

MC SYLLABUS

2.1

MC SYLLABUS

2.1

COLLOQUIUM STABILITEIT VAN DIFFERENTIESCHEMA'S

DEEL 1

DOOR

G.J.R. FÖRCH

P.J. VAN DER HOUWEN

R.P. VAN DE RIET

MATHEMATISCH CENTRUM AMSTERDAM

1967

Bewerkt door: P.J. van der Houwen

Inhoud

	blz.
Voorwoord	1
1. Inleiding	3
R.P. van de Riet, Math. Centrum Amsterdam	
2. Differentieschema's voor gewone differentiaalvergelijkingen	18
P.J. van der Houwen, Math. Centrum Amsterdam	
3. Differentieschema's voor partiële differentiaalvergelijkingen	39
P.J. van der Houwen, Math. Centrum Amsterdam	
4. Zwakke en sterke stabiliteit	55
G.J.R. Förch, Math. Centrum Amsterdam	
5. Enkele methoden voor stabiliteitsonderzoek	60
G.J.R. Förch, Math. Centrum Amsterdam	
Literatuur	77

Voorwoord

Deze tweede uitgave in de Syllabusreeks van het Mathematisch Centrum is samengesteld uit de stof van de voordrachten gegeven in het Colloquium "Stabiliteit van differentieschema's", dat in 1965/'66 plaats vond op het Mathematisch Centrum onder leiding van Prof.dr. H.A. Lauwerier en Prof.dr.ir. A. van Wijngaarden.

Het eerste deel is gewijd aan de algemene theorie van differentieschema's, ontleend aan de voordrachten van Drs. G.J.R. Förch, Drs. P.J. van der Houwen en Drs. R.P. van de Riet. In hoofdstuk 1 worden de verschillende aspecten van de differentiërmethode aan de hand van voorbeelden toegelicht. De volgende twee hoofdstukken geven voor gewone respectievelijk partiële differentiaalvergelijkingen een meer algemene en strengere formulering van differentieschema's. De stabiliteit van differentieschema's blijkt een belangrijke plaats in te nemen. Door de invoering van enkele nieuwe notaties kan de theorie bijzonder compact weergegeven worden. In hoofdstuk 4 worden de begrippen zwakke en sterke stabiliteit, die hun oorsprong in de rekenpraktijk hebben, nader onderzocht. Het eerste deel wordt besloten met enkele methoden om de stabiliteit van lineaire differentieschema's te onderzoeken.

De stof voor het tweede deel is ontleend aan de voordrachten van Dr. L. Dekker, Dr. T.J. Dekker, Drs. P.J. van der Houwen en Drs. M.N. Spijker. Er worden een aantal afzonderlijke onderwerpen besproken, waarbij aan de orde komen een parabolische vergelijking (diffusieproblemen), een hyperbolische vergelijking (Noordzeeprobleem) en elliptische randwaardeproblemen. Voorts de numerieke bepaling van eigenwaarden en eigenvectoren en een methode, waarmee het mogelijk is het stabiliteitsonderzoek uit te breiden tot niet-lineaire vergelijkingen.

1. Inleiding

De problemen uit de mathematische fysica leiden vaak tot het oplossen van differentiaalvergelijkingen met rand- en beginvoorwaarden. De analytische oplossingsmethoden zijn in het algemeen slechts toepasbaar voor een beperkte klasse van differentiaalvergelijkingen en bovendien wordt de oplossing dikwijls verkregen in de vorm van een reeks met weinig toegankelijke speciale functies. Het is dan moeilijk de getallen te berekenen waarin de fysicus geïnteresseerd is.

De methoden die in deze syllabus ter sprake zullen komen, de zogenaamde differentiemethoden, zijn toepasbaar op een veel grotere klasse van differentiaalvergelijkingen en zijn meer geschikt om kwantitatieve gegevens te verkrijgen.

In deze inleiding zullen begrippen als differentieschema, convergentie en stabiliteit aan de hand van voorbeelden toegelicht worden.

Differentieschema's

Stel dat gevraagd wordt de functie $U(t)$, die voldoet aan de differentiaalvergelijking

$$(1.1) \quad \frac{dU}{dt} + U = F(t)$$

voor $t \geq 0$, en aan de beginconditie

$$(1.2) \quad U(0) = 1.$$

Men kan dan als volgt te werk gaan. De afgeleide $\frac{dU}{dt}$ wordt vervangen door de uitdrukking

$$(1.3) \quad \frac{U(t+\tau) - U(t)}{\tau},$$

waarin $\tau > 0$, zodat (1.1) overgaat in de vergelijking

$$(1.4) \quad \frac{U(t+\tau) - U(t)}{\tau} + U(t) = F(t).$$

De oplossing van (1.1) zal in het algemeen niet aan (1.4) voldoen. Stel namelijk dat $\tilde{U}$ aan (1.1) voldoet en een continue tweede afgeleide bezit, dan vindt men met de Taylor-ontwikkeling voor $\tilde{U}$

$$(1.5) \quad \tilde{U}(t+\tau) = \tilde{U}(t) + \frac{dU}{dt} + \frac{1}{2} \tau^2 \frac{d^2 \tilde{U}}{dt^2} \Big|_{t+\theta\tau},$$

met $0 \leq \theta \leq 1$, voor $\tilde{U}$ de vergelijking

$$(1.6) \quad \frac{\tilde{U}(t+\tau) - \tilde{U}(t)}{\tau} + \tilde{U}(t) = \tilde{U}_t + \tilde{U} + \frac{1}{2} \tau \tilde{U}_{tt} \Big|_{t+\theta\tau} \\ = F(t) + \frac{1}{2} \tau \tilde{U}_{tt} \Big|_{t+\theta\tau}.$$

Voor $\tau \rightarrow 0$ gaat vergelijking (1.4) echter in (1.1) over. Men noemt (1.4) consistent met (1.1).

We zullen vergelijking (1.4) iets anders schrijven; daartoe markeren we op het t interval $(0, \infty)$ de equidistante punten $t = t_k = k\tau$ en we noteren de getallen $U(t_k)$ met u_k . Het differentieschema voor (1.1) is nu

$$(1.7) \quad \frac{u_{k+1} - u_k}{\tau} + u_k = F(t_k) = f_k.$$

De verzameling getallen u_k zullen we de differentieoplossing noemen.

Combineren we vergelijking (1.7) met de beginconditie

$$(1.8) \quad u_0 = 1,$$

dan zijn we in staat alle u_k 's te berekenen volgens de eenvoudige recurrente betrekking

$$(1.9) \quad u_{k+1} = u_k + \tau(f_k - u_k).$$

We zullen nog enkele algemene eigenschappen van differentieschema's vermelden. Voor een grondige behandeling van differentieschema's zij echter verwezen naar Milne [6].

Beschouw het iets algemenere schema

$$(1.10) \quad au_{k+1} + bu_k = f_k$$

waarin $k \geq 0$. Als $f_k \equiv 0$ dan is de algemene oplossing voor $k \geq 0$

$$(1.11) \quad u_k = \left(-\frac{b}{a}\right)^k u_0.$$

Wanneer $f_k \neq 0$ is een algemene oplossing aan te geven zodra een particuliere oplossing $\bar{u}_k$ bekend is; voor $k > 0$ geldt

$$(1.12) \quad u_k = \bar{u}_k + \left(-\frac{b}{a}\right)^k u_0.$$

Voor een n^{de} orde differentievergelijking

$$(1.13) \quad \sum_{i=0}^n a_i u_{k+i} = f_k$$

is op analoge wijze uit de algemene oplossing van de homogene vergelijking en een particuliere oplossing $\bar{u}_k$, de algemene oplossing samen te stellen. De oplossing van de homogene vergelijking vinden we door te substitueren

$$u_k = \alpha^k.$$

u_k is een oplossing wanneer α voldoet aan

$$(1.14) \quad \sum_{i=0}^n a_i \alpha^i = 0.$$

Wanneer (1.12) n verschillende wortels α_i heeft geldt voor de algemene oplossing

$$(1.15) \quad u_k = c_1 \alpha_1^k + c_2 \alpha_2^k + \dots + c_n \alpha_n^k,$$

waarin $c_1, c_2, \dots, c_n$ constanten zijn.

Wanneer er samenvallende wortels zijn is (1.3) niet de algemene oplossing.

Voor dit geval verwijzen we naar de literatuur.

De oplossing van de inhomogene vergelijking wordt

$$(1.16) \quad u_k = \sum_{i=1}^n c_i \alpha_i^k + \bar{u}_k.$$

De n constanten c_i kunnen bepaald worden door n extra voorwaarden op te leggen b.v. door $u_0, u_1, \dots, u_{n-1}$ voor te schrijven.

Convergentie van de differentieoplossing naar de analytische oplossing

Aan de hand van vergelijking (1.1) met $F(t) \equiv 0$ en beginvoorwaarde (1.2) zullen we het begrip convergentie aanschouwelijk maken.

Veronderstel dat we de oplossing voor het tijdsinterval $0 \leq t \leq 1$ willen berekenen. Met behulp van schema (1.9), waarin $f_k \equiv 0$, en formule (1.11), vinden we als $u_0 = 1$

$$(1.17) \quad u_k = (1 - \tau)^k.$$

Kiezen we $\tau = 1/N$, dan geldt

$$\begin{aligned} u_k &= \left(1 - \frac{1}{N}\right)^k = \exp\left[k \ln\left(1 - \frac{1}{N}\right)\right] \\ &= \exp\left[k \left(-\frac{1}{N} - \frac{1}{2N^2} + O(N^{-3})\right)\right] \\ &= \exp\left(-\frac{k}{N}\right) \cdot \exp\left(-\frac{k}{2N^2} + kO(N^{-3})\right) \\ &= \exp\left(-\frac{k}{N}\right) \cdot \left(1 - \frac{k}{2N^2} + kO(N^{-3})\right). \end{aligned}$$

De analytische oplossing luidt

$$U(t) = \tilde{U}(t) = \exp(-t),$$

dus

$$\begin{aligned} (1.18) \quad \tilde{U}(k\tau) - u_k &= \exp\left(-\frac{k}{N}\right) \cdot \left(-\frac{k}{2N^2} + kO(N^{-3})\right) \\ &= \exp\left(-\frac{t}{N}\right) \cdot \left(-\frac{t}{2N} + O(N^{-2})\right). \end{aligned}$$

De differentieoplossing convergeert dus naar de analytische oplossing, wanneer $N \rightarrow \infty$ of $\tau \rightarrow 0$. De nauwkeurigheid, ofwel de discretiseringsfout, is van de orde $\frac{1}{N} = \tau$. We merken op dat ook de nauwkeurigheid van het differentieschema (benaderingsfout) volgens (1.6) van de eerste orde in τ was.

Vervolgens bespreken we een differentieschema voor vergelijking (1.1) waarvoor de benaderingsfout en de discretiseringsfout van de oplossing beide van de tweede orde in τ zijn.

Beschouw het schema

$$(1.19) \quad \frac{u_{k+1} - u_{k-1}}{2\tau} + u_k = 0.$$

De benaderingsfout vinden we door de analytische oplossing $\tilde{U}(t)$ te substitueren:

$$(1.20) \quad \frac{\tilde{U}(t_k + \tau) - \tilde{U}(t_k - \tau)}{2} + \tilde{U}(t_k) = \tilde{U}_t(t_k) + \tilde{U}(t_k) + \frac{1}{6} \tau^2 \tilde{U}_{ttt}(t_k + \theta_k \tau) \\ = \frac{1}{6} \tau^2 \tilde{U}_{ttt}(t_k + \theta_k \tau),$$

waarin $0 \leq \theta_k \leq 1$. De nauwkeurigheid is dus van de tweede orde in τ .

De differentieoplossing van (1.19) vinden we door substitutie van $u_k = \alpha^k$. Dit geeft voor α de vergelijking

$$\alpha^2 - 1 + 2\tau\alpha = 0$$

met oplossingen

$$(1.21) \quad \alpha_{1,2} = -\tau \pm \sqrt{\tau^2 + 1}.$$

De algemene oplossing is

$$(1.22) \quad u_k = c_1 \alpha_1^k + c_2 \alpha_2^k.$$

De constanten c_1 en c_2 moeten gevonden worden uit twee extra voorwaarden, b.v. $u_0 = 1$ en $u_1 = a$. De waarde voor a kan gevonden worden door schema (1.9) voor $k = 0$ toe te passen. Substitutie van deze voorwaarden in (1.22) geeft voor c_1 en c_2 de betrekkingen

$$c_1 + c_2 = 1, \quad c_1 \alpha_1 + c_2 \alpha_2 = a,$$

zodat

$$(1.23) \quad c_1 = \frac{\alpha_2 - a}{\alpha_2 - \alpha_1}, \quad c_2 = \frac{a - \alpha_1}{\alpha_2 - \alpha_1}.$$

We ontwikkelen α_1 en α_2 in machten van τ

$$(1.24) \quad \left\{ \begin{array}{l} \alpha_1 = 1 - \tau + \frac{1}{2} \tau^2 + O(\tau^3) \\ \alpha_2 = -1 - \tau - \frac{1}{2} \tau^2 + O(\tau^3), \end{array} \right.$$

zodat

$$(1.25) \quad \left\{ \begin{array}{l} c_1 = \frac{1}{2} (1 + a) + \frac{1}{2} \tau + O(\tau^2) \\ c_2 = \frac{1}{2} (1 - a) - \frac{1}{2} \tau + O(\tau^2). \end{array} \right.$$

De oplossing u_k ontwikkeld in τ is dan

$$(1.26) \quad u_k = \left(\frac{1}{2} (1 + a) + \frac{1}{2} \tau + O(\tau^2) \right) (1 + O(\tau^2)) \cdot \exp(-t_k) \\ + (-1)^k \left(\frac{1}{2} (1 - a) - \frac{1}{2} \tau + O(\tau^2) \right) (1 + O(\tau^2)) \cdot \exp(+t_k).$$

u_k bestaat uit een dalende e-macht en een stijgende oscillerende functie. Ten gevolge van deze laatste term zal de discretiseringsfout voor willekeurige a niet naar nul convergeren als $\tau \rightarrow 0$. Uit (1.23) volgt echter dat als $a = \alpha_1$ de oscillerende functie niet optreedt en dan is de discretiseringsfout

$$(1.27) \quad \tilde{U}(t_k) - u_k = O(\tau^2).$$

Berekenen we a volgens schema (1.9) dan vinden we met behulp van (1.17)

$$a = 1 - \tau,$$

en dus

$$(1.28) \quad \tilde{U}(t_k) - u_k = (1 + (-1)^k \exp t_k) O(\tau^2).$$

Voor berekeningen over eindige intervallen is dit eveneens $O(\tau^2)$.

Stabiliteit van differentie schema's

We beschouwen nogmaals schema (1.19), maar we veronderstellen nu niet dat u_k exact berekend wordt, maar dat er afrondingsfouten gemaakt worden. We vinden dus niet u_k maar de numerieke oplossing u_k^* . De afron-

dingsfouten kunnen als een inhomogene term d_k opgevat worden, zodat u_k^* aan een schema voldoet van de vorm

$$(1.29) \quad u_{k+1}^* - u_{k-1}^* + 2\tau u_k^* = d_k.$$

De numerieke fout $u_k - u_k^*$ voldoet ook aan (1.29), maar heeft homogene beginvoorwaarden.

$$(1.30) \quad u_0 - u_0^* = u_1 - u_1^* = 0.$$

We kunnen de oplossing $u_k - u_k^*$ vinden door te stellen

$$(1.31) \quad u_k - u_k^* = C_{1,k} \alpha_1^k + C_{2,k} \alpha_2^k,$$

waarin α_1 en α_2 door (1.21) gegeven zijn.

Introduceren we de differentie-operator Δ gedefiniëerd door

$$\Delta f_k = f_{k+1} - f_k,$$

dan kunnen we (1.29) ook schrijven als

$$(1.32) \quad \Delta^2 u_{k-1}^* + 2\Delta u_{k-1}^* + 2\tau u_k^* = d_k.$$

Gebruik makend van de identiteit

$$\Delta(f_k \cdot g_k) = f_k \cdot \Delta g_k + g_{k+1} \cdot \Delta f_k,$$

vinden we

$$\Delta u_{k-1}^* = \alpha_1^{k-1} \Delta C_{1,k-1} + C_{1,k} \Delta \alpha_1^{k-1} + \alpha_2^{k-1} \Delta C_{2,k-1} + C_{2,k} \Delta \alpha_2^{k-1}.$$

We eisen nu dat

$$(1.33) \quad \alpha_1^k \Delta C_{1,k} + \alpha_2^k \Delta C_{2,k} = 0.$$

Voor vergelijking (1.32) kunnen we dan schrijven

$$\begin{aligned} & C_{1,k} \Delta^2 \alpha_1^{k-1} + (\Delta \alpha_1^k) \Delta C_{1,k} + C_{2,k} \Delta^2 \alpha_2^{k-1} + (\Delta \alpha_2^k) \Delta C_{2,k} + \\ & + 2C_{1,k} \Delta \alpha_1^{k-1} + 2C_{2,k} \Delta \alpha_2^{k-1} + 2\tau C_{1,k} \alpha_1^k + 2\tau C_{2,k} \alpha_2^k = d_k. \end{aligned}$$

Nu zijn α_1^k en α_2^k oplossingen van de homogene vergelijking (1.19), zodat we vinden

$$(1.34) \quad (\Delta \alpha_1^k) \cdot \Delta C_{1,k} + (\Delta \alpha_2^k) \cdot \Delta C_{2,k} = d_k.$$

Combinatie van (1.33) en (1.34) geeft

$$\Delta C_{1,k} = \frac{-d_k \alpha_2^k}{\alpha_1^k \cdot \Delta \alpha_2^k - \alpha_2^k \cdot \Delta \alpha_1^k} = \frac{d_k \alpha_1^{-k}}{\alpha_1^{-\alpha_2}}$$

en

$$\Delta C_{2,k} = \frac{-d_k \alpha_1^{-k}}{\alpha_1^{-\alpha_2}}.$$

Dus

$$(1.35) \quad \begin{cases} C_{1,k} = C_{1,0} + \sum_{i=0}^{k-1} \frac{d_i \alpha_1^{-i}}{\alpha_1^{-\alpha_2}}, \\ C_{2,k} = C_{2,0} - \sum_{i=0}^{k-1} \frac{d_i \alpha_2^{-i}}{\alpha_1^{-\alpha_2}}. \end{cases}$$

Blijkens (1.30) moet nu gelden

$$C_{1,0} + C_{2,0} = 0,$$

$$(C_{1,0} + \frac{d_0}{\alpha_1^{-\alpha_2}}) \alpha_1 + (C_{2,0} - \frac{d_0}{\alpha_1^{-\alpha_2}}) \alpha_2 = 0.$$

Aangezien $d_0 = 0$, volgt $C_{1,0} = C_{2,0} = 0$, dus

$$(1.36) \quad u_k - u_k^* = \sum_{i=1}^{k-1} \frac{d_i \alpha_1^{-i}}{\alpha_1^{-\alpha_2}} \alpha_1^k - \sum_{i=0}^{k-1} \frac{d_i \alpha_2^{-i}}{\alpha_1^{-\alpha_2}} \alpha_2^k.$$

Als voor alle k geldt $d_k = d$, dan is

$$u_k - u_k^* = \frac{d}{\alpha_1^{-\alpha_2}} \left\{ \alpha_1 \frac{1 - \alpha_1^{k-1}}{1 - \alpha_1} - \alpha_2 \frac{1 - \alpha_2^{k-1}}{1 - \alpha_2} \right\}.$$

Met behulp van (1.24) volgt tenslotte

$$(1.37) \quad u_k - u_k^* = \frac{1}{2} d \left[\frac{1-\tau}{\tau} (1 - e^{-t_k}) + \frac{1+\tau}{2+\tau} (1 + (-1)^k e^{t_k}) + O(t_k) \right].$$

In tegenstelling tot de discretiseringsfout (gegeven door (1.28)), welke $O(\tau^2)$ was, kan de numerieke fout niet zo klein gemaakt worden als men wil door τ maar klein te kiezen. Men moet er zelfs voor zorgen dat $\tau \gg d$.

