq 0$.* □

Deze definitie is slechts bruikbaar indien hetzelfde resultaat bekomen wordt bij ontwikkeling rond een ander punt. Veronderstel dat we $z(x + j\,h)$ en $z'(x + j\,h)$ ontwikkelen rond $x + t\,h$ i.p.v. rond x. Dan bekomen we :

$$\mathcal{L}[z(x);\, h] = D_0\, z(x + t\,h) + D_1\, h\, z^{(1)}(x + t\,h) + \ldots + D_q\, h^q\, z^{(q)}(x + t\,h) + \ldots . \quad (4.18)$$

De functies $z^{(q)}(x + t\,h), q = 0, 1, 2, \ldots$ (waarbij $z^{(0)}(x) \equiv z(x)$) kunnen elk ontwikkeld worden in Taylor-reeks rond x :

$$z^{(q)}(x + t\,h) = z^{(q)}(x) + t\,h\, z^{(q+1)}(x) + \ldots + \frac{t^s\, h^s}{s!}\, z^{(q+s)}(x) + \ldots , \quad q = 0, 1, 2, \ldots (4.19)$$

Introduceren we deze ontwikkelingen in (4.18), en vergelijken we de bekomen uitdrukking met de rechterzijde van (4.16), dan vinden we

$$\begin{aligned}
C_0 &= D_0 \\
C_1 &= D_1 + t\, D_0 \\
C_2 &= D_2 + t\, D_1 + \frac{t^2}{2!}\, D_0 \\
&\vdots \\
C_p &= D_p + t\, D_{p-1} + \ldots + \frac{t^p}{p!}\, D_0 \\
C_{p+1} &= D_{p+1} + t\, D_p + \ldots + \frac{t^{p+1}}{(p+1)!}\, D_0 \\
C_{p+2} &= D_{p+2} + t\, D_{p+1} + \ldots + \frac{t^{p+2}}{(p+2)!}\, D_0 \, .
\end{aligned}$$

Hieruit volgt dat $C_q = 0$, $q = 0, 1, \ldots, p$ a.s.a. $D_q = 0$, $q = 0, 1, \ldots, p$ en dan is bovendien ook $C_{p+1} = D_{p+1}$, wat aantoont dat de definitie inderdaad onafhankelijk is van de gekozen t. Merk echter wel op dat, gezien $D_{p+2} = C_{p+2} - t\, C_{p+1}$, geen belang meer kan gehecht worden aan deze coëfficiënt.

Voorbeeld 4.3.1

Beschouw de volgende twee-parameter familie van lineaire 2-stapsmethoden :

$$y_{n+2} - (1+\alpha)\, y_{n+1} + \alpha\, y_n = h\,[(1+\beta)\, f_{n+2} - (\alpha + \beta + \alpha\,\beta)\, f_{n+1} + \alpha\,\beta\, f_n]$$

met $\alpha\, (\neq 0)$ en β vrije parameters. Voor deze familie geldt

$$\begin{aligned}
C_0 &= 0 \\
C_1 &= 0 \\
C_2 &= (\alpha - 1)\left(\beta + \frac{1}{2}\right) \\
C_3 &= \begin{cases} -\beta - \frac{1}{2} & \text{als } \alpha = 1 \\ \frac{1}{12}\,(\alpha - 1) & \text{als } \beta = -\frac{1}{2} \end{cases} \\
C_4 &= -\frac{1}{12} \text{ als } \alpha = 1 \text{ en } \beta = -\frac{1}{2} \, .
\end{aligned}$$

Hieruit volgt:

(a) $\alpha \neq 1,\, \beta \neq -\dfrac{1}{2}$: orde $p = 1$, foutconstante $C_2 = (\alpha - 1)\left(\beta + \frac{1}{2}\right)$

(b) $\alpha \neq 1,\, \beta = -\dfrac{1}{2}$: orde $p = 2$, foutconstante $C_3 = \dfrac{1}{12}(\alpha - 1)$

(c) $\alpha = 1,\, \beta \neq -\dfrac{1}{2}$: orde $p = 2$, foutconstante $C_3 = -\beta - \dfrac{1}{2}$

(d) $\alpha = 1,\, \beta = -\dfrac{1}{2}$: orde $p = 3$, foutconstante $C_4 = -\dfrac{1}{12}$.