Men kan de numerieke fout wel verkleinen door de rekenprecisie te verhogen ($d \rightarrow 0$). De conclusie is dus dat de totale fout $\tilde{u}_k - u_k^*$ over een begreind interval $0 \leq t \leq T$ zo klein gemaakt kan worden als men wil door τ en d maar klein te maken. We noemen schema (1.19) stabiel voor berekeningen over een eindig tijdsinterval, maar instabiel voor berekeningen over een oneindig tijdsinterval.

We ontleen aan Godunov en Rjabenki [4] het volgende differentieschema voor vergelijking (1.1), waarin weer $F(t) \equiv 0$.

$$(1.38) \quad 4 \frac{u_{k+1} - u_{k-1}}{2\tau} - 3 \frac{u_{k+1} - u_k}{\tau} + u_k = 0.$$

De analytische oplossing $\tilde{U}(t)$ voldoet aan het schema

$$(1.39) \quad 4 \frac{\tilde{u}_{k+1} - \tilde{u}_{k-1}}{2\tau} - 3 \frac{\tilde{u}_{k+1} - \tilde{u}_k}{\tau} + \tilde{u}_k = -\frac{3}{2} \tilde{U}_{tt}(t_k + \theta_k \tau) \cdot \tau + \frac{2}{3} \tilde{U}_{ttt}(t_k + \bar{\theta}_k \tau) \cdot \tau^2,$$

waarin $0 \leq \theta_k \leq 1$ en $0 \leq \bar{\theta}_k \leq 1$ is. De nauwkeurigheid is blijkbaar van de eerste orde in τ . We stellen weer $u_k = \alpha^k$ en vinden voor α de vergelijking

$$\alpha^2 - (3 + \tau)\alpha + 2 = 0$$

met oplossingen

$$(1.40) \quad \begin{cases} \alpha_1 = \frac{1}{2} \left[3 + \tau - \sqrt{1 + 6\tau + \tau^2} \right] = 1 - \tau + 2\tau^2 + O(\tau^3) \\ \alpha_2 = \frac{1}{2} \left[3 + \tau + \sqrt{1 + 6\tau + \tau^2} \right] = 2 + 2\tau - 2\tau^2 + O(\tau^3). \end{cases}$$

De algemene oplossing is van de vorm

$$u_k = c_1 \alpha_1^k + c_2 \alpha_2^k.$$

Stellen we $u_0 = 1$ en $u_1 = a$ dan worden c_1 en c_2 volgens (1.23)

$$(1.41) \quad \begin{aligned} c_1 &= (2-a) + (3a-4)\tau + O(\tau^2) \\ c_2 &= (a-1) + (3a-2)\tau + O(\tau^2). \end{aligned}$$

Voor u_k vinden we

$$(1.42) \quad \begin{aligned} u_k &= (2 - a + (3a-4)\tau + O(\tau^2))(1 + O(\tau) \cdot t_k) \cdot \exp(-t_k) + \\ &\quad (a - 1 + (3a-2)\tau + O(\tau^2))(1 + O(\tau) \cdot t_k) \cdot 2^k \exp t_k. \end{aligned}$$

Kiezen we voor a weer $1-\tau$, dan wordt de discretiseringsfout

$$(1.43) \quad \tilde{u}_k - u_k = 2\tau(1 + O(\tau^2))\exp(-t_k) + O(\tau^2)2^k \exp t_k.$$

Het is blijkbaar niet mogelijk de fout te verkleinen door τ naar nul te laten gaan, want k gaat dan naar oneindig in elk punt $t_k = \text{constant} \neq 0$ en de discretiseringsfout wordt van de orde 2^k . Alleen wanneer a exact gelijk aan α_1 gekozen wordt en er geen afrondingsfouten in de verdere berekeningen optreden, is de discretiseringsfout $O(\tau)$. In de praktijk is dit niet te verwezenlijken. Stel dat er een fout d gemaakt wordt in a , dus $a = \alpha_1 + d$, dan is

$$(1.44) \quad \tilde{u}_k - u_k = 2^k \cdot d \cdot \exp t_k,$$

zodat voor vaste t_k en $\tau \rightarrow 0$ de discretiseringsfout naar oneindig gaat. We noemen schema (1.38) instabiel voor berekeningen over een willekeurig tijdsinterval $0 \leq t \leq T$.

Partiële differentiaalvergelijkingen

Tot dusver zijn steeds gewone differentiaalvergelijkingen beschouwd. We zullen nu wat nader ingaan op partiële differentiaalvergelijkingen.

Voorbeelden van partiële differentiaalvergelijkingen zijn de diffusie-vergelijking

$$(1.45) \quad U_t = U_{xx},$$

de golf-vergelijking

$$(1.46) \quad U_{tt} = U_{xx},$$

en de Laplace-vergelijking

$$(1.47) \quad U_{xx} + U_{yy} = 0.$$

De vergelijkingen zijn resp. van het parabolische, het hyperbolische en het elliptische type. In het tweede deel van deze syllabus zullen deze of soortgelijke vergelijkingen uitvoerig behandeld worden. We volstaan in deze inleiding met enkele algemene opmerkingen.

In het (x, t) vlak kiezen we de punten $(x_j, t_k) = (j \xi, k\tau)$, waarin ξ en τ positieve getallen zijn, j de waarden $0, 1, 2, \dots, m$ en k de waarden $0, 1, 2, \dots, N$ doorloopt. We geven het getal $U(x_j, t_k)$ aan met $u_{j,k}$.

De diffusie-vergelijking

We beschouwen het schema

$$(1.48) \quad \frac{u_{j,k+1} - u_{j,k}}{\tau} - \frac{u_{j+1,k} - 2u_{j,k} + u_{j-1,k}}{2\xi} = 0.$$

Dit wordt het klassieke expliciete differentieschema voor de diffusie-vergelijking genoemd. De analytische oplossing $\tilde{U}(x, t)$ voldoet aan het schema

$$(1.49) \quad \frac{\tilde{u}_{j,k+1} - \tilde{u}_{j,k}}{\tau} - \frac{\tilde{u}_{j+1,k} - 2\tilde{u}_{j,k} + \tilde{u}_{j-1,k}}{2\xi} = \frac{1}{2} \tilde{U}_{tt}(x_j, t_k + \theta_k \tau) \cdot \tau - \frac{1}{12} \tilde{U}_{xxxx}(x_j + \bar{\theta}_j \xi, t_k) \cdot \xi^2,$$

waarin $0 \leq \theta_k \leq 1$ en $0 \leq \bar{\theta}_j \leq 1$. De benaderingsfout is dus $O(\tau) + O(\xi^2)$, zodat (1.48) consistent is met (1.45).

Schema (1.48) kan geschreven worden als

$$(1.50) \quad u_{j,k+1} = ru_{j-1,k} + (1-2r)u_{j,k} + ru_{j+1,k},$$

waarin $r = \tau/\xi^2$. Stel dat $u_{j,0}$, $u_{0,k}$ en $u_{m,k}$ voor alle j en k gegeven zijn, dan is $u_{j,k}$ in alle punten (j,k) met (1.50) te berekenen.

We gaan na wat er kan gebeuren wanneer men in (1.50) $r > \frac{1}{2}$ kiest.

Stel dat we als beginfunctie kiezen

$$u_{j,0} = \delta_{i,0} \varepsilon,$$

waarin we de j waarden uitbreiden tot $j = 0, \pm 1, \pm 2, \dots, \pm m$ en waarin δ de Kronecker delta is. Het is eenvoudig met behulp van (1.50) in te zien dat als $r > \frac{1}{2}$ de tekens van $u_{j+1,k}$, $u_{j,k}$ en $u_{j-1,k}$ en de tekens van $u_{j,k}$ en $u_{j,k+1}$ voor $k \geq 1$ en $-k \leq j \leq k$ alterneren. Met deze eigenschap bewijzen we de relatie

$$\sum_{j=-(k+1)}^{k+1} |u_{j,k+1}| = (4r-1) \sum_{j=-k}^k |u_{j,k}|.$$

Immers

$$\begin{aligned} \sum_{j=-(k+1)}^{k+1} |u_{j,k+1}| &= u_{-k-1,k+1} - u_{-k,k+1} + u_{-k+1,k+1} - \dots = \\ &= ru_{-k,k} - (1-2r)u_{-k,k} + ru_{-k,k} \\ &+ ru_{-k,k} + (1-2r)u_{-k+1,k} + ru_{-k+2,k} + \dots \\ &= (4r-1) \sum_{j=-k}^k |u_{j,k}|. \end{aligned}$$

Hieruit volgt

$$\sum_{j=-k}^k |u_{j,k}| = (4r-1)^k \varepsilon.$$

Dus

$$\max_{-k \leq j \leq k} |u_{j,k}| \geq (2k+1)^{-1} \sum_{j=-k}^k |u_{j,k}| = \frac{(4r-1)^k}{2k+1} \varepsilon$$

en zeker

$$(1.51) \quad |||u_{j,k}||| \geq \frac{(4r-1)^k}{2k+1} \varepsilon,$$

waarin $|||$ de maximum norm van $u_{j,k}$ voor constante k voorstelt.

We zien dat voor $k \rightarrow \infty$ en $r > \frac{1}{2}$ geen goede resultaten te verwachten zijn.

We noemen het schema instabiel voor $r > \frac{1}{2}$.

De golf-vergelijking

We beschouwen het schema

$$(1.52) \quad \frac{u_{j,k+1} - 2u_{j,k} + u_{j,k-1}}{\tau^2} - \frac{u_{j+1,k} - 2u_{j,k} + u_{j-1,k}}{\xi^2} = 0$$

waarvoor de benaderingsfout is

$$(1.53) \quad \frac{1}{12} (\tilde{U}_{tttt}(x_j, t_k + \theta_k \tau) \cdot \tau^2 - \tilde{U}_{xxxx}(x_j + \bar{\theta}_j \xi, t_k) \cdot \xi^2)$$

waarin θ_k en $\bar{\theta}_j$ weer tussen 0 en 1 liggen. Deze fout is van de orde 2 in τ en ξ . $u_{j,k}$ berekenen we volgens de formule

$$(1.54) \quad u_{j,k+1} = ru_{j-1,k} + 2(1-r)u_{j,k} + ru_{j+1,k} - u_{j,k-1},$$

waarin $r = \tau/\xi$. De waarden $u_{0,k}$, $u_{m,k}$, $u_{j,0}$ en $u_{j,1}$ vormen de rand- en beginvoorwaarden van het probleem.

De waarde van r wordt weer bepaald door stabiliteitsoverwegingen.

Karakteristiekencriterium

Uit de theorie van hyperbolische differentiaalvergelijkingen ontle-
nen we een eigenschap, welke ook van belang is voor differentieschema's,
welke hyperbolische vergelijkingen benaderen. We lichten deze eigenschap
toe aan de golf-vergelijking.

De karakteristieken van de golf-vergelijking worden gegeven door

$$x \pm t = \text{constant},$$

zodat de karakteristieken door het punt (x_j, t_k) gegeven worden door

$$x \pm t = x_j \pm t_k.$$

Er geldt nu de eigenschap dat $U(x_j, t_k)$ afhangt van de beginwaarden op het interval

$$x_j - t_k \leq x \leq x_j + t_k,$$

het zogenaamde afhankelijkheidsgebied van het punt (x_0, t_0) .

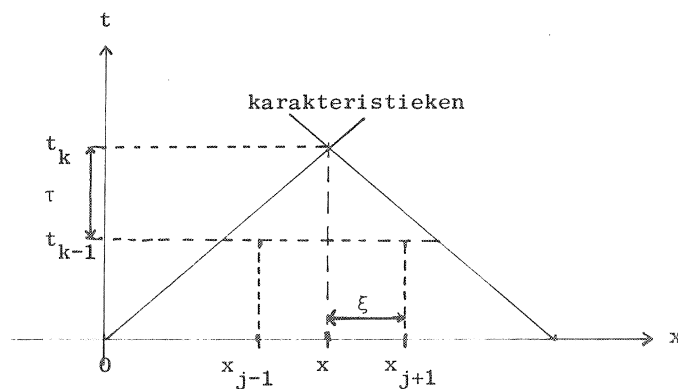


fig. 1.1

Karakteristiekencriterium

Voor het differentieschema betekent dit dat u_{j_0, k_0} berekend wordt uit waarden $u_{j, k}$ welke gegeven zijn in punten (j, k) binnen of op de karakteristieken door het punt (j_0, k_0) (zie figuur 1.1). Toepassing op schema (1.52) geeft

$$\frac{\tau}{\xi} = r \leq 1.$$

Laplace-vergelijking

Het oplossen van elliptische vergelijkingen, zoals de vergelijking van Laplace, leidt tot zuivere randwaardeproblemen, in tegenstelling tot de vergelijkingen van het parabolische en hyperbolische type, welke tot begin-randwaardeproblemen leiden. In het tweede deel van deze syllabus

zullen we zien dat dergelijke zuivere randwaardeproblemen te formuleren zijn als begin-randwaardeproblemen voor vergelijkingen van het parabolische of hyperbolische type. We zullen ons dus in dit deel uitsluitend bezig houden met begin-randwaardeproblemen.

2. Differentieschema's voor gewone differentiaalvergelijkingen

In dit hoofdstuk worden de begrippen, die in het eerste hoofdstuk aan de hand van voorbeelden zijn toegelicht, voor gewone lineaire differentiaalvergelijkingen gedefiniëerd. Verder worden er enkele stellingen gegeven.

Benadering van de differentiaalvergelijking

We beschouwen een lineaire ruimte E van vectorfuncties U , welke afhangen van een variabele t uit het reële interval $[0, T]$. T is een vast gekozen getal. Alle vectoren $U(t_k)$, waarin t_k een getal uit het interval $[0, T]$ is, worden verondersteld in een Euclidische vectorruimte E_0 te liggen. We noteren $U(t_k)$ met U_k .

We zullen differentiaalvergelijkingen beschouwen voor onbekende vectorfuncties $U \in E$ met beginvoorwaarden $U(0) = U_0 \in E_0$ en inhomogene termen $F \in E$. Bovendien zullen deze vergelijkingen lineair zijn. Bijvoorbeeld de in de inleiding behandelde vergelijking

$$U_t + U = F, \quad 0 < t \leq T,$$

$$U = U_0, \quad t = 0.$$

Hiervoor is E een ruimte van scalarfuncties en E_0 de ruimte der reële of complexe getallen. We kunnen deze vergelijking opvatten als een lineaire afbeelding L van $U \in E$ op het paar $(U_0, F) \in E_0 * E$, dus

$$(2.1) \quad LU = (U_0, F).$$

Alle hier beschouwde differentiaalvergelijkingen zullen afbeeldingen van het type (2.1) zijn. We geven het definitiegebied van L met D_L aan, het beeldgebied met Δ_L en we kiezen $E_0 * E = \Delta_L$.

We normeren de ruimten E_0 , E en $E_0 * E$ en geven de normen aan met

$$||U_0||, ||U||, ||(U_0, F)||.$$

Voor $||(U_0, F)||$ kunnen we kiezen

$$(2.2) \quad ||(U_0, F)|| = ||U_0|| + ||F||.$$

Uit de norm-eigenschappen van $||U_0||$ en $|||F|||$ volgt direct dat (2.2) inderdaad een norm voor de elementen van $E_0^* \times E$ definiëert.

Er zal steeds verondersteld worden dat L een éénduidige inverse bezit, welke continu is in (U_0, F) , dus

$$(2.3) \quad ||(U_0, F) - (U'_0, F')|| \rightarrow 0 \implies |||U - U'||| \rightarrow 0.$$

We gaan nu over op het differentie-analogon van de differentiaal-vergelijking (2.1). Daartoe definiëren we eerst een rooster Q van $N+1$ punten t_k , $k = 0, 1, \dots, N$, met $0 = t_0 < t_1 < t_2 < \dots < t_N = T$ en een discretiseringsoperator $[]_d$, welke aan een op $[0, T]$ gedefiniëerde functie U de waarden $U_k = U(t_k)$ toevoegt. De verkregen roosterfunctie noteren we met u , dus

$$u = [U]_d.$$

u wordt ook wel de discretisering van U genoemd. De discretisering van de functies $U \in E$ liggen in een ruimte E^* . De vectoren in de roosterpunten t_k van Q liggen in E_0 en vallen samen met de vectoren $U(t_k)$. Om alvast dezelfde notatie te gebruiken, welke voor partiële differentiaal-vergelijkingen ingevoerd zal worden, zullen we de vectoren in de punten t_k met u_k aangeven en de ruimte, waartoe ze behoren met E_0^* . De ruimten E^* en E_0^* normeren we met normen $|||u|||$ en $||u_k||$, waarbij we eisen

$$(2.4) \quad |||u||| \rightarrow |||U|||, \quad ||u_k|| \rightarrow ||U_k|| \text{ als } \tau_k \rightarrow 0 \text{ voor alle } k=1, 2, \dots, N.$$

Hierin is τ_k gedefiniëerd als $\tau_k = t_k - t_{k-1}$. In het vervolg laten we de toevoeging "voor alle k " weg. Het Cartesisch product $E_0^* \times E^*$ normeren we door middel van

$$(2.5) \quad ||(u_0, f)|| = ||u_0|| + |||f|||.$$

We zijn nu in staat differentieschema's te definiëren en wel in analogie met (2.1) zullen die zijn lineaire afbeeldingen R van $u \in E^*$ op $(u_0, f) \in E_0^* \times E^*$, dus

$$(2.6) \quad Ru = (u_0, f).$$

Het schema (1.10) bijvoorbeeld, in de vorm

$$au_k + bu_{k-1} = f_{k-1}, \quad k = 1, 2, \dots, N,$$

$$u_k = u_0, \quad k = 0,$$

laat zien, dat door lineaire bewerkingen op u toe te passen het paar (u_0, f) verkregen wordt.

We eisen steeds dat R uit algebraïsche operaties bestaat en dat R een éénduidige inverse bezit. Deze inverse vinden is dan een algebraïsche kwestie geworden, in tegenstelling tot het vinden van L^{-1} dat in de analyse thuishoort. In het vervolg veronderstellen we R^{-1} expliciet bekend.

Tot dusver is het differentieschema formeel ingevoerd. Het doel is echter een differentieschema te vinden, zodanig dat de differentieoplossing u van (2.6) voor $\tau_k \rightarrow 0$ naar de oplossing $\tilde{u}$ van (2.1) convergeert:

$$(2.7) \quad |||\tilde{u} - u||| \rightarrow 0 \quad \text{als} \quad \tau_k \rightarrow 0.$$

We illustreren dit met de volgende figuur.

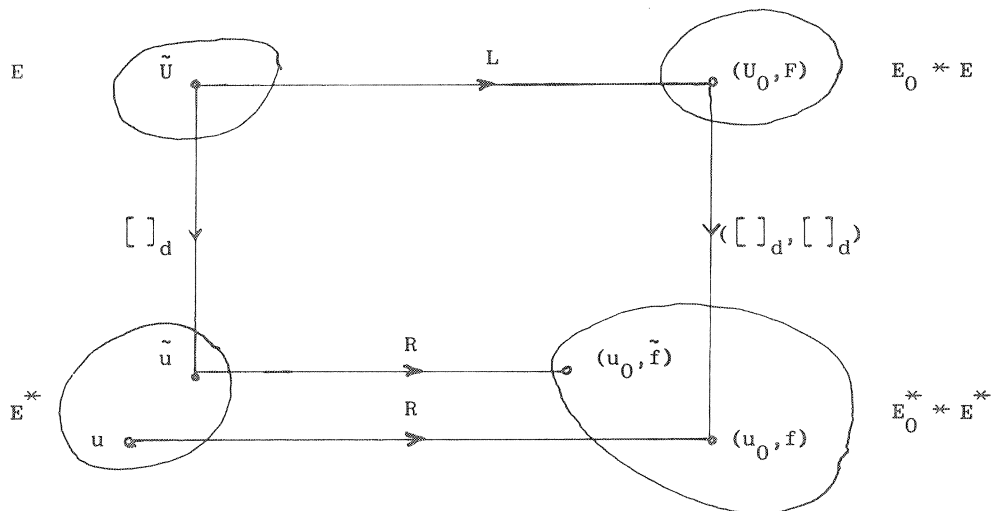


fig. 2.1

Uitgaande van de gegevens (U_0, F) vinden we enerzijds door toepassing van L^{-1} de exacte oplossing $\tilde{u}$ (waarbij we opmerken dat $U_0 = \tilde{u}_0$), anderzijds vinden we na discretisering van (U_0, F) tot (u_0, f) en toepassing van de differentieoperator R^{-1} , de differentieoplossing u . In de ruimte E^* kunnen

de analytische oplossing en de differentieoplossing vergeleken worden, dus we discretiseren $\tilde{U}$ tot $\tilde{u}$. Stel dat $\tilde{u}$ oplossing is van het schema

$$(2.8) \quad R\tilde{u} = (u_0, \tilde{f}),$$

dan geldt (zie figuur 2.1)

$$(2.9) \quad \begin{aligned} \tilde{u} - u &= R^{-1}(u_0, \tilde{f}) - R^{-1}(u_0, f) \\ &= R^{-1}((u_0, \tilde{f}) - (u_0, f)) \\ &= R^{-1}(R\tilde{u} - [L\tilde{U}]_d). \end{aligned}$$

Wanneer we voor R^{-1} definiëren de norm

$$(2.10) \quad ||R^{-1}|| = \sup_{\tilde{U} \in C} \frac{||R^{-1}(R\tilde{u} - [L\tilde{U}]_d)||}{||R\tilde{u} - [L\tilde{U}]_d||},$$

waarin C de klasse van de beschouwde functies $\tilde{U}$ is, dan vinden we voor het verschil $||\tilde{u} - u||$:

$$(2.11) \quad ||\tilde{u} - u|| \leq ||R^{-1}|| \cdot ||R\tilde{u} - [L\tilde{U}]_d||.$$

Naar aanleiding van deze ongelijkheid geven we een aantal definities.

Definitie 2.1

De waarde van $||R\tilde{u} - [L\tilde{U}]_d||$ is de benaderingsfout van het schema $Ru = (u_0, f)$ voor de vergelijking $LU = (U_0, F)$.

Definitie 2.2

De waarde van $||\tilde{u} - u||$ is de discretiseringsfout van de differentieoplossing.

Definitie 2.3

Een differentieschema is consistent met een differentiaalvergelijking over een klasse $C \subset D_L \cap E$, wanneer voor alle $\tilde{U} \in C$ de benaderingsfout van het differentieschema voor $\tau_k \rightarrow 0$ naar nul gaat.

Definitie 2.4

De differentieoplossingen convergeren over de klasse $C \subset D_L \cap E$ naar de analytische oplossingen $\tilde{U} \in C$, wanneer voor alle $\tilde{U} \in C$ de discretiseringsfout van de differentieoplossingen voor $\tau_k \rightarrow 0$ naar nul gaat.

Uit (2.11) volgt nu dat de differentieoplossingen naar de analytische oplossing convergeren, wanneer het rechterlid voor $\tau_k \rightarrow 0$ naar nul convergeert; dit hangt dus af van $\|R^{-1}\|$ en de benaderingsfout. We zullen eerst de benaderingsfout wat nader beschouwen en wel aan de hand van het volgende voorbeeld:

Voorbeeld 2.1

Beschouw het al eerder genoemde beginwaarde-probleem

$$\tilde{U}_t + \tilde{U} = F, \quad 0 < t \leq T,$$

$$\tilde{U} = U_0, \quad t = 0.$$

We kiezen hierbij een rooster Q met punten $t_k = k \frac{T}{N}$, $k = 0, \dots, N$ en een differentieschema $Ru = (u_0, f)$ dat beschreven wordt door de vergelijkingen

$$\alpha u_{k+1} + \beta u_k + \gamma u_{k-1} = f_k, \quad k = 1, 2, \dots, N,$$

$$u_0 = [U_0]_d, \quad u_1 = [\tilde{U}_1]_d + O(\tau^p).$$

Hierin is $p > 0$ en $\tau = T/N$. De waarde van u_1 kan verkregen worden door met een eerste orde schema met $\bar{\tau} = \tau^p$, u op $t = \tau$ uit te rekenen (zie inleiding).

We willen α , β en γ zodanig kiezen, dat de benaderingsfout bij een gegeven τ zo klein mogelijk is. Daartoe substitueren we $\tilde{u}$ in het differentieschema en ontwikkelen met behulp van de Taylor-formule, waarbij we veronderstellen dat D_L uit oneindig aantal malen differentiëerbare functies bestaat. Dus

$$(2.12) \quad \alpha(1 + \tau \frac{\partial}{\partial t} + \frac{1}{2} \tau^2 \frac{\partial^2}{\partial t^2} + \dots) \tilde{u}_k + \beta \tilde{u}_k + \\ + \gamma(1 - \tau \frac{\partial}{\partial t} + \frac{1}{2} \tau^2 \frac{\partial^2}{\partial t^2} - \dots) \tilde{u}_k = \tilde{f}_k,$$

waarin $k = 1, 2, \dots, N$ en $\frac{\partial^n}{\partial t^n} \tilde{u}_k = \left(\left[\frac{\partial^n}{\partial t^n} \tilde{U} \right]_{d k} \right)$.

Hieruit volgt

$$(2.13) \quad \tilde{f}_k - f_k = (\alpha + \beta + \gamma - 1) \tilde{u}_k + (\alpha - \gamma - \frac{1}{\tau}) \tau \frac{\partial}{\partial t} \tilde{u}_k \\ + \frac{1}{2} (\alpha + \gamma) \tau^2 \frac{\partial^2}{\partial t^2} \tilde{u}_k + \frac{1}{3!} (\alpha - \gamma) \tau^3 \frac{\partial^3}{\partial t^3} \tilde{u}_k + \dots$$

De benaderingsfout zal blijken af te hangen van de klasse C, waartoe we de functies $\tilde{U}$ beperken. In het bijzonder geldt dit wanneer C al of niet bevat is in de ruimte N_L . De ruimte N_L bestaat uit alle $\tilde{U}$ met

$$L\tilde{U} = (U_0, 0),$$

dus N_L bevat de oplossingen van de homogene differentiaalvergelijking.

(a) C is uitgebreider dan N_L

We stellen

$$(2.14) \quad \begin{cases} \alpha + \beta + \gamma - 1 = 0, \\ \alpha - \gamma - \frac{1}{\tau} = 0, \end{cases}$$

waarmee (2.13) overgaat in

$$(2.13a) \quad \tilde{f}_k - f_k = \frac{1}{2} (2^{\alpha\tau} - 1) \tau \frac{\partial^2}{\partial t^2} \tilde{u}_k + O(\tau^2), \quad k = 2, 3, \dots, N.$$

Kiezen we $\alpha = O(\tau^{-1})$ en $p = 1$ dan hebben we een hele klasse van differentieschema's verkregen welke over C een benaderingsfout van $O(\tau)$ hebben. Kiezen we echter $2^{\alpha\tau} - 1 = O(\tau)$ en $p = 2$, dan is de benaderingsfout $O(\tau^2)$. Dit geval komt voor $2^{\alpha\tau} - 1 = 0$ overeen met schema (1.19). We hebben dus

$$||\tilde{R}u - [L\tilde{U}]_d|| = O(\tau) \text{ over } C \text{ als } \alpha = O(\tau^{-1}) \text{ en } p \geq 1$$

$$||\tilde{R}u - [L\tilde{U}]_d|| = O(\tau^2) \text{ over } C \text{ als } \alpha = \frac{1}{2\tau} + O(1) \text{ en } p \geq 2.$$

(b) C ligt in N_L

We mogen nu gebruik maken van het feit dat $\tilde{U}_t = -\tilde{U}$. Uit (2.12) volgt dan

$$(2.13b) \quad \tilde{f}_k - f_k = \tilde{f}_k = ((\alpha + \beta + \gamma - 1) - (\alpha - \gamma - \frac{1}{\tau})\tau + \frac{1}{2}(\alpha + \gamma)\tau^2 - \frac{1}{3!}(\alpha - \gamma)\tau^3 + \dots)\tilde{u}_k.$$

In dit eenvoudige geval is het mogelijk een keuze voor α , β en γ te maken, zodanig dat de coëfficiënt van $\tilde{u}_k$ nul wordt nl.

$$\alpha = 1, \quad \beta = -e^{-\tau}, \quad \gamma = 0.$$

Het differentieschema wordt dan beschreven door

$$u_{k+1} = e^{-\tau} u_k.$$

Dit schema levert de analytische oplossing van de homogene vergelijking.

Ook echter met de onder (a) gedane keuze voor α , β en γ is er een schema te construeren, dat de differentiaalvergelijking exact oplost. Substitutie van de relaties (2.14) in (2.13b) geeft

$$\tilde{f}_k = \left(\frac{2\alpha\tau - 1}{2!} - \frac{1}{3!}\tau^2 + \frac{2\alpha\tau - 1}{4!}\tau^3 - \frac{1}{5!}\tau^4 + \dots \right) \tilde{u}_k.$$

De benaderingsfout voor de homogene vergelijking is nul wanneer we kiezen

$$2\alpha\tau - 1 = \frac{\frac{\tau^2}{3!} + \frac{\tau^4}{5!} + \dots}{\frac{\tau}{2!} + \frac{\tau^3}{4!} + \dots} = \frac{\sinh \tau - \tau}{\cosh \tau - 1}.$$

Definiëren we $\rho = 2\alpha\tau - 1$ en drukken we met behulp van (2.14) β en

γ in ρ uit, dan is de conclusie voor dit voorbeeld, dat het schema

$$(2.15) \quad (\rho+1)u_{k+1} + 2(\tau-\rho)u_k + (\rho-1)u_{k-1} = 2\tau f_k$$

voor

$$(2.16) \quad \rho = \frac{\sinh \tau - \tau}{\cosh \tau - 1}$$

de homogene differentiaalvergelijking exact oplost, en de inhomogene vergelijking tot op $O(\tau^2)$ benadert.

Stabiliteit van differentieschema's

In het volgende zullen we aandacht besteden aan de operator R^{-1} . Daartoe beschouwen we eerst de vergelijking

$$L\tilde{U} = (U_0, F).$$

Er is verondersteld dat de oplossing $\tilde{U}$ continu afhangt van (U_0, F) . Wanneer $\tilde{U}$ de toestand van een systeem beschrijft, U_0 de begintoestand is en F een uitwendige beïnvloeding voorstelt, dan betekent deze continuïteit dat een verstoring van U_0 en F een toestandsverandering van de toestand $\tilde{U}$ tot gevolg heeft, welke zo klein gemaakt kan worden als men wil door er voor te zorgen dat de verstoringen voldoende klein zijn. Men zou de toestand $\tilde{U}$ stabiel kunnen noemen. Het ligt voor de hand differentieoplossingen u , of het schema $Ru = (u_0, f)$, stabiel te noemen wanneer u continu is in het paar (u_0, f) . We zullen eisen dat deze continuïteit uniform geldig is voor $\tau_k \rightarrow 0$. (Voor eindige waarden van T en vaste τ_k 's, is hier in het algemeen wel aan voldaan, omdat R dan uit een eindig aantal algebraïsche bewerkingen bestaat.)

Definitie 2.5

Het differentieschema $Ru = (u_0, f)$ is over het interval $[0, T]$ uniform stabiel t.o.v. het paar (u_0, f) , wanneer u continu is in $(u_0, f) \in E_0^{**} E$ en wel uniform op $0 < \tau_k \leq \bar{\tau}$.

In deze vorm wordt de stabiliteit door Rjabenki en Filippow [8] gedefinieerd. De toevoeging uniform laten we in het vervolg weg en we zullen spreken van R-F stabiliteit.

Men spreekt ook wel van R-F stabiliteit t.o.v. de beginvoorwaarden of t.o.v. de inhomogene term f , wanneer men continuïteit in $(u_0, 0)$ resp. $(0, f)$ bedoelt. We merken op dat ten gevolge van de norm definitie (2.5) voor (u_0, f) deze beide laatste vormen van stabiliteit tezamen de stabiliteit t.o.v. (u_0, f) garanderen.

Een equivalente, maar meer concrete stabiliteitsdefinitie verkrijgen we door gebruik te maken van de stelling, dat continue lineaire operatoren begrensd zijn. Nu is R lineair, dus ook R^{-1} is lineair en bovendien continu, dus $\|R^{-1}\|$ is begrensd en wel uniform in τ_k . Uit de norm-definitie (2.5) volgt dan voor de R-F stabiliteit t.o.v. (u_0, f)

$$(2.17) \quad \|u\| \leq \|R^{-1}\| \cdot \|(u_0, f)\| \leq \text{constante} \cdot (\|u_0\| + \|f\|),$$

waarin de constante uniform begrensd blijft voor $\tau_k \rightarrow 0$.

In deze vorm wordt de stabiliteit in Godunov en Rjabenki [4] gedefiniëerd.

We kunnen weer stabiliteit t.o.v. de beginvoorwaarden en t.o.v. de inhomogene term onderscheiden, wanneer de operator R^{-1} uniform begrensd is voor $\tau_k \rightarrow 0$ voor de elementen $(u_0, 0)$ resp. $(0, f)$.

Convergentie van de differentieoplossingen

Uit (2.11) volgt direct dat een consistent differentieschema dat R-F stabiel is t.o.v. de inhomogene term, een naar de analytische oplossing convergerende differentieoplossing heeft. Er geldt dus

Stelling 2.1

Een consistent differentieschema, dat R-F stabiel is t.o.v. de inhomogene term, heeft een oplossing welke voor $\tau_k \rightarrow 0$ naar de analytische oplossing convergeert.

We merken op dat de stabiliteitsvoorwaarde niet noodzakelijk is voor convergentie, omdat de verstoring, welke de benaderingsfout vormt van een gegeven orde in τ_k is. Wanneer $\|R^{-1}\|$ nu maar van een lagere orde in τ_k^{-1} is, kan men de toestandsverandering, zoals de discretiseringsfout opgevat kan worden, toch wel zo klein maken als men wil door τ_k voldoende klein te kiezen. De stabiliteitsvoorwaarde is echter niet alleen

gesteld om convergentie te verzekeren, maar ook om een zekere ongevoeligheid te verkrijgen voor verstoringen van een ander soort dan benaderingsfouten. Hierop gaan we later nog nader in.

Differentieschema's van de vorm $u_{k+1} = A_k u_k + a \tau f_k$

We zullen wat nader ingaan op schema's van de vorm

$$(2.18) \quad u_{k+1} = A_k u_k + a \tau f_k, \quad k = 0, 1, \dots, N-1.$$

u_0 wordt weer bekend verondersteld. De matrices A_k zijn de z.g. amplificatiematrices en a is een constante. De tot dusver besproken voorbeelden waren alle van dit type of als zodanig te schrijven. Bijvoorbeeld het eerste orde schema (1.9) is te schrijven als

$$u_{k+1} = (1 - \tau)u_k + \tau f_k,$$

en het tweede orde schema (2.15), waarvan de schema's (1.19) en (1.38) bijzondere gevallen zijn, is equivalent met

$$u_{k+1} = 2 \frac{\rho-1}{\rho+1} u_k + \frac{1-\rho}{1+\rho} b v_k + \frac{2\tau}{\rho+1} f_k$$

$$v_{k+1} = \frac{1}{b} u_k.$$

Hierin worden $\rho \neq -1$ en $b \neq 0$ verondersteld en zijn u_0 en v_0 de beginvoorwaarden. In matrixvorm krijgen we

$$(2.19) \quad \begin{pmatrix} u_{k+1} \\ v_{k+1} \end{pmatrix} = \begin{pmatrix} 2 \frac{\rho-1}{\rho+1} & b \frac{1-\rho}{1+\rho} \\ \frac{1}{b} & 0 \end{pmatrix} \begin{pmatrix} u_k \\ v_k \end{pmatrix} + \frac{2}{\rho+1} \tau \begin{pmatrix} f_k \\ 0 \end{pmatrix},$$

en dit is van de vorm (2.18).

Voor schema's van het type (2.18) is eenvoudig een bovengrens voor de norm van R^{-1} af te leiden. Daartoe voeren we een functie $\alpha(\tau)$ in, die het maximum is van alle normen van de amplificatie-matrices, dus

$$(2.20) \quad \alpha(\tau) = \max_{0 \leq k \leq N-1} \|A_k\|.$$

Verder zullen we de normen $||| \cdot |||$ en $||| \cdot |||$ met elkaar moeten kunnen vergelijken. In het vervolg zullen we kiezen:

$$(2.21) \quad |||u||| = \max_{0 \leq k \leq N} ||u_k||.$$

Stelling 2.2

Op de verzameling $\{(u_0, 0) \mid u_0 \in E_0^*\}$ wordt R^{-1} begrensd door $\max_{0 \leq k \leq N} (\alpha^k)$ en op de verzameling $\{(0, f) \mid f \in E^*\}$ wordt R^{-1} begrensd door $\max_{0 \leq k \leq N} (|a| \tau \frac{1-\alpha^k}{1-\alpha})$.