Merk wel op dat de methode niet nulstabiel is voor $\alpha = 1$. □

Opmerking 4.3.1

Gelet op de voorgaande definities kunnen we stellen : een LMM is consistent a.s.a. de orde van de geassocieerde differentieoperator tenminste 1 is. □

Stelling 4.3.1 *Een lineaire differentieoperator $\mathcal{L}$ is van de orde p a.s.a.*

$$\mathcal{L}[x^q;\, h] \equiv 0, \quad q = 0, 1, \ldots, p \quad \textit{en} \quad \mathcal{L}[x^{p+1};\, h] \not\equiv 0\,. \tag{4.20}$$

Bovendien geldt

$$C_{p+1} = \frac{1}{(p+1)!\, h^{p+1}}\, \mathcal{L}[x^{p+1};\, h]\,. \tag{4.21}$$

□

Bewijs. Zij $\mathcal{L}$ van de orde p. Dan is

$$\mathcal{L}[z(x);\, h] = C_{p+1}\, h^{p+1}\, z^{(p+1)}(x) + C_{p+2}\, h^{p+2}\, z^{(p+2)}(x) + \ldots$$

Vermits $z^{(p+1)} \equiv 0$ voor $z(x) = x^q$ met $q \leq p$ en $z^{(p+1)} = (p+1)!$ voor $z(x) = x^{p+1}$ is het $\Rightarrow$ gedeelte bewezen.
Omgekeerd, zij $\mathcal{L}[x^q;\, h] \equiv 0$ voor $q = 0, 1, \ldots, p$ en $\mathcal{L}[x^{p+1};\, h] \not\equiv 0$. Als $z(x) = x^q$ is $z^{(q)}(x) = q!$, zodat de coëfficiënt van $\mathcal{L}[x^q;\, h]$ bij h^q gelijk is aan $C_q\, q!$. Gezien de veronderstellingen volgt hieruit dat $C_q = 0$ voor $q \leq p$ en $C_{p+1} \neq 0$, m.a.w. $\mathcal{L}$ is van de orde p. ■

Voorbeeld 4.3.2

Stel dat we een impliciete lineaire 2-stapsmethode van zo groot mogelijke orde willen construeren die één vrije parameter bevat. We kunnen vertrekken van de vorm

$$y_{n+2} + \alpha_1\, y_{n+1} + \alpha_0\, y_n = h\,[\beta_2\, f_{n+2} + \beta_1\, f_{n+1} + \beta_0\, f_n]\,.$$

Deze vorm bevat vijf parameters : α_0, α_1, β_0, β_1 en β_2. Door vier voorwaarden op te leggen kunnen we vier van hen uitdrukken in termen van de vijfde. Kiezen we $\alpha_0 = a$ als overblijvende parameter, dan volgt door $y(x) = x^q$, $q = 0, 1, 2, 3$ en bijvoorbeeld $x_n = 0$ te stellen :

$$\begin{cases} 0!\,C_0 = 0 = a + \alpha_1 + 1 \\ 1!\,C_1 = 0 = \phantom{a+{}} \alpha_1 + 2 - 1\,(\beta_0 + \beta_1 + \beta_2) \\ 2!\,C_2 = 0 = \phantom{a+{}} \alpha_1 + 4 - 2\,(\phantom{\beta_0+{}} \beta_1 + 2\,\beta_2) \\ 3!\,C_3 = 0 = \phantom{a+{}} \alpha_1 + 8 - 3\,(\phantom{\beta_0+{}} \beta_1 + 4\,\beta_2) \end{cases} \tag{4.22}$$

De oplossing van (4.22) luidt

$$\alpha_1 = -(a+1) \quad \beta_0 = -\frac{1}{12}\,(1+5\,a) \quad \beta_1 = \frac{2}{3}\,(1-a) \quad \beta_2 = \frac{1}{12}\,(5+a)\,,$$

m.a.w. we vinden de methode

$$y_{n+2} - (1+a)\,y_{n+1} + a\,y_n = \frac{h}{12}\,[(5+a)\,f_{n+2} + 8\,(1-a)\,f_{n+1} - (1+5\,a)\,f_n]\,. \tag{4.23}$$