Bewijs

Uit (2.18) volgt

$$u_{k+1} = A_k A_{k-1} \dots A_0 u_0 + a \tau (f_k + \sum_{n=1}^k A_k A_{k-1} \dots A_{k-n+1} f_{k-n}),$$

dus volgens (2.20) en (2.21) geldt

$$\begin{aligned} ||u_{k+1}|| &\leq \alpha^{k+1} ||u_0|| + |a| \tau \sum_{n=0}^k \alpha^n ||f_{k-n}|| \\ &\leq \alpha^{k+1} ||u_0|| + |a| \tau \frac{1-\alpha^k}{1-\alpha} |||f|||. \end{aligned}$$

In het bijzonder geldt

$$|||u||| \leq \max_{0 \leq k \leq N-1} (||u_0||, \alpha^{k+1} ||u_0|| + |a| \tau \frac{1-\alpha^k}{1-\alpha} |||f|||).$$

Vergelijking met (2.17) levert tenslotte bovenstaande bovengrenzen.

Met behulp hiervan bewijzen we de volgende stelling

Stelling 2.3

Schema (2.18) is R-F stabiel t.o.v. het paar (u_0, f) , wanneer $\alpha = 1 + O(\tau)$ voor $\tau \rightarrow 0$.

Bewijs

Indien $\alpha(\tau) \leq 1$ op het interval $0 < \tau \leq \bar{\tau}$, dan volgt uit

$$\lim_{\alpha \rightarrow 1} \frac{1 - \alpha^k}{1 - \alpha} = k,$$

dat

$$(2.22a) \quad ||R^{-1}|| = O(T).$$

Indien $\alpha > 1$, is $||R^{-1}||$ uniform begrensd voor $\tau \rightarrow 0$ als $\alpha^N = \alpha^{T/\tau}$ uniform begrensd is. Nu geldt

$$\begin{aligned} \alpha^N &= \exp [N \ln \alpha] = \exp [N \ln(1 + (\alpha - 1))] \\ &= \exp [N(\alpha - 1)(1 + O(\alpha - 1))] = \exp \left[T \cdot \frac{\alpha - 1}{\tau} \cdot (1 + O(\alpha - 1)) \right]. \end{aligned}$$

Wanneer dus $\alpha(\tau) = 1 + O(\tau)$, vinden we

$$(2.22b) \quad ||R^{-1}|| = \exp |O(T)|,$$

en dit is uniform begrensd, want T is constant verondersteld.

Alvorens deze stelling toe te passen geven we een criterium, waarvan in het volgende veel gebruik gemaakt zal worden in verband met het vinden van schattingen voor eigenwaarden.

Criterium 2.1

De wortels van de vergelijking $x^2 - Sx + P = 0$ liggen op of binnen de eenheidscirkel, wanneer S en P voldoen aan de ongelijkheden

$$P \leq 1, \quad 1 - S + P \geq 0, \quad 1 + S + P \geq 0.$$

Bewijs

Zijn de wortels complex dan worden deze gegeven door $x_1 = re^{i\phi}$ en $x_2 = re^{-i\phi}$. Nu stelt P het product der wortels voor dus $P = r^2$, waaruit de bewering volgt.

Zijn de wortels reëel, dan volgt uit figuur 2.2 dat aan de 2 laatste voorwaarden van het criterium voldaan moet worden.

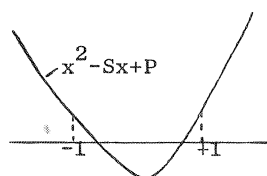


fig. 2.2

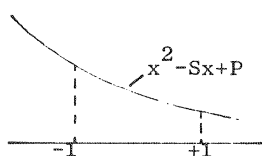


fig. 2.3

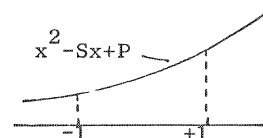


fig. 2.4

Omgekeerd volgen uit deze 2 voorwaarden één van de drie figuren 2.2, 2.3 en 2.4. Alleen de situatie in figuur 2.2 is consistent met $P \leq 1$, waarmee het criterium bewezen is.

We geven nu enkele toepassingen van stelling 2.3.

Voorbeeld 2.2

We beschouwen schema (1.9). Hiervoor zijn de amplificatie-matrices allen gelijk aan de factor $1-\tau$, dus $\alpha = 1-\tau$ m.a.w. schema (1.9) is R-F stabiel.

Voorbeeld 2.3

Schema (2.19) heeft de amplificatie-matrix

$$A_k = A = \begin{pmatrix} 2 \frac{\rho-\tau}{\rho+1} & b \frac{1-\rho}{1+\rho} \\ \frac{1}{b} & 0 \end{pmatrix}$$

Voor de berekening van $\alpha(\tau)$ moeten we een norm voor de matrix A kiezen. We zullen hier alleen beschouwen de z.g. maximum-norm $\|A\|_1$ en spectrale norm $\|A\|_2$ (zie hfdst. 1 deel 2). Deze worden respectievelijk gedefiniëerd

als

$$(2.23) \quad \|A\|_1 = \max_j \sum_i |a_{ij}|$$

$$(2.24) \quad \|A\|_2 = \sqrt{\sigma(A^* A)}.$$

Hierin stelt a_{ij} het matrixelement op de i^{de} rij en de j^{de} kolom voor, A^* de getransponeerde van A en $\sigma(A^* A)$ de grootste absolute waarde van de eigenwaarden van $A^* A$.

In het geval van de maximum-norm is eenvoudig na te gaan, dat $\alpha > 1$ voor $\tau \rightarrow 0$ als $\rho < 0$ of $\rho > 1$. Voor $0 \leq \rho \leq 1$ geldt, als we $b = 1$ kiezen

$$(2.25) \quad \begin{cases} \alpha(\tau) = \text{Max} \left(1, \frac{1-3\rho}{1+\rho} + \frac{2}{1+\rho} \tau \right) & \text{als } 0 \leq \rho \leq \tau \\ \alpha(\tau) = \text{Max} \left(1, 1 - \frac{2}{1+\rho} \tau \right) & \text{als } \tau \leq \rho \leq 1 \end{cases}$$

In het geval van de spectrale norm kiezen we $b = \sqrt{\frac{1+\rho}{1-\rho}}$, $\rho \neq 1$ en $\rho \neq -1$. De matrix A is nu symmetrisch en dus $\|A\|_2 = \sigma(A)$. De eigenwaardenvergelijking bij A is

$$\lambda^2 - 2 \frac{\rho-\tau}{\rho+1} \lambda - \frac{1-\rho}{1+\rho} = 0.$$

We gebruiken nu het criterium 2.1 met

$$(2.26) \quad S = 2 \frac{\rho-\tau}{\rho+1} \quad \text{en} \quad P = \frac{\rho-1}{\rho+1}.$$

Wanneer aan de voorwaarden van het criterium voldaan is voor $\tau = 0$, geldt $\alpha(\tau) \leq 1 + O(\tau)$. We vinden

$$(2.27) \quad \alpha(\tau) \leq 1 + O(\tau) \quad \text{als} \quad \rho \geq 0.$$

We zien in de eerste plaats dat stabiliteit afhankelijk is van de normen, die gebruikt worden: t.o.v. zowel de maximum-norm als de spectrale norm zijn de schema's met $0 \leq \rho \leq 1$ R-F stabiel, terwijl de schema's met $\rho > 1$ alleen t.o.v. de spectraal-norm R-F stabiel zijn.

In de tweede plaats volgt uit deze kwalitatieve analyse direct dat de in de Inleiding behandelde schema's (1.19) en (1.38) welke ontstaan als men $\rho = 0$ en $\rho = -3$ kiest, respectievelijk stabiel en instabiel zijn.

Stabiliteit van berekeningen over lange tijdsintervallen

In de Inleiding is reeds ter sprake gekomen, dat de stabiliteit over eindige en oneindige tijdsintervallen verschillend is. De kwantitatieve behandeling van schema (1.19) toonde aan dat zowel de discretiseringsfout (t.g.v. de benaderingsfout) als de numerieke fout (t.g.v. afrondingsfouten) exponentiëel met T samenhangen. Inderdaad volgt uit (2.22), dat een dergelijk exponentiëel gedrag niet uitgesloten wordt door de R-F stabiliteitsvoorwaarden. Wanneer er dus berekeningen gemaakt moeten worden over zeer lange tijdsintervallen (T groot) dan zou men de tijdstap heel klein moeten kiezen. We zullen nu een stabiliteitsdefinitie geven, welke zo'n exponentiëel gedrag bij voorbaat uitsluit.

Definitie 2.6

Het differentieschema $Ru = (u_0, f)$ is over het interval $[0, \infty]$ stabiel in de zin van O'Brien, Hyman en Kaplan [1], wanneer u continu is in $(u_0, 0)$ voor vaste waarden van de tijdstappen τ_k en wel uniform op het interval $0 \leq T \leq \infty$.

We zullen een dergelijke vorm van stabiliteit in het vervolg O'B-H-K stabiliteit noemen. Uit de equivalentie van continue en begrensde lineaire operatoren volgt dat de O'B-H-K stabiliteit betekent dat $\|R^{-1}\|$ uniform begrensd is voor $T \rightarrow \infty$ (i.p.v. $\tau_k \rightarrow 0$). Uit (2.20) en stelling 2.2 volgt dan voor schema (2.18)

Stelling 2.4

Schema (2.18) is O'B-H-K stabiel t.o.v. het paar $(u_0, 0)$, wanneer $\alpha \leq 1$ voor $T \rightarrow \infty$.

We merken op dat er voor $\alpha \leq 1$ ($T \rightarrow \infty$) geen O'B-H-K stabiliteit t.o.v. de inhomogene term mogelijk is, omdat volgens (2.22a) voor R^{-1} geldt

$$||R^{-1}|| = O(T),$$

maar i.p.v. het exponentiële gedrag, dat optreedt voor $\alpha > 1$, hebben we nu te maken met een lineaire aangroeiing.

Voorbeeld 2.4

We beschouwen nogmaals schema (2.19) en in het bijzonder de stabiliteitsvoorwaarden (2.25). We zien hieruit dat $\alpha \leq 1$ wanneer ρ en τ voldoen aan

$$(2.28a) \quad 2\rho - \tau \geq 0, \quad \rho \leq 1, \quad \tau \geq 0.$$

T.o.v. de spectraal-norm vinden we (toepassing van criterium 2.1 met $\tau \neq 0$) de voorwaarden

$$(2.28b) \quad 2\rho - \tau \geq 0, \quad \tau \geq 0.$$

Schema (1.19), dat ontstaat voor $\rho = 0$, is dus niet stabiel voor berekeningen over oneindige tijdsintervallen. Dit volgde reeds uit de kwantitatieve analyse uit hoofdstuk 1 (zie formule (1.37)). Ook het schema met (zie voorbeeld 2.1)

$$\rho = \frac{\sinh \tau - \tau}{\cosh \tau - 1} \sim \frac{1}{3}\tau$$

is niet O'B-H-K stabiel.

Elk schema echter met $\rho > 0$ voor $\tau = 0$ is stabiel te maken voor berekeningen over oneindige tijdsintervallen, door $\tau \leq 2\rho$ te kiezen.

Voorbeeld 2.5

Beschouw de differentiaalvergelijking

$$(2.29) \quad \begin{cases} \tilde{U}_{tt} - \tilde{U} = 0 & 0 < t \leq T \\ \tilde{U} = U_0, \quad \tilde{U}_t = V_0 & t = 0. \end{cases}$$

We schrijven dit als

$$(2.30) \quad \vec{w}_t = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \vec{w}, \quad \vec{w} = \vec{w}_0 \quad \text{als } t = 0.$$

Een consistent schema hiervoor is

$$(2.31) \quad \vec{w}_{k+1} = (1 + \tau \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}) \vec{w}_k.$$

De amplificatie-matrices hebben de vorm

$$A_k = A = \begin{pmatrix} 1 & \tau \\ \tau & 1 \end{pmatrix}$$

met eigenwaarden $\lambda = 1 \pm \tau$.

Het is duidelijk dat

$$\alpha = \|A\|_1 = \|A\|_2 = 1 + \tau,$$

zodat schema (2.30) R-F stabiel is, maar niet O'B-H-K stabiel.

Opmerking 2.1

De stabiliteitsstellingen 2.3 en 2.4 stellen in sommige gevallen een te zware voorwaarde. In het tweede deel van deze syllabus zullen problemen behandeld worden, waarvoor verscherpte stabiliteitsstellingen nodig zullen zijn.

Opmerking 2.2

In de literatuur komt men vaak stabiliteit in de zin van Lax en Richtmyer [7] tegen. Deze vorm van stabiliteit is equivalent met de R-F stabiliteit t.o.v. de beginvoorwaarden en wordt gedefiniëerd voor schema's van de vorm $u_{k+1} = Au_k$. Men eist dan dat de operatoren $(A)^{T/\tau}$ voor $\tau \rightarrow 0$ en T constant, uniform begrensd blijven.

Stabiliteit in de zin van Forsythe en Wasow

In de praktijk zal de differentieoplossing u nooit exact berekend kunnen worden, omdat er afrondingsfouten zullen optreden. Dit geeft aanleiding tot een numerieke oplossing u^* . In analogie met de benaderingsfout en de discretiseringsfout definiëren we:

Definitie 2.7

De norm van de roosterfunctie d , welke gevormd wordt door alle afrondingsfouten noemen we de totale afrondingsfout.

Definitie 2.8

De waarde van $|||u - u^*|||$ noemen we de rekenfout van het differentieschema t.g.v. de afrondingsfouten.

Forsythe en Wasow geven m.b.v. deze begrippen een meer op de praktijk afgestemde stabiliteitsdefinitie, welke op het volgende neerkomt:

Definitie 2.9

Het differentieschema $Ru = (u_0, f)$ is over het interval $[0, T]$ stabiel in de zin van Forsythe en Wasow [2], wanneer u^* continu is in d en wanneer de rekenfout niet sneller aangroeit dan een of andere macht van τ^{-1} voor $\tau \rightarrow 0$.

Forsythe en Wasow [2] betogen, dat in alle problemen, die mathematisch geanalyseerd zijn, de grootteorde van de rekenfout of een lage macht van τ^{-1} of een exponentiële functie van τ^{-1} is. In het laatste geval zijn de differentieschema's ongeschikt voor werkelijke berekening; dergelijke schema's worden door Forsythe en Wasow instabiel genoemd.

In analogie met de stellingen 2.3 en 2.4 voor de R-F en de O'B-H-K stabiliteit bewijzen we voor de F-W stabiliteit (stabiliteit in de zin van Forsythe en Wasow)

Stelling 2.5

Schema (2.18) is F-W stabiel, wanneer $\alpha \leq 1 + O(\tau \ln \tau)$.

Bewijs

De numerieke oplossing u^* van schema (2.18) kunnen we beschrijven met het schema

$$(2.32) \quad \begin{aligned} u_{k+1}^* &= A_k u_k^* + a\tau f_k + d_k \\ &= A_k u_k^* + a\tau(f_k + (a\tau)^{-1}d_k). \end{aligned}$$

Hierin is d_k de afrondingsfout, die op het tijdstip $t_k = k\tau$ gemaakt wordt. We kunnen deze afrondingsfouten interpreteren als een verstoring $(a\tau)^{-1}d$ van de inhomogene term, dus

$$(2.32') \quad u^* = R^{-1}(u_0, f + (a\tau)^{-1}d).$$

Voor de rekenfout vinden we

$$u - u^* = R^{-1}(0, (a\tau)^{-1}d),$$

zodat analoog aan (2.11) geldt

$$(2.33) \quad |||u - u^*||| \leq |||R^{-1}||| \cdot |a|^{-1} \cdot \tau^{-1} \cdot |||d|||.$$

Volgens stelling 2.2 wordt R^{-1} in dit geval begrensd door

$$\max_{0 \leq k \leq T/\tau} |a|_{\tau} \frac{1 - \alpha^k}{1 - \alpha}.$$

Substitutie van $\alpha = 1 + O(\tau \ln \tau)$ geeft analoog aan het bewijs van stelling 2.3, dat

$$|||R^{-1}||| \leq |a|_{\tau} \cdot \frac{T}{\tau} \quad \text{als} \quad \alpha \uparrow 1 \quad \text{voor} \quad \tau \rightarrow 0$$

$$|||R^{-1}||| = O(\tau^{-q}) \quad \text{als} \quad \alpha \uparrow 1 \quad \text{voor} \quad \tau \rightarrow 0,$$

waarin $q = O(T)$.

Hieruit en uit (2.33) volgt de stelling.

Opmerking 2.3

Stel dat

$$||R^{-1}|| = O(\tau^{-q}), \quad \tau \rightarrow 0$$

en dat

$$||\tilde{f} - f|| = O(\tau^p), \quad \tau \rightarrow 0,$$

dan geldt volgens (2.11) voor de discretiseringsfout

$$||\tilde{u} - u|| = O(\tau^{p-q}).$$

Dus wanneer $q < p$ sluit F-W stabiliteit convergentie van de differentie-oplossingen niet uit (vergelijk de toelichting bij stelling 2.1).

Met behulp van (2.33) kunnen we een bovengrens geven voor, waar het ons in de praktijk uiteindelijk om gaat, het verschil $\tilde{u} - u^*$. Hiervoor definiëren we

Definitie 2.10

De waarde van $||\tilde{u} - u^*||$ noemen we de totale fout van het differentieschema.

Uit (2.11) en (2.33) volgt met behulp van de driehoeksongelijkheid

$$(2.34) \quad ||\tilde{u} - u^*|| = ||(\tilde{u} - u) + (u - u^*)|| \leq ||\tilde{u} - u|| + ||u - u^*||$$

dus

$$(2.35) \quad ||\tilde{u} - u^*|| \leq ||R^{-1}|| \cdot (||\tilde{f} - f|| + |a|^{-1} \tau^{-1} ||d||),$$

derhalve

$$(2.35') \quad ||\tilde{u} - u^*|| = ||R^{-1}|| \cdot (O(\tau^p) + ||d|| \cdot O(\tau^{-1})),$$

waarin p de orde van benadering is.

Het verkleinen van τ reduceert de benaderingsfout maar vergroot de rekenfout. In de praktijk is $||d||$ echter zo klein, dat men in het algemeen τ geen beperkingen hoeft op te leggen, mits het schema F-W stabiel is.

Men kan zich echter voorstellen, dat voor een exponentieel gedrag in τ^{-1} van $||R^{-1}||$ (instabiliteit), òf de benadering zó grof is òf $||R^{-1}||$ zó groot is, dat de oplossing u^* onbruikbaar is.

Voorbeeld 2.6

In de Inleiding is voor schema (1.19) een expliciete uitdrukking afgeleid voor de rekenfout als functie van $|||d|||$, τ en T (vergelijk (1.37)). We zien hieruit dat de rekenfout $O(\tau^{-1})$ is.

We besluiten dit hoofdstuk met een overzicht van de gegeven stabiliteitsdefinities in de vorm van voorwaarden voor $\alpha(\tau)$.

Schema (2.18) is stabiel in de zin van

- | | | |
|---|--|--|
| 1 ^o . Rjabenki en Filippow | als $\alpha \leq 1 + O(\tau)$ | voor $\tau \rightarrow 0$ en T constant |
| 2 ^o . O'Brien, Hyman en Kaplan | als $\alpha \leq 1$ | voor $T \rightarrow \infty$ en τ constant |
| 3 ^o . Forsythe en Wasow | als $\alpha \leq 1 + O(\tau \ln \tau)$ | voor $\tau \rightarrow 0$ en T constant |
| 4 ^o . Godunov en Rjabenki | als het R-F stabiel is | |
| 5 ^o . Lax en Richtmyer | als het R-F stabiel is t.o.v. beginvoorwaarden | |

3. Differentieschema's voor partiële differentiaalvergelijkingen

De voor gewone differentiaalvergelijkingen gegeven beschouwingen worden in dit hoofdstuk uitgebreid tot partiële differentiaalvergelijkingen. Daar waar de theorie vrijwel analoog is zal volstaan worden met het geven van de overeenkomstige definities en stellingen.

Definities

Beschouw een reëel interval $[0, T]$, een Euclidische ruimte $\mathcal{R}_m$ van de dimensie m , en een gebied $G \subset \mathcal{R}_m$ met rand Γ . Analoog aan hoofdstuk 2 voeren we in, een lineaire ruimte E van op $(G \cup \Gamma) \times [0, T]$ gedefiniëerde vectorfuncties U , en een lineaire ruimte E_0 van op $G \cup \Gamma$ gedefiniëerde vectorfuncties U_0 . Voorts definiëren we een lineaire deelruimte E_Γ van E , welke alle elementen $U \in E$ bevat, die nul zijn op $G \times [0, T]$. We noteren de elementen uit E_Γ met Griekse hoofdletters ($\neq \Gamma$). Interpreteren we de variabele $t \in [0, T]$ als de tijd en de punten $x \in G \cup \Gamma$ als plaatsvector, dan zou men de elementen van E_0 als niveaufuncties behorend bij E kunnen opvatten en E_Γ als de verzameling randfuncties uit E .

Met behulp van deze ingevoerde ruimten willen we de begin-randwaardeproblemen voor lineaire partiële differentiaalvergelijkingen abstract definiëren.

Definitie 3.1

Een begin-randwaardeprobleem voor een stelsel lineaire partiële differentiaalvergelijkingen is een lineaire afbeelding L van een (onbekend) element $U \in E$ op een (bekend) element $(U_0, F, \phi) \in E_0 \times E \times E_\Gamma$.

We noemen U_0 de beginfunctie, F de inhomogene term en ϕ de randfunctie. Indien Γ leeg is en G uit één punt bestaat, definiëert definitie 3.1 een beginwaardeprobleem voor een gewone differentiaalvergelijking. We noteren in het vervolg een partiële differentiaalvergelijking als

$$(3.1) \quad LU = (U_0, F, \phi).$$

Voorbeeld 3.1

Het volgende begin-randwaardeprobleem voor de diffusie-vergelijking is op te vatten als een lineaire toevoeging van de oplossing U aan de drie functies U_0 , F en ϕ

$$U_t - U_{xx} = F \quad \text{voor } (x,t) \in (G \cup \Gamma) \times [0,T]$$

$$U_x = \phi \quad \text{voor } (x,t) \in \Gamma \times [0,T]$$

$$U = U_0 \quad \text{voor } (x,t) \in (G \cup \Gamma) \times [0,0].$$

Hierin zijn U , F en ϕ scalarfuncties van de reële variabelen x en t en is U_0 een functie van x . G bestaat uit een open interval van de x -as en Γ bestaat uit de 2 randpunten van G .

We normeren de ruimten E en E_0 met normen $|||U|||$ en $||U_0||$ en met behulp hiervan maken we ook $E_0 \times E \times E_\Gamma$ tot een genormeerde ruimte door te definiëren:

$$(3.2) \quad ||(U_0, F, \phi)|| = ||U_0|| + |||F||| + |||\phi|||.$$

Aan L leggen we twee eisen op:

- 1°. L heeft een éénduidige inverse L^{-1} .
- 2°. L^{-1} is continu in (U_0, F, ϕ) .

In het algemeen kennen we L^{-1} niet, maar zijn éénduidige existentie is in vele gevallen wel te bewijzen. De continuïteit van L^{-1} is op te vatten, zoals in het voorgaande hoofdstuk al opgemerkt is, als een stabiliteits-eigenschap van de oplossing U t.o.v. de gegevens U_0 , F en ϕ .

In het vervolg geven we het definitiegebied van L met D_L aan, het beeldgebied met Δ_L en we veronderstellen

$$\Delta_L = L(D_L) = E_0 \times E \times E_\Gamma.$$

Op analoge wijze willen we nu abstract differentieschema's definiëren. In plaats van het interval $[0,T]$ beschouwen we de verzameling punten uit $[0,T]$, welke gegeven worden door

$$\{t_k \mid t_0 = 0, t_N = T, k = 0, 1, \dots, N\}.$$

We noteren $t_{k+1} - t_k$ met τ_k en de grootste τ_k met τ . We kiezen bij elke rij $\{\tau_k\}$ in G een verzameling punten G_τ en op Γ een verzameling punten Γ_τ .

Definitie 3.2

De puntverzamelingen $Q_\tau = \{t_k\}_0^N \ast (G_\tau \cup \Gamma_\tau)$, met $0 < \tau \leq \bar{\tau}$, noemen we een verzameling roosters, wanneer ze de eigenschap hebben, dat in elke omgeving van een punt uit $G \cup \Gamma$ voor voldoende kleine τ een punt van $G_\tau \cup \Gamma_\tau$ ligt.

Verder definiëren we een lineaire ruimte $E^\ast$ van op Q_τ gedefiniëerde vectorfuncties u , een lineaire ruimte $E_0^\ast$ van op $G_\tau \cup \Gamma_\tau$ gedefiniëerde vectorfuncties u_0 en een lineaire deelruimte $E_\Gamma^\ast$ van $E^\ast$ van vectorfuncties met de eigenschap dat ϕ nul is op $\{t_k\}_0^N \ast G_\tau$. Wanneer $U \in E$, $U_0 \in E_0$ of $\phi \in E_\Gamma$, dan noteren we

$$[U]_d = u, \quad [U_0]_d = u_0, \quad [\phi]_d = \phi.$$

Hierin is $[]_d$ een discretiseringsoperator, welke aan functies gedefiniëerd op "niet-discrete" verzamelingen, de roosterwaarden toekent.

In analogie met definitie 3.1 volgt nu

Definitie 3.3

Een differentieschema voor een begin-randwaardeprobleem is een lineaire afbeelding R van een (onbekend) element $u \in E^\ast$ op een (bekend) element $(u_0, f, \phi) \in E_0^\ast \ast E^\ast \ast E_\Gamma^\ast$.

We veronderstellen dat R^{-1} bestaat, éénduidig is en expliciet bekend is.

De oplossing van het begin-randwaardeprobleem $LU = (U_0, F, \phi)$ wordt in het vervolg de exacte of analytische oplossing genoemd en met $\tilde{U}$ aangegeven, de oplossing u van $Ru = (u_0, f, \phi)$ wordt differentieoplossing genoemd en de tengevolge van het afrondingsfoutenveld d berekende oplossing $u^\ast$ zal numerieke oplossing genoemd worden.

De ruimten $E^\ast$, $E_0^\ast$ en $E_\Gamma^\ast \ast E^\ast \ast E_\Gamma^\ast$ normeren we met normen

$$(3.3) \quad |||u|||, ||u_0|| \text{ en } ||(u_0, f, \phi)|| = ||u_0|| + |||f||| + |||\phi|||$$

en we eisen dat

$$(3.4) \quad |||u||| \rightarrow |||U|||, ||u_0|| \rightarrow ||U_0|| \text{ als } \tau \rightarrow 0.$$

We kunnen nu alle nodige begrippen definiëren. De definities lopen parallel met de in hoofdstuk 2 gegeven definities.

Definitie 3.4

De waarde van $||Ru - [L\tilde{U}]_d||$ is de benaderingsfout van het schema $Ru = (u_0, f, \phi)$ voor het begin-randwaardeprobleem $LU = (U_0, F, \Phi)$.

Definitie 3.5

De waarde $|||\tilde{u} - u|||$ is de discretiseringsfout van de differentie-oplossing.

Definitie 3.6

Een differentieschema is consistent met een differentiaalvergelijking over een klasse $C \subset D_L \cap E$, wanneer voor alle $\tilde{U} \in C$ de benaderingsfout van het differentieschema voor $\tau \rightarrow 0$ naar nul gaat.

Definitie 3.7

De differentieoplossingen convergeren over de klasse $C \subset D_L \cap E$ naar de analytische oplossingen $\tilde{U} \in C$, wanneer voor alle $\tilde{U} \in C$ de discretiseringsfout van de differentieoplossingen voor $\tau \rightarrow 0$ naar nul gaat.

Definitie 3.8

Het differentieschema $Ru = (u_0, f, \phi)$ is over het interval $[0, T]$ stabiel in de zin van Rjabenki en Filippow, wanneer u continu is in $(u_0, f, \phi) \in E_0^* * E^* * E_\Gamma^*$ en wel uniform in $0 < \tau_k \leq \bar{\tau}$.

Definitie 3.9

Het differentieschema $Ru = (u_0, f, \phi)$ is over het interval $[0, T]$ stabiel in de zin van Godunov en Rjabenki, wanneer geldt

$$(3.5) \quad |||u||| \leq N_1 |||u_0||| + N_2 |||f||| + N_3 |||\phi|||,$$

waarin N_1, N_2, N_3 onafhankelijk van τ_k en (u_0, f, ϕ) zijn.

Definitie 3.10

Het differentieschema $Ru = (u_0, f, \phi)$ is over het interval $[0, \infty]$ stabiel in de zin van O'Brien, Hyman en Kaplan, wanneer u continu is in $(u_0, 0, 0)$ voor vaste waarden van de tijdstappen τ_k en wel uniform op het interval $0 \leq T \leq \infty$.

Definitie 3.11

Het differentieschema $Ru = (u_0, f, \phi)$ is over het interval $[0, T]$ stabiel in de zin van Forsythe en Wasow, wanneer de rekenfout $u - u^*$ voldoet aan het schema $R^*(u - u^*) = d$ met $|(R^*)^{-1}| = O(\tau^{-q})$, $q \geq 0$ voor $\tau \rightarrow 0$.

Stellingen

De in hoofdstuk 2 gevonden resultaten zijn zonder meer over te dragen op differentieschema's van het type $Ru = (u_0, f, \phi)$. Zonder bewijs vermelden we de volgende drie stellingen

Stelling 3.1

Voor de benaderingsfout en de discretiseringsfout geldt

$$(3.6) \quad |||\tilde{u} - u||| \leq |||R^{-1}||_C \cdot |||R\tilde{u} - [L\tilde{u}]_d|||,$$

waarin C de beschouwde klasse van functies $\tilde{U}$ is (vergelijk (2.10)).

Stelling 3.2

De stabiliteitsdefinities van Rjabenki-Filippow en Godunov-Rjabenki zijn equivalent.

Stelling 3.3

De differentieoplossingen van een consistent R-F stabiel schema convergeren naar de exacte oplossing als $\tau \rightarrow 0$.

Vervolgens bewijzen we de volgende zeer belangrijke stelling.

Stelling 3.4

Stel dat voldaan wordt aan de volgende voorwaarden

(1) Voor de benaderingsfout over de klasse C geldt

$$\tau^{-p}(\tilde{R}u - [\tilde{L}\tilde{U}]_d) = [(0, H, \Psi)]_d + o(1), \quad \tau \rightarrow 0$$

waarin $p > 0$ en $(0, H, \Psi) \in L(C)$. ($L(C)$ is het beeld van C onder L.)

(2) W is de oplossing van $LW = (0, H, \Psi)$.

(3) Het schema is R-F stabiel.

Onder deze voorwaarden geldt

$$(3.7) \quad \tilde{u} = u + \tau^p w + o(\tau^p), \quad \tau \rightarrow 0.$$

Bewijs

Uit stelling 3.1 volgt (zie ook formule (2.11))

$$\tilde{u} - u = \tau^p R^{-1} \tau^{-p}(\tilde{R}u - [\tilde{L}\tilde{U}]_d),$$

zodat volgens voorwaarde (1)

$$\tilde{u} - u = \tau^p R^{-1} ([(0, H, \Psi)]_d + o(1)).$$

Volgens voorwaarde (2)

$$\tilde{u} - u = \tau^p R^{-1} ([LW]_d + o(1)).$$

Uit voorwaarde (1) volgt dat de benaderingsfout van de orde τ^p is, dus

$$Rw - [LW]_d = O(\tau^p).$$

Hieruit volgt

$$\begin{aligned} \tilde{u} - u &= \tau^p R^{-1} (Rw + O(\tau^p) + o(1)) \\ &= \tau^p w + \tau^p R^{-1} (O(\tau^p) + o(1)) \\ &= \tau^p w + \tau^p R^{-1} (o(1)). \end{aligned}$$

Volgens voorwaarde (3) is R^{-1} continu, dus $R^{-1}(o(1)) = o(1)$ voor $\tau \rightarrow 0$. Hiermee is de stelling bewezen.

Opmerking 3.1

De stelling beschrijft een asymptotische voorstelling voor $\tilde{u} - u$. We zien dat de locale discretiseringsfout van dezelfde orde in τ is als de benaderingsfout.

Opmerking 3.2

Voorwaarde (1) is geenszins triviaal; de bewering is dat het linkerlid voor $\tau \rightarrow 0$ een discretisering is van een element uit $L(C)$ op de limiet van de verzameling roosters $\{Q_\tau\}_{\tau=\tau}^{\tau=0}$. Aangezien Q_0 een echte deelverzameling van $G \cup \Gamma$ kan zijn (Q_0 ligt dan dicht in $G \cup \Gamma$ volgens definitie (3.2)), is niet altijd aan voorwaarde (1) voldaan. Bijvoorbeeld, wanneer $L(C)$ uit continue functie-tripels bestaat, moet $\tau^{-p}(Ru - [L\tilde{u}]_d)$ voldoende "glad" zijn op Q_0 om een discretisering te kunnen zijn van een element uit $L(C)$.

Opmerking 3.3

We kunnen w elimineren wanneer we twee differentieoplossingen u_1 en u_2 kennen. Op de doorsnede van de roosters Q_1 en Q_2 geldt dan

$$(3.8) \quad \tilde{u} = u_2 + \frac{u_2 - u_1}{(\tau_1/\tau_2)^p - 1} + o(\tau_1^p \tau_2^p).$$

Door deze extrapolatie kan een aanzienlijke verhoging van de nauwkeurigheid verkregen worden.

Differentieschema's van de vorm $u_{k+1} = A_k u_k + \tau Df_k + B\phi_k$

We zullen het volgende schema nader beschouwen

$$(3.9) \quad \begin{cases} u_{k+1} = A_k u_k + \tau Df_k + B\phi_k, & k = 0, 1, \dots, N-1 \\ A_k, D \text{ en } B \text{ lineaire operatoren} \\ Df_k = 0 \text{ op } \Gamma_\tau \end{cases}$$

De meeste lineaire partiële differentiaalvergelijkingen geven aanleiding tot dergelijke schema's.

Voorbeeld 3.2

We beschouwen het volgende begin-randwaardeprobleem voor de diffusie-vergelijking

$$U_t - U_{xx} = F \quad \text{voor } 0 \leq x \leq 1, 0 \leq t \leq T$$

$$U_x = -\phi \quad \text{voor } x = 0, 0 \leq t \leq T$$

$$U_x = \phi \quad \text{voor } x = 1, 0 \leq t \leq T$$

$$U = U_0 \quad \text{voor } 0 \leq x \leq 1, t = 0.$$

We kiezen een rooster Q_τ bestaande uit punten $(x_j, t_k) = (j\xi, k\tau)$, waarin j de waarden $0, 1, 2, \dots, m$ aanneemt en $\xi = \frac{1}{m}$ is. We drukken ξ in τ uit door middel van de relatie

$$\xi = \sqrt{\frac{\tau}{r}}$$

waarin r constant verondersteld wordt. De puntverzamelingen Q_τ vormen nu inderdaad een rij roosters in de zin van definitie 3.2.

Op het rooster Q_τ definiëren we het schema

$$u_{k+1} = X_+ u_{k+1} + \sqrt{\frac{\tau}{r}} \cdot \phi_{k+1} \quad \text{in het punt } x_j = 0$$

$$u_{k+1} = r X_- u_k + (1 - 2r) u_k + r X_+ u_k + \tau f_k \quad \text{in de punten } x_j, j=1, 2, \dots, m-1$$

$$u_{k+1} = X_- u_{k+1} + \sqrt{\frac{\tau}{r}} \phi_{k+1} \quad \text{in het punt } x_j = x_m.$$

Hierin stelt $X_\pm$ een verschuivingsoperator voor welke aan de functiewaarde in een punt (x_j, t_k) de functiewaarde in het punt $(x_{j\pm 1}, t_k)$ toevoegt. Dit schema verkrijgen we door vervangen van de afgeleiden door differentiequotienten (vergelijk schema (1.48) uit de Inleiding).

We vormen bovenstaand schema om tot

$$(3.10) \begin{cases} u_{k+1} = ru_k + (1 - 2r)X_+ u_k + rX_+^2 u_k + \tau X_+ f_k + \sqrt{\frac{\tau}{r}} E_+ \phi_k, & j = 0 \\ u_{k+1} = rX_- u_k + (1 - 2r)u_k + rX_+ u_k + \tau f_k, & j = 1, \dots, m-1 \\ u_{k+1} = rX_-^2 u_k + (1 - 2r)X_- u_k + ru_k + \tau X_- f_k + \sqrt{\frac{\tau}{r}} E_+ \phi_k, & j = m, \end{cases}$$

waarin E_+ gedefiniëerd wordt door $E_+ \phi_k = \phi_{k+1}$.

Uit deze differentieformules kan men direct de operatoren A, D en B aflezen.