Deze familie levert slechts nulstabiele methoden voor $-1 \le a < 1$. Bovendien vinden we $z(x) = x^4$ en $z(x) = x^5$:

$$4!\,C_4 = -(a+1) \quad \text{en} \quad 5!\,C_5 = -\frac{1}{3}\,(17+13\,a)\,.$$

Dit betekent dat voor $a \ne -1$ de methode van orde 3 is en van orde 4 voor $a = -1$. In dat geval vinden we trouwens

$$y_{n+2} - y_n = \frac{h}{3}\,[f_{n+2} + 4\,f_{n+1} + f_n]\,,$$

wat de impliciete Milne-methode is van orde 4. Voor $a = 0$ vinden we de 2-staps Adams–Moulton-methode :

$$y_{n+2} - y_{n+1} = \frac{h}{12}\,[5\,f_{n+2} + 8\,f_{n+1} - f_n]$$

Voor $a = -5$ vinden we een niet-nulstabiele expliciete methode van orde 3. □

Voorbeeld 4.3.3

Stel dat we nu impliciete lineaire 3-stapsmethode van zo groot mogelijke orde willen construeren die één vrije parameter bevat. Dan vertrekken we van de vorm

$$y_{n+3} + \alpha_2\,y_{n+2} + \alpha_1\,y_{n+1} + \alpha_0\,y_n = h\,[\beta_3\,f_{n+3} + \beta_2\,f_{n+2} + \beta_1\,f_{n+1} + \beta_0\,f_n]\,.$$

Deze vorm bevat zeven parameters : α_0, α_1, α_2 β_0, β_1, β_2 en β_3. Door zes voorwaarden op te leggen kunnen we zes van hen uitdrukken in termen van de zevende. Kiezen we $\alpha_0 = a$ als overblijvende parameter, dan volgt uit $y(x) = x^q$, $q = 0, 1, \ldots, 5$ de oplossing

$$\alpha_1 = \frac{1}{11}\,(8\,a - 19) \quad \alpha_2 = \frac{1}{11}\,(8 - 19\,a) \quad \beta_0 = -\frac{1}{33}\,(1 + 10\,a)$$

$$\beta_1 = \frac{1}{11}\,(8 - 19\,a) \quad \beta_2 = \frac{1}{11}\,(19 - 8\,a) \quad \beta_3 = \frac{1}{33}\,(10 + a)\,,$$

m.a.w. we vinden de vijfde-orde methode

$$y_{n+3} + \frac{1}{11}\,(8 - 19\,a)\,y_{n+2} + \frac{1}{11}\,(8\,a - 19)\,y_{n+1} + a\,y_n =$$
$$\frac{h}{33}\,[(10 + a)\,f_{n+3} + 3\,(19 - 8\,a)\,f_{n+2} + 3\,(8 - 19\,a)\,f_{n+1} - (1 + 10\,a)\,f_n]\,.$$

Geen enkele methode uit deze klasse is echter nulstabiel, gezien de wortels van de eerste karakteristieke veelterm te schrijven zijn als

$$r_1 = 1\,, \quad r_2 = \frac{19}{22}\,w + \sqrt{\frac{361}{484}\,w^2 + w + 1}\,, \quad r_3 = \frac{19}{22}\,w - \sqrt{\frac{361}{484}\,w^2 + w + 1}\,,$$

waarbij $a = w + 1$. Als $w < 0$ is $r_3 < -1$, als $w = 0$ is $r_1 = r_2$ en als $w > 0$ is $r_2 > 1$. Er bestaan dus geen nulstabiele lineaire 3-stapsmethoden van orde 5. $\square$

Stelling 4.3.2 *Als $\mathcal{L}$ een lineaire differentie-operator is die geassocieerd is aan een symmetrische LMM van orde p, dan legt*

$$\mathcal{L}[x^{2\,q};\,h] \equiv 0 \text{ voor } q \in \mathbb{N} \text{ met } 2\,q \leq p + 1$$

geen voorwaarde op aan de α- en β-coëfficiënten, m.a.w. de coëfficiënten worden alleen bepaald door

$$\mathcal{L}[x^{2\,q+1};\,h] \equiv 0 \text{ voor } q \in \mathbb{N} \text{ met } 2\,q + 1 \leq p\,.$$

$\square$

Bewijs. Beschouw de lineaire differentie-operator die geassocieerd is aan de methode

$$\sum_{j=0}^{k} \alpha_j\,y_{n+j} = h\,\sum_{j=0}^{k} \beta_j\,f_{n+j} \;:$$

$$L[y(x);\,h] = \sum_{j=0}^{k} \alpha_j\,y(x + j\,h) - h\,\sum_{j=0}^{k} \beta_j\,y'(x + j\,h)\,.$$

We weten dat het punt x waarrond ontwikkeld wordt geen rol speelt voor machten h^i met $i \leq p + 1$. Stel daarom $x = -k\,h/2$:

$$L[y(-k\,h/2);\,h] = \sum_{j=0}^{k} \alpha_j\,y((j - k/2)\,h) - h\,\sum_{j=0}^{k} \beta_j\,y'((j - k/2)\,h)\,.$$

Gelet op de symmetrie is dan ofwel $\alpha_{k-j} = -\alpha_j$ en $\beta_{k-j} = \beta_j$, ofwel $\alpha_{k-j} = \alpha_j$ en $\beta_{k-j} = -\beta_j$. De tweede mogelijkheid valt echter niet te combineren met de voorwaarden voor convergentie. Voor $k = 1$ is dit makkelijk na te rekenen : in dat geval immers zou $\rho(r) = r + 1$, wat zou betekenen dat $\rho(1) \neq 0$. Voor $k > 1$ geldt de volgende redenering : als $\rho(r) = r^k\,\rho(\frac{1}{r})$, dan is $\rho'(r) = k\,r^{k-1}\,\rho(\frac{1}{r}) - r^{k-2}\,\rho'(\frac{1}{r})$ en in het bijzonder is dan $\rho'(1) = -\rho'(1)$, wat betekent dat $\rho(1) = \rho'(1) = 0$, zodat de methode niet nul-stabiel is. Mogelijkheid (ii) moet dus uitgesloten worden.

In de bovenstaande som nemen we nu de termen voor j en $k - j$ samen voor $y = x^{2\,q}$ en we vinden dat die som nul is :

- $\alpha_j\, y((j-k/2)\,h) + \alpha_{k-j}\, y((k-j-k/2)\,h) = 0$
- $\beta_j\, y'((j-k/2)\,h) + \beta_{k-j}\, y'((k-j-k/2)\,h) = 0$

Hierbij werd rekening gehouden met het feit dat, als $y(x)$ even is, $y(-x) = y(x)$ en $y'(-x) = -y'(x)$. Merk op dat als k even is, het aantal termen in elke som oneven is. Geen nood echter : ook in dat geval verdwijnt alles vermits $\alpha_{k/2} = 0$ en $y'(0) = 0$ als $y(x)$ even is. ■

4.4 De eerste Dahlquist-barrière

De voorbeelden van de voorbije paragraaf doen de volgende vraag ontstaan : wat is de grootste orde die kan bereikt worden bij een convergente lineaire k-stapsmethode ?

De lineaire k-stapsmethode (4.1) bezit $2\,k+2$ parameters, nl. α_j, β_j, $j = 0, 1, \ldots, k$, waarbij evenwel steeds $\alpha_k = 1$. Bovendien is voor een expliciete methode $\beta_k = 0$. Dit betekent dat een impliciete LMM $2\,k+1$ vrije parameters bezit, een expliciete $2\,k$.

Nu volgt uit (4.16) dat $p+1$ lineaire vergelijkingen in α_j, β_j , $j = 0, 1, \ldots, k$ moeten voldaan zijn als de methode de orde p moet bezitten. De hoogste orde die we aldus met een lineaire k-stapsmethode kunnen bereiken, is $2\,k$ als ze impliciet is en $2\,k-1$ als ze expliciet is. Een LMM die deze bovengrens bereikt wordt *maximaal* genoemd.

Het is echter evident dat we slechts geïnteresseerd zijn in convergente methoden en de voorbeelden van de vorige paragraaf maken duidelijk dat die convergentie niet altijd gewaarborgd is. De consistentievoorwaarden zijn steeds voldaan als $p > 1$ wordt geëist, maar er ontstaat een bovengrens wanneer de nulstabiliteit aan de orde komt. Deze grens staat bekend als de *eerste Dahlquist-barrière*.