Voorbeeld 3.3

Vervolgens beschouwen we een zuiver beginwaarde probleem (Cauchy-probleem) voor de golf-vergelijking.

$$U_{tt} - U_{xx} = F \quad \text{voor} \quad -\infty \leq x \leq \infty, \quad 0 \leq t \leq T$$

$$U = U_0, \quad U_t = V_0 \quad \text{voor} \quad -\infty \leq x \leq \infty, \quad t = 0.$$

We voeren $(U, U_t) = (U, V)$ als nieuwe afhankelijke (vector-)variabele in en definiëren op het rooster $Q_\tau = \{(x_j, t_k) \mid x_j = j\xi = j \frac{\tau}{r}, \quad t_k = k\tau, \quad r \text{ constant}\}$ het differentieschema:

$$u_{k+1} = u_k + \tau v_k$$

$$v_{k+1} = v_k + r(X_- - 2 + X_+)u_k + \tau f_k$$

ofwel

$$(3.11) \quad \begin{pmatrix} u \\ v \end{pmatrix}_{k+1} = \begin{pmatrix} 1 & \tau \\ r(X_- - 2 + X_+) & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}_k + \tau \begin{pmatrix} 0 \\ f_k \end{pmatrix}$$

Hierin stelt $X_\pm$ weer dezelfde verschuivingsoperator voor als in voorbeeld 3.2.

Opmerking 3.4

Sommige schema's zijn van de vorm

$$C_k u_{k+1} = H_k u_k + \dots,$$

dus

$$A_k = C_k^{-1} H_k.$$

Wanneer men C_k^{-1} niet expliciet kent, wordt het schema impliciet genoemd. In alle andere gevallen spreekt men van expliciete schema's.

Stabiliteit van schema (3.9)

We brengen de norm van R^{-1} in verband met de normen van A_k , D en B . Door iteratie verkrijgen we uit (3.9)

$$\begin{aligned} u_{k+1} &= A_k A_{k-1} \dots A_0 u_0 \\ &+ \tau(Df_k + \sum_{n=1}^k A_k \dots A_{k-n+1} Df_{k-n}) \\ &+ B\phi_k + \sum_{n=1}^k A_k \dots A_{k-n+1} B\phi_{k-n}. \end{aligned}$$

We definiëren de operatoren P_k , Q_k en S_k door de formules

$$(3.12) \quad P_k u_0 = A_k \dots A_0 u_0$$

$$(3.13) \quad Q_k f = \tau(Df_k + \sum_{n=1}^k A_k \dots A_{k-n+1} Df_{k-n})$$

$$(3.14) \quad S_k = B\phi_k + \sum_{n=1}^k A_k \dots A_{k-n+1} B\phi_{k-n}.$$

Hiermee kan schema (3.9) in de volgende vorm geschreven worden

$$(3.15) \quad u_{k+1} = P_k u_0 + Q_k f + S_k \phi.$$

We kiezen weer (vergelijk (2.21))

$$|||u||| = \max_{0 \leq k \leq N} ||u_k||$$

en in analogie voor de operatornormen $||P_k||$, $||Q_k||$ en $||S_k||$ de bovengrenzen $|||P|||$, $|||Q|||$ en $|||S|||$ volgens de formule:

$$(3.16) \quad |||P||| = \max_{0 \leq k \leq N} |||P_k|||.$$

Aldus verkrijgen we de ongelijkheid

$$(3.17) \quad |||u||| \leq |||P||| \cdot |||u_0||| + |||Q||| \cdot |||f||| + |||S||| \cdot |||\phi|||.$$

Analoog aan stelling 2.2 geldt nu de stelling

Stelling 3.5

Voor schema (3.9) geldt

$$(a) \quad |||R^{-1}|||_C \leq \max_{0 \leq k \leq N} \alpha^k \quad \text{als } C = \{(u_0, 0, 0) \mid u_0 \in E_0^*\}$$

$$(b) \quad |||R^{-1}|||_C \leq \max_{0 \leq k \leq N} (\tau |||D||| \frac{1 - \alpha^k}{1 - \alpha}) \quad \text{als } C = \{(0, f, 0) \mid f \in E^*\}$$

$$(c) \quad |||R^{-1}|||_C \leq \max_{0 \leq k \leq N} (|||B||| \frac{1 - \alpha^k}{1 - \alpha}) \quad \text{als } C = \{(0, 0, \phi) \mid \phi \in E_\Gamma^*\}$$

waarin α weer de maximale amplificatie-factor voorstelt.

Stabiliteit t.o.v. de beginvoorwaarden

Uit stelling 3.5 (a) volgen direct de volgende twee stellingen (vergelijk de stellingen 2.3 en 2.4)

Stelling 3.6

Schema (3.9) is R-F stabiel t.o.v. de beginvoorwaarden, wanneer $\alpha = 1 + O(\tau)$ voor $\tau \rightarrow 0$.

Stelling 3.7

Schema (3.9) is O'B-H-K stabiel, wanneer $\alpha \leq 1$ voor $T \rightarrow \infty$.

Voorbeeld 3.4

We voegen in voorbeeld 3.2 nog een term aU aan het linkerlid toe. In de punten x_j , $j = 0, 1, \dots, m$ geldt dan

$$(3.18) \left\{ \begin{array}{l} u_{k+1} = ru_k + (1 - 2r - a\tau)X_+ u_k + rX_+^2 u_k, \quad j = 0 \\ u_{k+1} = rX_- u_k + (1 - 2r - a\tau)u_k + rX_+ u_k, \quad j = 1, \dots, m-1 \\ u_{k+1} = rX_-^2 u_k + (1 - 2r - a\tau)X_- u_k + ru_k, \quad j = m, \end{array} \right.$$

waarin de inhomogene term en de randfunctie buiten beschouwing gelaten zijn.

Wordt voor $||u_k||$ en $||A_k||$ de maximumnorm gekozen,

$$||u_k|| = \max_{0 \leq j \leq m} |u_k(x_j)|, \quad ||A_k|| = ||A_k||_1,$$

waarin $||\cdot||_1$ door (2.23) gedefiniëerd wordt, dan geldt

$$\alpha(\tau) = r + |1 - 2r - a\tau| + r.$$

We passen stelling 3.6 toe. Dit levert

$$|1 - 2r - a\tau| \leq 1 - 2r + c\tau,$$

waarin c een constante is. Uitwerking geeft

$$(3.19) \quad a \geq -c, \quad r \leq \frac{1}{2} + \frac{1}{4}(c - a)\tau.$$

Dit moet gelden voor $\tau \rightarrow 0$, dus we vinden de R-F stabiliteitsconditie

$$(3.20) \quad r \leq \frac{1}{2}.$$

Vervolgens passen we stelling 3.7 toe, dus $c = 0$ en $\tau = \text{constant}$.

Uit (3.19) volgen de O'B-H-K stabiliteitscondities

$$(3.21) \quad a \geq 0, \quad \tau \leq 2 \frac{1 - 2r}{a}.$$

Voorbeeld 3.5 .

We beschouwen voorbeeld 3.4, maar kiezen een andere randvoorwaarde in $x = 0$, namelijk

$$U_x + bU = 0.$$

Een consistente differentieformule hiervoor is

$$u_{k+1} = \frac{1}{1 - b\sqrt{\frac{\tau}{r}}} X_+ u_{k+1}, \quad x_j = 0.$$

Dit kan m.b.v. de in x_1 geldende differentieformule geschreven worden als

$$u_{k+1} = \frac{\sqrt{r}}{\sqrt{r - b\sqrt{\tau}}} \cdot (ru_k + (1 - 2r - a\tau)X_+ u_k + rX_+^2 u_k).$$

Behalve de voorwaarden (3.20) of (3.21) moet nu ook gelden

$$(3.22) \quad |1 - 2r - a\tau| \leq \frac{|\sqrt{r - b\sqrt{\tau}}|}{\sqrt{r}} (1 + c\tau) - 2r.$$

We beschouwen de R-F stabiliteit van het schema als $r = \frac{1}{2}$. Voorwaarde (3.22) luidt dan

$$|a|\tau \leq |1 - b\sqrt{2\tau}| (1 + c\tau) - 1$$

ofwel voor $\tau \rightarrow 0$

$$1 - b\sqrt{2\tau} \geq 1 + O(\tau)$$

waaruit volgt

$$(3.23) \quad b \leq 0.$$

Vervolgens beschouwen we de O'B-H-K stabiliteit. Uit (3.21) volgt dat $1 - 2r - a\tau$ niet negatief is, zodat (3.22) overgaat in ($c = 0$)

$$\sqrt{r} (1 - a\tau) \leq |\sqrt{r - b\sqrt{\tau}}|.$$

We werken dit uit voor $b \leq 0$ en vinden dan voor $a \neq 0$

$$(3.24) \quad \tau \geq \frac{b^2}{a^2 r}.$$

Indien $a = 0$ wordt er geen verdere voorwaarde gevonden.

Stabiliteit t.o.v. de inhomogene term

Uit stelling 3.5 (a) en (b) volgt direct de volgende stelling.

Stelling 3.8

Schema (3.9) is stabiel t.o.v. de inhomogene term, wanneer het stabiel is t.o.v. de beginvoorwaarden en wanneer $\|D\|$ uniform begrensd is voor $\tau \rightarrow 0$.

Voorbeeld 3.6

Aan Rjabenki en Filippow [8, p. 38] ontleen we het volgende beginrandwaardeprobleem

$$\frac{\partial^3 U}{\partial t \partial x^2} + \frac{\partial U}{\partial t} = F, \quad 0 \leq x \leq \pi, \quad 0 \leq t \leq \pi$$

$$U = 0, \quad x = 0, \pi, \quad 0 \leq t \leq \pi$$

$$U = U_0(x), \quad 0 \leq x \leq \pi, \quad t = 0.$$

Op het rooster Q_τ , bestaande uit punten $(x_j, t_k) = (j\xi, k\tau)$ met $\xi = \frac{\pi}{m}$, definiëren we het differentieschema:

$$u_{k+1} = 0 \quad \text{in het punt } x_j = 0,$$

$$X_+ u_{k+1} + (\tau^2 - 2)u_{k+1} + X_- u_{k+1} = X_+ u_k + (\tau^2 - 2)u_k + X_- u_k + \tau^3 f_k$$

$$\text{in de punten } x_j = j\xi, \quad j = 1, \dots, m-1,$$

$$u_{k+1} = 0 \quad \text{in het punt } x_j = \pi.$$

We hebben met een impliciet schema te maken

$$Cu_{k+1} = Hu_k + \tau^3 f_k.$$

Op operator D is van de vorm $\tau^2 C^{-1}$. We zullen hiervan bewijzen dat deze niet uniform begrensd blijft als τ naar nul gaat.

Beschouw de roosterfunctie

$$g_0 = (0, \sin \tau, \sin 2\tau, \dots, \sin m\tau),$$

waarin de tussen haakjes voorkomende waarden de waarden van g_0 in $x_j = j\xi$ voorstellen. Uit bovenstaand schema zien we direct dat

$$Cg_0 = (0, (\tau^2 - 2)\sin \tau + \sin 2\tau, \dots, \sin(j-1)\tau + (\tau^2 - 2)\sin j\tau + \\ + \sin(j+1)\tau, \dots, \sin m\tau).$$

Uitwerking geeft

$$Cg_0 = (\tau^2 - 2 + 2 \cos \tau)g_0.$$

Blijkbaar is g_0 een eigenfunctie van C met eigenwaarde $\lambda = (\tau^2 - 2 + 2 \cos \tau)$. Hieruit volgt dat λ^{-1} eigenwaarde van C^{-1} is en aangezien $\|C^{-1}\|$ groter dan of gelijk aan zijn eigenwaarden is t.o.v. geschikt gekozen normen, geldt

$$\|D\| = \tau^2 \|C^{-1}\| \geq \tau^2 (\tau^2 - 2 + 2 \cos \tau)^{-1} = O(\tau^{-2}).$$

D is dus niet uniform begrensd voor $\tau \rightarrow 0$ en daarmee is het schema instabiel t.o.v. de inhomogene term. Het schema is wel stabiel t.o.v. de beginvoorwaarden (zie [8, p. 38]).

Stabiliteit t.o.v. de randvoorwaarden

Aan deze vorm van stabiliteit is over het algemeen heel moeilijk te voldoen. In de operator S_k ontbreekt namelijk de factor τ zoals die in de operator Q_k optreedt.

Stelling 3.9

Schema (3.9) is stabiel t.o.v. de randvoorwaarden in de volgende twee gevallen:

- (a) De randvoorwaarden zijn van de eerste soort en $\|B\|$ is uniform begrensd voor $\tau \rightarrow 0$.
- (b) $\alpha < 1$ en $\|B\|$ is uniform begrensd voor $\tau \rightarrow 0$.

Bewijs

(a) Wanneer de randvoorwaarden van de eerste soort zijn hebben de operatoren A_k in de randpunten een "nulwerking", zodat uit de definitie van S_k (formule (3.14)) volgt

$$S_k \phi = B \phi_k.$$

Uit (3.17) en de uniforme begrensdsheid van B voor $\tau \rightarrow 0$ volgt het eerste deel van de stelling.

(b) Dit volgt direct uit stelling 3.5 (c).

Voorbeeld 3.7

We beschouwen nogmaals het begin-randwaardeprobleem voor de diffusie-vergelijking uit voorbeeld 3.2. Uit schema (3.10) volgt, dat voor de randoperatoren geldt

$$||B|| = O(\sqrt{\tau}),$$

zodat de stabiliteit t.o.v. de randvoorwaarden niet vast staat.

4. Zwakke en sterke stabiliteit

In dit hoofdstuk zullen we de invloed van afrondingsfouten op het rekenproces onderzoeken, en wel aan de hand van een artikel van O'Brien, Hyman en Kaplan [1]. Een en ander zal in verband gebracht worden met de stabiliteitsdefinities die we hebben leren kennen. We zullen dit weer abstract formuleren met behulp van de in hoofdstuk 2 en 3 ingevoerde notaties.

Gedrag van de rekenfout

Evenals in hoofdstuk 2 voor het differentie-analogon van gewone differentiaalvergelijkingen, zullen we in het geval van partiële differentiaalvergelijkingen definiëren de totale afrondingsfout $|||d|||$, de rekenfout $|||u - u^*|||$ en de totale fout $|||\tilde{u} - u^*|||$. Analooq aan (2.34) kunnen we de totale fout begrenzen door de som van de benaderingsfout en de rekenfout:

$$(2.34) \quad |||\tilde{u} - u^*||| \leq |||\tilde{u} - u||| + |||u - u^*|||.$$

We gaan nu nader in op het gedrag van de rekenfout als functie van τ_k en T . In de praktijk zou men wensen dat voor vaste tijdstappen τ_k en een willekeurig foutenveld d , de rekenfout uniform begrensd blijft voor $T \rightarrow \infty$ en dat deze grens naar nul gaat als $|||d||| \rightarrow 0$. Hieraan beantwoordt de volgende definitie (O'Brien, Hyman en Kaplan).

Definitie 4.1

Het differentieschema $Ru = (u_0, f, \phi)$ is sterk stabiel [1], wanneer de rekenfout $|||u - u^*|||$ voor een willekeurig foutenveld d en voor vaste tijdstappen τ_k uniform begrensd blijft als $T \rightarrow \infty$.

Naast deze sterke stabiliteit voeren O'Brien, Hyman en Kaplan ook het begrip zwakke stabiliteit in.

Definitie 4.2

Het differentieschema $Ru = (u_0, f, \phi)$ is zwak stabiel [1], wanneer de rekenfout $|||u - u^*|||$ voor het foutenveld d , met $d_k = 0$ als $k \neq 0$, en vaste tijdstappen τ_k , uniform begrensd blijft als $T \rightarrow \infty$.

Het is duidelijk dat sterke stabiliteit een veel zwaardere eis is dan zwakke stabiliteit. Slechts begin-randwaarde problemen welke een stationnaire oplossing hebben als $t \rightarrow \infty$ laten zich door sterk stabiele schema's beschrijven. Hierop komen we in het tweede deel van deze syllabus nog terug.

O'Brien, Hyman en Kaplan bepalen zich tot een onderzoek naar zwakke stabiliteit. Zij rechtvaardigen dit door de aanname dat de afrondingsfouten "at random" optreden, zodat geen accumulatie verwacht hoeft te worden.

Stelling 4.1

Schema (3.9) is zwak stabiel, wanneer het O'B-H-K stabiel is.

Bewijs

In het differentieschema (3.9)

$$(3.9) \quad u_{k+1} = A_k u_k + \tau Df_k + B\phi_k$$

introduceren we voor $k = 0$ de afrondingsfoutenfuntie d_0 , dus

$$(4.1) \quad u_1^* = A_0 u_0 + \tau Df_0 + B\phi_0 + d_0 = u_1 + d_0.$$

Voor de verdere berekeningen mogen we de nu ontstane u_1^* als beginfuntie zien, zodat het effect van de eenmalige afrondingsfout opgevat kan worden als een verstoring van de beginvoorwaarden.

We kunnen dit ook inzien door (4.1) te schrijven als

$$(4.1') \quad u_1^* = A_0(u_0 + A_0^{-1}d_0) + \tau Df_0 + B\phi_0,$$

waarin $A_0^{-1}d_0$ een verstoring van u_0 betekent.

Uit definitie 3.10 volgt nu de stelling.

Tot zover de beschouwingen van O'Brien, Hyman en Kaplan.

We zullen nu de sterke stabiliteit van differentieschema's onderzoeken; daartoe voeren we een deelruimte E_G^* van E^* in, bestaande uit roosterfuntties welke op de randpunten Γ_T nul zijn. We bewijzen nu de volgende stelling

Stelling 4.2

Schema (3.9) is sterk stabiel, wanneer de maximale amplificatiefactor α kleiner is dan 1.

Bewijs

Bij het begrip sterke stabiliteit neemt men als uitgangspunt de situatie, dat bij elke berekening afzonderlijk, afrondingsfouten gemaakt worden. We kunnen dit weergeven door in schema (3.9) een term d_k toe te voegen; we krijgen dan

$$(4.2) \quad u_{k+1}^* = A_k u_k^* + \tau D f_k + B \phi_k + d_k.$$

We kunnen nu de invloed van de afrondingsfouten niet zien als een verstoring van de beginvoorwaarden. We splitsen het foutenveld d volgens

$$d = b + \beta,$$

waarin $b \in E_G^*$ en $\beta \in E_\Gamma^*$, dan kan vergelijking (4.2) worden geschreven als

$$(4.3) \quad u_{k+1}^* = A_k u_k^* + \tau D f_k + b_k + B \phi_k + \beta_k.$$

Indien we de operatoren D beperken tot niveaufuncties waarvan de bijbehorende roosterfuncties in E_G^* liggen, kunnen we D inverteren. Omdat ook B inverteerbaar is, kunnen we vergelijking (4.3) schrijven als

$$(4.3') \quad u_{k+1}^* = A_k u_k^* + \tau D(f_k + \tau^{-1} D^{-1} b_k) + B(\phi_k + B^{-1} \beta_k).$$

We voeren in de roosterfuncties e en ϵ uit E^* , door de volgende relaties voor de bijbehorende niveaufuncties uit E_0^* :

$$(4.4) \quad e_k = \tau^{-1} D^{-1} b_k, \quad \epsilon_k = B^{-1} \beta_k.$$

We zien dat het foutenveld d op te vatten is als een verstoring e van de inhomogene term en een verstoring ϵ van de randvoorwaarden.

In de notatie van het voorgaande hoofdstuk heeft vergelijking (4.3') de vorm

$$(4.5) \quad Ru^* = (u_0, f+e, \phi+\epsilon).$$

De differentieoplossing u voldeed aan een schema van de vorm

$$Ru = (u_0, f, \phi),$$

dus voor de discretiseringsfout vinden we (R is lineair en inverteerbaar)

$$u - u^* = R^{-1}(0, e, \varepsilon),$$

derhalve

$$(4.6) \quad |||u - u^*||| \leq |||R^{-1}||| \cdot |||(0, e, \varepsilon)||| = |||R^{-1}||| \cdot (|||e||| + |||\varepsilon|||).$$

We zullen de norm in E^* in die van E_0^* uitdrukken met de relatie (vergelijk (2.21))

$$|||f||| = \max_k ||f_k||.$$

Met behulp van stelling 3.5 (b) en (c) vinden we

$$(4.7) \quad |||u - u^*||| \leq \max_k \left(\frac{1-\alpha^k}{1-\alpha} \right) \cdot (\tau ||D|| + ||B||) \cdot (|||e||| + |||\varepsilon|||).$$

Substitutie van e en ε geeft tenslotte

$$(4.7') \quad |||u - u^*||| \leq \max_k \left(\frac{1-\alpha^k}{1-\alpha} \right) \cdot (\tau ||D|| + ||B||) \cdot (\tau^{-1} ||D^{-1}|| + ||B^{-1}||) \cdot |||d|||,$$

waaruit de stelling direct volgt.

Opmerking 4.1

De ongelijkheid (4.7') geeft geen uniforme bovengrens als $\tau \rightarrow 0$ en T constant. We kunnen bewijzen, dat een schema $Ru = (u_0, f, \phi)$ met een in τ uniform begrensde rekenfout $|||u - u^*|||$ niet consistent kan zijn met een differentiaalvergelijking. Daartoe schrijven we

$$\begin{aligned} u - u^* &= R^{-1}(0, e, \varepsilon) = \tau^{-1} R^{-1}(0, \tau e, \tau \varepsilon) \\ &= \tau^{-1} R^{-1}(0, \bar{e}, \bar{\varepsilon}), \end{aligned}$$

waarin $\bar{e} = \tau e$ en $\bar{\varepsilon} = \tau \varepsilon$. Nu geldt

$$\bar{e}_k = D^{-1} b_k,$$

dus voor $\tau \rightarrow 0$ gaat $\bar{e}_k$ niet naar nul.

Stel dat het differentieschema consistent is met het begin-rand-waardeprobleem

$$L\tilde{U} = (U_0, F, \Phi).$$

Verder veronderstellen we het foutenveld d zodanig, dat voor $\tau \rightarrow 0$ de foutenvelden convergeren naar een continue functie. We zullen de continue functies waartoe $\bar{e}$ en $\bar{\varepsilon}$ naderen voor $\tau \rightarrow 0$ resp. $\bar{E}$ en $\bar{\varepsilon}$ noemen. Uit de vergelijking

$$\tau(u - u^*) = R^{-1}(0, \bar{e}, \bar{\varepsilon})$$

volgt dat dit voor $\tau \rightarrow 0$ overgaat in de relatie

$$L0 = (0, \bar{E}, \bar{\varepsilon}),$$

waarin in elk geval $\bar{E} \neq 0$.

Nu is L lineair, zodat er een tegenspraak bereikt is.

We kunnen dus besluiten dat een willekeurig foutenveld een verstoring van het rekenproces vormt, welke geen R-F stabiliteit toelaat. Daarentegen laat een éénmalige afrondingsfout d_0 wel R-F stabiliteit toe.

Opmerking 4.2

Uit (4.7') volgt voor de rekenfout een gedrag als functie van τ (T constant) van de vorm

$$|||u - u^*||| \sim \tau^{-q} \max_{0 \leq k \leq N} \frac{1 - \alpha^k}{1 - \alpha} \cdot |||d|||,$$

waarin $q > 0$. Indien $\alpha \leq 1 + O(\tau \ln \tau)$ is dit het gedrag dat nog net de F-W stabiliteit garandeert (vergelijk definitie 2.9 en 3.10).

5. Enkele methoden voor stabiliteitsonderzoek

In het voorgaande hoofdstuk hebben we gezien wat voor verandering in de differentieoplossing u optreedt ten gevolge van een verstoring door afrondingsfouten. Het bleek dat de maximale amplificatie-factor α hierin een belangrijke rol speelde. We zullen nu enkele methoden beschrijven, waarmee deze α geanalyseerd kan worden m.a.w. we zullen ons met de stabiliteit t.o.v. de beginvoorwaarden bezighouden.

(a) Directe methode van Todd

Zij gegeven een homogene partiële differentiaalvergelijking met begin- en randvoorwaarden. We denken hierbij in het bijzonder aan een parabolische vergelijking met homogene randvoorwaarden van de eerste soort, bijvoorbeeld de diffusie-vergelijking

$$U_t = U_{xx}$$

op het gebied $0 < x < 1$ en $0 < t \leq T$ met de beginvoorwaarde

$$U(x, 0) = U_0(x)$$

en de randvoorwaarde

$$U(0, t) = U(1, t) = 0.$$

Een dergelijk probleem geeft aanleiding tot een schema van de vorm

$$u_{k+1} = A u_k, \quad k = 0, 1, \dots, N,$$

waarin A een lineaire differentie operator is.

Met behulp van de directe methode (Todd [10], Godunov-Rjabenki [4]) tracht men de norm van de operator A zonder omwegen te bepalen.

Wanneer het interval op de x -as $(m+1)$ roosterpunten bevat, kan A weergegeven worden door een $(m-1) \times (m-1)$ matrix. Bij een expliciet schema is dit bovendien in het algemeen een bandmatrix.

We zullen hier uitsluitend z.g. normale matrices beschouwen.

Definitie 5.1

Een normale matrix is een matrix die commuteert met zijn geconjugueerd getransponeerde.

Bijvoorbeeld alle symmetrische matrices zijn normale matrices.

T.o.v. normen, welke afgeleid zijn van een inwendig product, hebben symmetrische matrices de eigenschap dat de matrix-norm gelijk is aan de spectrale norm (grootste eigenwaarde in absolute waarde). Iets dergelijks geldt voor normale matrices.

Stelling 5.1

T.o.v. een inwendig-product-norm is de norm van een normale matrix gelijk aan zijn spectrum-norm.

Verder zullen we gebruik maken van de stelling:

Stelling 5.2

Indien de matrix A eigenvectoren η_v heeft en eigenwaarden λ_v , dan heeft de matrix A^k dezelfde eigenvectoren en eigenwaarden λ_v^k voor alle gehele waarden van k .

Voor het bewijs van deze stellingen zij verwezen naar de literatuur [7].

De directe methode is toe te passen, wanneer we A kunnen uitdrukken in eenvoudige standaardmatrices, waarvan we de eigenwaarden kennen. In de theorie van Todd [10] neemt de symmetrische $(n \times n)$ matrix

$$(5.1) \quad T_n = (\dots, 1 \quad -2 \quad 1, \dots)$$

een centrale plaats in. We ontlelen aan Todd [10] de volgende stelling:

Stelling 5.3

De matrix T_n heeft eigenvectoren η_v en eigenwaarden λ_v , welke gegeven worden door

$$(5.2) \quad \eta_v = \sin \frac{vj\pi}{n+1} \quad \lambda_v = -4 \sin^2 \frac{v\pi}{2(n+1)}$$

$$j = 1, 2, \dots, n; \quad v = 1, 2, \dots, n.$$

Uit stelling 5.2 en 5.3 volgt nu direct dat de matrix T_n^2 eigenvectoren η_v en eigenwaarden μ_v heeft, welke gegeven worden door

$$(5.3) \quad \eta_v = \sin \frac{vj\pi}{n+1} \quad \mu_v = \lambda_v^2 = 16 \sin^4 \frac{v\pi}{2(n+1)} \quad j, v = 1, 2, \dots, n.$$

Volledigheidshalve geven we hier ook de expliciete vorm voor T_n^2 :

$$(5.4) \quad T_n^2 = (\dots, 1 \quad -4 \quad 6 \quad -4 \quad 1, \dots)$$

waarbij het eerste en laatste diagonaalelement van T_n^2 niet 6 maar 5 is.

Bij de toepassing van de directe methode zullen we de matrices welke door T_n gegenereerd worden laten werken op vectoren, die gevormd worden door de waarden in de inwendige roosterpunten (G_i) van het beschouwde niveau. Dit is mogelijk omdat we ons hier beperken tot randwaarden van de eerste soort, dus de differentieoplossing is op de rand gegeven.

Wanneer we b.v. T_n bekijken, dan zien we dat de werking van T_n overeenkomt met het differentieschema

$$u_{k+1} = (X_- - 2 + X_+)u_k \quad \text{in } x_1, x_2, \dots, x_n.$$

De homogeniteit van de randvoorwaarden zorgt ervoor dat de waarden van u_k in de randpunten x_0 en x_{n+1} , die nodig zijn voor de berekening van u_k in x_1 en x_n , weggelaten kunnen worden. Hierdoor is het differentieschema inderdaad door de symmetrische matrix T_n voor te stellen.

De toepassingsmogelijkheden van de directe methode laten zich het best duidelijk maken aan de hand van enkele voorbeelden.

Voorbeeld 5.1

We beschouwen weer het klassieke expliciete differentieschema voor de diffusievergelijking met homogene randvoorwaarden in $x = 0$ en $x = 1$. Dit kan geschreven worden als

$$u_{k+1} = (rX_- + (1 - 2r) + rX_+)u_k,$$

waarin $r = \frac{\tau}{\xi^2}$ (vergelijk voorbeeld 3.2) en waarin X_- en X_+ in de punten x_1 resp. x_{m-1} aan u_k de waarde nul toekennen.

We schrijven dit schema als matrix-vergelijking:

$$u_{k+1} = A u_k,$$

waarin

$$A = I + r T_{m-1}.$$

Volgens stelling 5.3 zijn de eigenwaarden α_v van deze matrix

$$\alpha_v = 1 - 4r \sin^2 \frac{v\pi}{2m}, \quad v = 1, 2, \dots, m-1.$$

Voor $\alpha = \max_v |\alpha_v|$ vinden we

$$\begin{aligned} \alpha &= \max \left(\left| 1 - 4r \sin^2 \frac{\pi}{2m} \right|, \left| 1 - 4r \sin^2 \frac{m-1}{2m} \pi \right| \right) \\ &= \max \left(\left| 1 - 4r \sin^2 \frac{\pi}{2m} \right|, \left| 1 - 4r \cos^2 \frac{\pi}{2m} \right| \right). \end{aligned}$$

Voor $\alpha \leq 1$ is het schema zowel R-F stabiel als O'B-H-K stabiel. Dit betekent

$$(5.5) \quad r \leq \frac{1}{2 \cos^2 \frac{\pi}{2m}}.$$

Voorbeeld 5.2

We zullen nu een impliciet differentieschema voor de diffusie-vergelijking analyseren:

$$(5.6) \quad \frac{u_{k+1} - u_k}{\tau} = \frac{(X_- - 2 + X_+)u_{k+1} + (X_- - 2 + X_+)u_k}{2\xi^2}.$$

We stellen weer $r = \frac{\tau}{\xi^2}$. Schema (5.6) kan geschreven worden als

$$(5.6') \quad -(rX_- - 2(1+r) + rX_+)u_{k+1} = (rX_- + 2(1-r) + rX_+)u_k.$$

We definiëren een matrix B door

$$(5.7) \quad B = 2I - rT_{m-1}.$$

Hiermee kan (5.6') als de volgende matrix-vergelijking geformuleerd worden

$$Bu_{k+1} = (4I - B)u_k,$$

dus

$$u_{k+1} = (4B^{-1} - I)u_k.$$

De eigenwaarden α_v van $(4B^{-1} - I)$ zijn volgens stelling 5.3

$$\alpha_v = \frac{4}{2 + 4r \sin^2 \frac{v\pi}{2m}} - 1 = \frac{1 - 2r \sin^2 \frac{v\pi}{2m}}{1 + 2r \sin^2 \frac{v\pi}{2m}}, \quad v=1,2,\dots,m-1.$$

Blijkbaar is dit schema onvoorwaardelijk stabiel.

We merken op dat men deze stabiliteit moet bekopen met het inverteren van de matrix B . In numerieke berekeningen betekent dit dat er meer rekentijd per tijdstap nodig is, maar men kan de tijdstappen veel groter kiezen dan in het klassieke expliciete schema. In het tweede deel van deze syllabus zal de diffusievergelijking nauwkeuriger onderzocht worden.

Voorbeeld 5.3

We passen nu de directe methode toe op een 3-niveau-schema. Beschouw de differentiaalvergelijking

$$(5.8) \quad U_{xxxx} + U_{tt} = 0,$$

welke optreedt bij de beschrijving van transversale trillingen in een staaf. Als beginvoorwaarden kiezen we

$$U(x,0) = U_0(x), \quad U_t(x,0) = V_0(x)$$

en voor de randvoorwaarden nemen we

$$U(0,t) = U_{xx}(0,t) = U(1,t) = U_{xx}(1,t) = 0.$$

De afgeleiden in (5.8) vervangen we door centrale differentie quotiënten, zodat de volgende differentievergelijking ontstaat.

$$(5.9) \quad u_{k+1} = -r^2 \left[X_-^2 - 4X_- + (6-2r^{-2}) - 4X_+ + X_+^2 \right] u_k - u_{k-1},$$

waarin $r = \frac{\tau}{\xi^2}$.

In matrix-vorm wordt dit

$$(5.9') \quad u_{k+1} = -r^2 D u_k - u_{k-1}$$

waarin

$$(5.10) \quad D = T_{m-1}^2 - 2r^{-2} I.$$

Met behulp van (5.9') kunnen we u_k uitdrukken in u_0 en u_1 (u_1 volgt uit $v_0 = [v_0]_d$). We vinden

$$(5.11) \quad u_k = p_0^{(k)}(D)u_0 + p_1^{(k)}(D)u_1.$$

$p_0^{(k)}$ en $p_1^{(k)}$ zijn polynomen van de graad $(k-2)$ resp. $(k-1)$ in D .

We zullen eerst deze polynomen nader onderzoeken.

Er geldt:

$$u_k = p_0^{(k)}(D)u_0 + p_1^{(k)}(D)u_1$$

$$u_{k-1} = p_0^{(k-1)}(D)u_0 + p_1^{(k-1)}(D)u_1$$

$$u_{k-2} = p_0^{(k-2)}(D)u_0 + p_1^{(k-2)}(D)u_1.$$

Substitutie hiervan in (5.9') geeft

$$\begin{aligned} p_0^{(k)}(D)u_0 + p_1^{(k)}(D)u_1 &= -r^2 D p_0^{(k-1)}(D)u_0 - r^2 D p_1^{(k-1)}(D)u_1 \\ &\quad - p_0^{(k-2)}(D)u_0 - p_1^{(k-2)}(D)u_1. \end{aligned}$$

Voor de beide polynomen $p_0^{(k)}$ en $p_1^{(k)}$ volgt hieruit de recurrente betrekking

$$p^{(k)}(x) + r^2 x p^{(k-1)}(x) + p^{(k-2)}(x) = 0, \quad k = 2, 3, \dots$$

$$p_0^{(0)}(x) \equiv 1, \quad p_0^{(1)}(x) \equiv 0.$$

$$p_1^{(0)}(x) \equiv 0, \quad p_1^{(1)}(x) \equiv 1.$$

Deze recurrente betrekking heeft oplossingen van de vorm (vergelijk inleiding)

$$\begin{aligned}
 (5.12) \quad p_0^{(k)}(x) &= a_0 \mu_1^k(x) + b_0 \mu_2^k(x) \\
 p_1^{(k)}(x) &= a_1 \mu_1^k(x) + b_1 \mu_2^k(x),
 \end{aligned}$$

waarin $\mu_1(x)$ en $\mu_2(x)$ de wortels zijn van de vergelijking

$$(5.13) \quad \mu^2 + r^2 x \mu + 1 = 0.$$

Het differentieschema (5.9) is stabiel (zie (5.11)) wanneer de operatoren $p_0^{(k)}(D)$ en $p_1^{(k)}(D)$ uniform begrensd zijn voor alle k . Uit (5.10) volgt dat D een symmetrische operator is, zodat $p_0^{(k)}(D)$ en $p_1^{(k)}(D)$ ook symmetrisch zijn. Beschouwen we de inwendig-product-norm voor deze operatoren dan wordt de stabiliteitseis de uniforme begrensdheid van de spectrale normen, dus

$$(5.14) \quad \sigma(p_0^{(k)}(\delta_v)) \text{ en } \sigma(p_1^{(k)}(\delta_v)) \text{ uniform begrensd voor alle } k.$$

Hierin zijn δ_v de eigenwaarden van D .

Uit (5.12) volgt dat μ_1 en μ_2 binnen of op de eenheidscirkel moeten liggen. Toepassing van criterium 2.1 met $S = -r^2 \delta_v$ en $P = 1$ geeft

$$(5.15) \quad r^2 \delta_v - 2 \leq 0 \text{ voor } v = 1, 2, \dots, m-1.$$

Nu wordt δ_v gegeven door

$$\delta_v = 16 \sin^4 \frac{v\pi}{2m} - 2r^{-2}.$$

Substitutie in (5.15) geeft

$$(5.15') \quad 16r^2 \sin^4 \frac{v\pi}{2m} - 4 \leq 0 \text{ voor } v = 1, 2, \dots, m-1.$$

Hieraan is voldaan voor

$$(5.16) \quad r \leq \frac{1}{2 \cos^2 \frac{\pi}{2m}}.$$

Onder deze voorwaarden is het schema zowel R-F stabiel als O'B-H-K stabiel (vergelijk voorbeeld 5.1).

We merken op dat deze directe methode alleen bruikbaar is in heel speciale gevallen (zie ook Godunov en Rjabenki [4]).

(b) Fourier methode van von Neumann

In het artikel "A study of the numerical solution of partial differential equations" van O'Brien, Hyman en Kaplan [1] wordt een basis gelegd voor het gebruik van Fourier methoden bij het onderzoek van de stabiliteit van differentieschema's. Dit artikel is voor een groot deel een uitwerking van de ideeën van J. von Neumann, zoals deze door hem tijdens de oorlog ontwikkeld zijn, maar die hij zelf niet heeft gepubliceerd. We geven hier een samenvatting van deze methode.

Zij gegeven een lineaire scalaire differentievergelijking van de vorm (3.9), waarin de coëfficiënten constant zijn:

$$u_{k+1} = A u_k + \tau D f_k + B \phi_k.$$

We zullen alleen stabiliteit t.o.v. de beginvoorwaarden beschouwen.

Veronderstel dat op $t = 0$ een fout d_0 geïntroduceerd wordt, zodat niet de differentieoplossing u met beginvoorwaarde u_0 , maar de numerieke oplossing u^* met beginvoorwaarde $u_0 + d_0$ verkregen wordt. De roosterfunctie $v = u - u^*$ voldoet wegens de lineariteit van het schema aan dezelfde differentievergelijking maar met beginvoorwaarde d_0 .