Stelling 4.4.1 *De maximale orde van een convergente k-staps methode is $k+1$ als k oneven is en $k+2$ als k even is.* □

Stelling 4.4.2 *Voor een LMM met k even die van de orde $k+2$ is, liggen alle wortels van $\rho(\xi)$ op de eenheidscirkel.* □

Voorbeeld 4.4.1

We bepalen de optimale LMM met $k = 4$. Rekening houdend met Stelling 4.4.2 geldt voor een zekere $0 < \theta < \pi$:

$$\begin{aligned}\rho(\xi) &= (\xi-1)\,(\xi+1)\,(\xi-\exp(i\,\theta))\,(\xi-\exp(-i\,\theta))\\ &= (\xi^2-1)\,(\xi-(\cos\theta+i\sin\theta))\,(\xi-(\cos\theta-i\sin\theta))\\ &= \xi^4-2\cos\theta\,\xi^3+2\,\cos\theta\,\xi-1\,.\end{aligned}$$

Stellen we $\cos\theta = \mu$, dan bekomen we

$$\alpha_4 = 1, \quad \alpha_3 = -2\,\mu, \quad \alpha_2 = 0, \quad \alpha_1 = 2\,\mu, \quad \alpha_0 = -1\;.$$

De veelterm $\sigma(\xi)$ bepalen we nu uit de eis dat de methode van de orde $k+2 = 6$ is. Omwille van de symmetrie van de α_i waarden, is het aangewezen $z(x) = x^q$ te ontwikkelen rond $x = -2\,h$. Voeren we de waarden voor α_i in, dan verkrijgen we de volgende vergelijkingen voor de β_i:

$$\left\{\begin{array}{rl} \beta_0 + \beta_1 + \beta_2 + \beta_3 + \quad \beta_4 = & 4 - 4\,\mu \\ -\ 2\,\beta_0 - \beta_1 \qquad\quad + \beta_3 + \ 2\,\beta_4 = & 0 \\ 4\,\beta_0 + \beta_1 \qquad\quad + \beta_3 + \ 4\,\beta_4 = & \frac{2}{3}\,(8 - 2\,\mu) \\ -\ 8\,\beta_0 - \beta_1 \qquad\quad + \beta_3 + \ 8\,\beta_4 = & 0 \\ 16\,\beta_0 + \beta_1 \qquad\quad + \beta_3 + 16\,\beta_4 = & \frac{2}{5}\,(32 - 2\,\mu) \\ -\,32\,\beta_0 - \beta_1 \qquad\quad + \beta_3 + 32\,\beta_4 = & 0 \end{array}\right.$$

De oplossing van dit stelsel van 6 vergelijkingen in 5 onbekenden volgt uit de keuze van vijf vergelijkingen. Het stelsel gevormd door de eerste vijf vergelijkingen bvb. heeft een unieke oplossing gezien de coëfficiëntenmatrix een Vandermonde-matrix is. De oplossing ervan luidt :

$$\beta_0 = \beta_4 = \frac{1}{45}\,(14 + \mu)\,, \quad \beta_1 = \beta_3 = \frac{1}{45}\,(64 - 34\,\mu)\,, \quad \beta_2 = \frac{1}{15}\,(8 - 38\mu)\,.$$

Controle leert dat deze oplossing ook een oplossing is van de zesde vergelijking. De foutconstante luidt

$$C_7 = -\frac{16 + 5\,\mu}{1890}\,.$$

Vermits $\mu = \cos\theta$ is $-1 < \mu < 1$, zodat geen aanvaardbare $\mu-$waarde kan gekozen worden waarvoor de foutconstante verdwijnt. Ook β_4 kan niet nul gemaakt worden, wat betekent dat de methode impliciet is. □

Een convergente k-staps methode van orde $k+2$ is steeds symmetrisch. Het aantal parameters in een dergelijke methode is k, nl. $k/2-1$ α's en $k/2+1$ β's. Om orde $k+2$ te bereiken moeten slechts $k/2+1$ voorwaarden opgelegd worden, nl. $C_{2\,i+1} = 0$ met $i = 0, 1, \ldots, k/2$. Dit betekent dat alle β-coëfficiënten kunnen vastgelegd worden in termen van de α-coëfficiënten, die als vrije parameters kunnen beschouwd worden.