We gaan nu voor elke k , v_k ontwikkelen in een Fourierreeks, waarbij we veronderstellen dat v_k van de één-dimensionale plaatsvariabele x_j afhangt. Op $t = 0$ zij

$$(5.17) \quad v_0(x_j) = d_0(x_j) = \sum_{\omega} A_n e^{i\omega x_j}.$$

Omdat d_0 reëel is, is deze reeks eigenlijk een sinus-cosinus reeks. Als bijvoorbeeld de randvoorwaarden van de eerste soort zijn, zodat in de randpunten geen afrondingsfouten optreden, is de reeks (5.17) equivalent met de reeks

$$(5.18) \quad v_0(x_j) = \sum_{n=1}^{m-1} a_n \sin \frac{n j \pi}{m}.$$

Om zo nauw mogelijk aan te sluiten aan het artikel van O'Brien, Hyman en Kaplan stellen we $v_0(x_j)$ steeds voor door (5.17).

In het algemeen zijn de frequenties ω en het aantal termen van de reeks (5.17) tamelijk willekeurig. Om de ontwikkeling in de tijd van de rekenfout v_k te onderzoeken is het vanwege het superpositiebeginsel voldoende het gedrag van de termen van (5.17) afzonderlijk na te gaan. We beschouwen dus een beginfunctie $e^{i\omega x_j}$ en onderzoeken het gedrag in de tijd van de bijbehorende oplossing.

Indien de randvoorwaarden van de eerste soort zijn, of algemener, te vervangen zijn door periodiciteitsvoorwaarden (zie Richtmyer [7]), wordt de door $e^{i\omega x_j}$ voortgebrachte oplossing gevonden in de vorm

$$(5.19) \quad e^{i\omega x_j} \cdot e^{k\gamma}$$

waarin γ een functie is van ω , maar niet van j afhangt. v_k wordt als een som van dergelijke termen gevonden.

O'Brien, Hyman en Kaplan stellen nu de stabiliteitsvoorwaarde:

$$(5.20) \quad |e^\gamma| \leq 1 \quad \text{voor alle } \omega.$$

Het is duidelijk dat dit O'B-H-K stabiliteit (zie definitie 3.10) garandeert.

Opmerking 5.1

Indien we (5.20) vervangen door

$$(5.21) \quad |e^{\frac{\tau}{T} \gamma(\omega)}| \text{ uniform begrensd voor alle } \omega \text{ en } \tau \rightarrow 0$$

krijgen we de R-F stabiliteitsvoorwaarde t.o.v. de beginvoorwaarden (zie Richtmyer [7]).

Opmerking 5.2

Men past de criteria (5.20) en (5.21) ook wel toe op schema's met variërende coëfficiënten en niet-periodieke randvoorwaarden door de partiële oplossing (5.19) in een vast punt x_{j_0} te beschouwen. De criteria (5.20) en (5.21) worden dan plaatsafhankelijk en zijn als het ware

locale stabiliteitscriteria. Alhoewel zulke locale stabiliteitscriteria geen "overall-stabiliteit" garanderen worden ze in numerieke berekeningen veelvuldig gebruikt. We komen hier in het tweede deel nog op terug.

Opmerking 5.3

O'Brien, Hyman en Kaplan passen de Fouriermethode ook toe op hogere orde (scalaire) differentieschema's. Omdat in deze syllabus alle schema's als eerste orde (vector) schema's geformuleerd worden, zullen we in de beschrijving hiervan (einde van dit hoofdstuk) afwijken van het oorspronkelijke artikel (zie ook Richtmyer [7]).

Voorbeeld 5.4

Evenals in de toepassing van de directe methode beschouwen we het klassieke expliciete differentieschema voor de homogene diffusievergelijking met homogene randvoorwaarden van de eerste soort (voorbeeld 5.1).

$$\frac{u_{k+1} - u_k}{\tau} = \frac{X_+ u_k - 2u_k + X_- u_k}{\xi^2}.$$

Substitutie van (5.19) voor $u_k(x_j)$ geeft

$$\frac{e^Y - 1}{\tau} = \frac{1}{\xi^2} (e^{i\omega\xi} - 2 + e^{-i\omega\xi})$$

dus

$$e^Y = 1 - 4r \sin^2 \frac{1}{2} \omega \xi$$

waarin $r = \frac{\tau}{\xi^2}$. We vinden de bekende voorwaarde $r \leq \frac{1}{2}$, wanneer we ω alle reële waarden laten aannemen.

(c) Methode van de scheiding der variabelen

De hierboven beschreven methode was in wezen gebaseerd op de scheiding der variabelen en ontwikkeling naar de eigenfuncties van de operator A . Deze eigenfuncties waren de complexe e machten $e^{i\omega x_j}$.

We zullen nu dit scheiden der variabelen algemener formuleren en

tevens uitbreiden tot het geval van een stelsel differentievergelijkingen.

We beschouwen eerst een scalair differentieschema van de vorm

$$(5.23) \quad u_{k+1} = A u_k$$

met een beginvoorwaarde u_0 .

We zoeken een particuliere oplossing voor (5.23) in de vorm

$$(5.24) \quad u_k = \theta_k \eta(j).$$

Substitutie in (5.23) geeft

$$\theta_{k+1} \eta(j) = \theta_k A \eta(j)$$

ofwel

$$(5.25) \quad \frac{\theta_{k+1}}{\theta_k} = \frac{A \eta(j)}{\eta(j)}.$$

Het linkerlid is alleen afhankelijk van k , het rechterlid alleen van j , dus beide leden zullen gelijk zijn aan een constante α :

$$\frac{\theta_{k+1}}{\theta_k} = \alpha = \frac{A \eta(j)}{\eta(j)}.$$

We zien dat α de groei van het tijdsafhankelijke deel van de onderzochte particuliere oplossing bepaalt. Verder is α een eigenwaarde van A en $\eta(j)$ een eigenfunctie van A :

$$(5.26) \quad A \eta(j) = \alpha \eta(j).$$

We zullen nu aannemen, dat A een stelsel eigenfuncties $\eta^{(n)}(j)$ heeft, zodanig dat we de beginfunctie u_0 kunnen schrijven als

$$(5.27) \quad u_0 = \sum_n \theta_0^{(n)} \eta_n(j),$$

waarin $\theta_0^{(n)}$ de coëfficiënten van de ontwikkeling zijn. We zien nu dat $u_1, u_2, \dots, u_k$ gegeven worden door:

$$\begin{aligned}
 u_1 &= \sum_n \theta_0^{(n)} \alpha_n \eta_n(j) = \sum_n \theta_1^{(n)} \eta_n(j) \\
 u_2 &= \sum_n \theta_1^{(n)} \alpha_n \eta_n(j) = \sum_n \theta_2^{(n)} \eta_n(j) \\
 (5.28) \quad & \dots\dots\dots \\
 u_k &= \sum_n \theta_k^{(n)} \eta_n(j).
 \end{aligned}$$

Het verband tussen de coëfficiënten $\theta_k^{(n)}$ en de amplificatie-factoren α_n wordt gegeven door

$$(5.29) \quad \theta_k^{(n)} = \alpha_n^k \theta_0^{(n)}.$$

Uit (5.28) en (5.29) volgt

$$(5.30) \quad |u_k| \leq \sum_n |\alpha_n|^k |\theta_0^{(n)} \eta_n(j)| \leq \sigma^k(A) \sum_n |\theta_0^{(n)} \eta_n(j)|,$$

waarin $\sigma(A)$ de spectraal-norm van A is.

We hebben dus de volgende stelling bewezen

Stelling 5.1

Wanneer de beginfunctie u_0 te schrijven is als een lineaire combinatie van de eigenfuncties van de operator A en wanneer de eigenwaarden van A binnen of op de eenheidscirkel liggen, dan is het schema

$$u_{k+1} = A u_k \text{ O'B-H-K stabiel.}$$

Opmerking 5.4

Het is niet zonder meer mogelijk om een equivalente stelling voor de R-F stabiliteit uit te spreken, omdat als $\tau \rightarrow 0$ (T constant) het aantal eigenfuncties $\eta_n(j)$ onbeperkt toe kan nemen, zodat in (5.30) een oneindige som optreedt. Indien echter

$$\eta_n(j) = e^{i\omega_n x_j},$$

dan is er wel een dergelijke stelling te formuleren, waarin dan de eigen-

waarden α_n binnen cirkels met stralen $1 + O(\tau)$ voor $\tau \rightarrow 0$ moeten liggen (vergelijk opmerking 5.1).

Verder merken we op dat de eis $\sigma(A) \leq 1$ in stelling 5.1 in het algemeen minder streng is dan de eis $\alpha = \|A\| \leq 1$, welke per definitie O'B-H-K stabiliteit impliceert; echter stelling 5.1 eist verder dat de eigenfuncties van A een volledig stelsel vormen.

Voorbeeld 5.5

Beschouw het differentieschema uit voorbeeld 5.4. De eigenfuncties van de differentieoperator zijn

$$\eta_n(j) = \sin \frac{nj\pi}{m}, \quad n = 1, 2, \dots, m-1.$$

Inderdaad kan elke beginfunctie $u_0(j)$ met $u_0(0) = u_0(m) = 0$ als een sinus-reeks geschreven worden, zodat men alleen hoeft te eisen dat de corresponderende eigenwaarden binnen of op de eenheidscirkel liggen. Overigens is dit een bijzonder geval van de Fourier-methode.

Voorbeeld 5.6

Vervangen we in voorbeeld 5.4 de differentieformules in de randpunten door

$$u_k(-1) = u_k(1) \quad \text{en} \quad u_k(m-1) = u_k(m+1)$$

dan is het stelsel

$$\eta_n(j) = \cos \frac{nj\pi}{m}, \quad n = 1, 2, \dots, m-1$$

een volledig stelsel eigenfuncties voor alle beginfuncties, die aan bovenstaande randvoorwaarden voldoen. Ook dit is een bijzonder geval uit de Fourier-methode.

We zullen nu het voorgaande generaliseren, zodat de theorie toepasbaar is op schema's van de vorm

$$(5.31) \quad \vec{u}_{k+1} = A \vec{u}_k,$$

waarin $\vec{u}_k$ nu een vectorfunctie voorstelt. Eenvoudigheidshalve beschouwen

we het geval van één plaatsvariabele x_j , dus $\vec{u}_k = \vec{u}_k(j)$. Het aantal componenten van $\vec{u}_k$ noemen we s .

In analogie met het voorgaande zoeken we weer naar een particuliere oplossing van (5.31) van de vorm

$$(5.32) \quad \vec{u}_k = \vec{\theta}_k \eta(j).$$

Hierin is $\vec{\theta}_k$ een vector met s componenten, welke alleen van k afhangen en $\eta(j)$ een scalarfunctie, die alleen van j afhangt.

Substitutie van (5.32) in (5.31) geeft

$$(5.33) \quad \vec{\theta}_{k+1} \eta(j) = A \vec{\theta}_k \eta(j).$$

We kunnen nu niet voor $A \vec{\theta}_k \eta(j)$ schrijven $\vec{\theta}_k A \eta(j)$, omdat de operator A zowel op $\vec{\theta}_k$ als op $\eta(j)$ werkt. Om deze werkingen te kunnen scheiden zullen we de structuur van A nader onderzoeken.

A is in wezen een operator die voor te stellen is door een matrix van plaatsdifferentieoperatoren A_{pq}

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1s} \\ A_{21} & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ A_{s1} & \dots & \dots & A_{ss} \end{pmatrix}$$

De operatoren A_{pq} werken uitsluitend op $\eta(j)$.

De plaatsfuncties $\eta(j)$ moeten nu zodanig zijn, dat de operatorenmatrix A in een gewone matrix B met elementen b_{pq} overgaat. Er moet dus gelden

$$(5.34) \quad A_{pq} \eta(j) = b_{pq} \eta(j),$$

m.a.w. de functie $\eta(j)$ moeten eigenfuncties zijn van de operatoren A_{pq} voor elke p en q . De matrix B wordt amplificatie-matrix genoemd en ook wel met $\hat{A}$ aangegeven.

We nemen nu aan dat de operatoren A_{pq} inderdaad een gemeenschappelijk stelsel eigenfuncties $\{\eta_n(j)\}$ hebben, zodanig dat we de oplossing

van het stelsel differentievergelijkingen (5.31) kunnen schrijven als

$$(5.35) \quad \vec{u}_k = \sum_n \vec{\theta}_k^{(n)} \eta_n(j).$$

Wegens de lineariteit van het probleem is het voldoende om de termen in (5.35) afzonderlijk te beschouwen. Voor elke n geldt

$$A_{pq} \eta_n(j) = b_{pq}^{(n)} \eta_n(j),$$

zodat

$$\vec{\theta}_{k+1}^{(n)} \eta_n(j) = B_n \vec{\theta}_k^{(n)} \eta_n(j),$$

waarin B_n de elementen $b_{pq}^{(n)}$ heeft.

Omdat B_n onafhankelijk van k is geldt

$$(5.36) \quad \vec{\theta}_k^{(n)} = B_n^k \vec{\theta}_0^{(n)}.$$

Deze betrekking is het analogon van (5.29) voor scalaire differentievergelijkingen. Uit (5.35) en (5.36) volgt de stelling:

Stelling 5.2

Wanneer de eigenfuncties van de operatoren A_{pq} een gemeenschappelijk stelsel eigenfuncties bevatten, welke een volledig stelsel vormen voor de beginfunctie $\vec{u}_0$ en wanneer de normen van de amplificatie-matrices B_n binnen of op de eenheidscirkel liggen, dan is het schema $\vec{u}_{k+1} = A \vec{u}_k$ O'B-H-K stabiel.

Opmerking 5.5

In het voorgaande hebben we steeds verondersteld dat we slechts met één plaats variabele x_j te doen hebben. Alle beschouwingen gaan echter door voor een meerdimensionaal plaatsrooster met roosterpunten $(x_1, \dots, x_r)$. We noteren zo'n roosterpunt met $\vec{x}_j$. Ook de index van de eigenfuncties van A_{pq} wordt een vector-index, dus $\eta_n(j)$ gaat over in $\eta_n(\vec{j})$, waarin $\vec{j}$ het vector-karakter van $\vec{x}_j$ aanduidt. Het in de praktijk veel optredende stelsel

$$\{\eta_n(j)\} = \{\exp i \omega_n x_j\}$$

zullen we in meerdimensionale roosters noteren met

$$(5.37) \quad \{\eta_{\vec{\omega}}(\vec{j})\} = \{\exp i \vec{\omega} \cdot \vec{x}_j\}.$$

Opmerking 5.6

Een verdere generalisatie bestaat uit het toelaten van schema's, waarin de operator A van k afhangt. Bovenstaande beschouwingen blijven vrijwel ongewijzigd, wanneer de eigenfuncties η niet van k afhangen. Relatie (5.36) gaat dan over in

$$(5.38) \quad \vec{\theta}_{k+1}^{(n)} = B_n(k) \vec{\theta}_k^{(n)},$$

zodat

$$\vec{\theta}_k^{(n)} = \prod_{l=k-1}^0 B_n(l) \cdot \vec{\theta}_0^{(n)}.$$

Dit leidt tot beschouwingen van de amplificatie-matrix met de grootste norm.

De von Neumann-voorwaarde

Indien voor de eigenfuncties van de operatoren A_{pq} het stelsel (5.37) gevonden wordt, dan is het schema R-F stabiel indien de normen van de amplificatie-matrices voldoen aan

$$(5.39) \quad \|B_{\vec{\omega}}^{(k)}\| \leq 1 + O(\tau), \quad \tau \rightarrow 0.$$

Voor een bewijs van deze bewering zij verwezen naar de literatuur [7].

Indien de operatoren B normaal zijn is het voldoende, dat de eigenwaarden β voldoen aan

$$(5.40) \quad |\beta| \leq 1 + O(\tau) \text{ voor } \tau \rightarrow 0 \text{ en alle } \vec{\omega}.$$

Dit is de z.g. von Neumann-voorwaarde. Voor niet-normale matrices is deze voorwaarde niet voldoende voor R-F stabiliteit, maar wel noodzakelijk.

Voorbeeld 5.7

Tot slot zullen we het in voorbeeld 5.3 onderzochte schema m.b.v. bovenstaande methoden onderzoeken.

Schema (5.9) schrijven we in de volgende equivalente vorm

$$\begin{pmatrix} u \\ v \end{pmatrix}_{k+1} = \begin{pmatrix} -r^2(X_-^2 - 4X_- + 6 - 2r^{-2} - 4X_+ + X_+^2) & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}_k$$

Omdat de randvoorwaarden periodiek zijn vormt het stelsel

$$\{\exp(i \omega x_j)\} = \{\exp(i \omega j \xi)\}$$

een volledig stelsel eigenfuncties van de operatoren A_{pq} . De amplificatiematrices worden:

$$B_\omega = \begin{pmatrix} -r^2(2 \cos 2\omega\xi - 8 \cos \omega\xi + 6 - 2r^{-2}) & -1 \\ 1 & 0 \end{pmatrix}$$

Deze matrices zijn niet-normaal, dus de von Neumann-voorwaarde geeft nog geen uitsluitel over R-F stabiliteit. Werken we (5.40) voor dit voorbeeld uit, dan vinden we

$$r \leq \frac{1}{2}.$$

We hebben met de directe methode gezien, dat dit voldoende is voor zowel R-F als O'B-H-K stabiliteit.

Literatuur

- [1] O'Brien, G.G., Hyman, M.A. and Kaplan, S.: A study of the numerical solution of partial differential equations. J. Math. Physics, vol. 29, p. 223 (1951).
- [2] Forsythe, G.E. and Wasow, W.R.: Finite difference methods for partial differential equations. John Wiley, New York (1960).
- [3] Fox, L.: Numerical solutions of ordinary and partial differential equations. Pergamon Press, Oxford (1962).
- [4] Godunov, J.K. and Rjabenki, V.S.: Theory of difference schemes - an introduction. North-Holland Publishing Company, Amsterdam (1964).
- [5] Kreiss, H.O.: Die Stabilitätsdefinitionen für Differenzengleichungen die partielle Differentialgleichungen approximieren. Nordisk Tidskrift for Informations Behandling, Bind 2, Hefte nr. 3 (1962).
- [6] Milne, W.E.: Numerical Calculus. Princeton Press (1949).
- [7] Richtmyer, D.R.: Difference methods for initial value problems. Interscience, New York (1960).
- [8] Rjabenki, V.S. und Filippow, A.F.: Über die Stabilität von Differenzengleichungen. VEB Deutscher Verlag der Wissenschaften, Berlin (1960).
- [9] Saul'yev, V.K.: Integration of equations of parabolic type by the method of nets. Pergamon Press, Oxford (1964).
- [10] Todd, J.: A direct approach to the problem of stability in the numerical solution of partial differential equations. Communications on Pure and Applied Mathematics, Vol. IX, nr. 3 (1956).

