

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O).

MC SYLLABUS 33

P.J. VAN DER HOUWEN

**BEREKENING
VAN WATERSTANDEN
IN ZEEËN EN RIVIEREN**

MATHEMATISCH CENTRUM AMSTERDAM 1977

AMS(MOS) subject classification scheme (1970): 65M05, 65M60

ISBN 90 6196 138 6

INHOUD

VOORWOORD

vii

1.	De waterbeweging in rivieren en zeeën	1
1.1	Bewegings-vergelijkingen voor stroming in een rivier	2
1.2	Bewegings-vergelijkingen voor stroming in een ondiepe zee	3
1.3	De continuïteitsvergelijking	4
1.4	Het volledige stelsel differentiaalvergelijkingen	4
1.5	Ruwe analyse van het gelineariseerde stelsel	5
2.	Numerieke aspecten van de methode van Euler	10
2.1	De methode van Euler	11
2.2	Stabiliteit van de methode van Euler	13
2.3	Stabiliteit t.o.v. bodemwrijving, Corioliskracht en convectieve traagheidsterman	16
2.4	De voortplanting van fouten in de methode van Euler	18
2.5	Het Routh-Hurwitz-criterium	23
2.6	De constante coëfficiënten methode	25
2.7	Stabiliteit t.o.v. bodemwrijving, Corioliskracht en convectieve traagheidstermen (vervolg)	26
2.8	Stabiliteit van het volledige stelsel ondiep-water-vergelijkingen	30
3.	Gegeneraliseerde Eulermethoden	36
3.1	Vergelijking van de stabiliteit van de meest expliciete en meest impliciete vorm	36
3.2	Het volledige differentieschema voor de meest impliciete vorm ...	39
3.3	Het Fischer-Hansen-schema	42
3.4	Het Sielecki-schema	45
3.5	Gesymmetriseerde schema	46
3.6	Stabiliteit van het gesymmetriseerde schema	48
3.7	Impliciete schema's	52

4.	Gestabiliseerde Runge-Kuttamethoden	58
4.1	Algemene Runge-Kuttamethoden	58
4.2	Runge-Kutta-methoden van de tweede orde	60
4.3	Het stabiliteitspolynoom	61
4.4	Optimale formules voor differentiaalvergelijkingen met imaginair spectrum	63
4.5	Gegeneraliseerde Runge-Kuttamethoden	67
5.	Meerstapsmethoden	69
5.1	Het "leap-frog" schema	69
6.	Dissipatieve termen en andere stabiliserende factoren	73
6.1	Effeningsoperatoren, de Shuman-operator	74
6.2	Dissipatieve termen, het Lax-Wendroff-schema	81
7.	Discretisering van $\partial/\partial x$ en $\partial/\partial y$	85
7.1	Gewogen centrale differenties	86
7.2	De afbreekfout van de gewogen centrale differentie-operator	87
7.3	Het spectrum van gewogen centrale differentie-operatoren	88
7.4	Hogere orde discretisaties	90
7.5	Niet-uniforme roosters	91
7.6	Het spectrum van de centrale differentie-operator op een niet-uniform rooster	94
7.7	Discretisering van de randvoorwaarden	96
8.	Programmatuur voor ondiep-water-problemen	98
	Referenties	98

VOORWOORD

De in deze syllabus behandelde stof stemt overeen met een aan de Universiteit van Amsterdam gegeven caput-college over numerieke oplossingsmethoden voor zogenaamde "ondiep-water-problemen". De omvang bedraagt ongeveer 1 jaaruur.

De veronderstelde voorkennis is inleidende analyse, lineaire algebra en elementaire numerieke wiskunde.

Na een ruwe verkenning van de analytische aspecten van de ondiep-water-vergelijkingen in hoofdstuk I, worden in de hoofdstukken II tot en met V achtereenvolgens behandeld de methode van Euler, gegeneraliseerde Euler-methoden, gestabiliseerde Runge-Kutta-methoden en meerstapsmethoden. In het bijzonder worden de voor- en nadelen van deze methoden besproken voor het ondiep-water-probleem.

In hoofdstuk VI worden technieken geanalyseerd om rekenschema's die bij daadwerkelijke toepassing instabiliteiten ontwikkelen, te stabiliseren.

Hoofdstuk VII is gewijd aan de discretisering van de plaatsafgeleiden op uniforme en niet-uniforme roosters.

Naar aanleiding van dit college zijn er twee programma's ontwikkeld voor de integratie van de ondiep-water-vergelijkingen. Voor een beschrijving van deze programma's zij verwezen naar ROZENHART [1977] en SOMMEIJER [1977].

Tenslotte wil ik mijn dank betuigen aan Mevr. E. Pins voor de zeer accurate wijze waarop deze syllabus getypt is, de heer R.T. Baanders voor het tekenen van de figuren en de heren D. Zwarts en J. Schipper voor het drukken en binden van dit boekje.

P.J. van der Houwen.

I. DE WATERBEWEGING IN ZEEËN EN RIVIEREN

In het door de wind in beweging gebrachte water kunnen we twee typen golfbewegingen waarnemen: *lange* golven en *korte* golven. In de lange golven is de golflengte veel groter dan de diepte van het water; de verticale versnellingen in het water kunnen over 't algemeen verwaarloosd worden ten opzichte van de horizontale versnellingen en het water dat aanvankelijk in een verticaal vlak zit, blijft in een verticaal vlak gedurende de oscillaties, dus elk verticaal vlak van de vloeistof beweegt in zijn geheel. In de Engelstalige literatuur wordt dit type golfbeweging *tidal wave* of ook wel *long waves in shallow water* genoemd.

In de korte golven zijn de verstoringen van het waterniveau relatief gering; de verticale versnelling mag niet langer verwaarloosd worden en de golflengte is veel kleiner dan de diepte van de rivier of de zee. Deze rimpelingen van het wateroppervlak worden *oppervlakte golven* genoemd. In deze syllabus zullen we ons beperken tot de *lange golven*.

Naast de wind en eventueel andere atmosferische krachten, speelt in de waterbeweging nog een tweede drijvende kracht een belangrijke rol en wel de *getijdewerking* tengevolge van de aantrekkingskracht van zon en maan. Dit effect zullen we niet beschouwen; een uitvoerige behandeling kan men b.v. in DRONKERS [1964] vinden.

Wanneer we uitgaan van een zee of rivier in rust, dus zonder windveld of atmosferische krachten, en we laten vervolgens deze krachten op het water inwerken, dan zal het water door de uitgeoefende schuifkracht in beweging komen waardoor een aantal andere krachten geactiveerd worden en wel een *wrijvingskracht* tengevolge van de waterbeweging over de rivier- of zeebodem en de *Corioliskracht* tengevolge van de draaiing van de aarde. Daarnaast hebben we nog de *vervalkracht* tengevolge van de zwaartekracht, welke uiteraard ook al in de stationaire toestand werkzaam was. Stel dat deze krachten te samen in een horizontale kracht $\vec{K}$ resulteren, dan geldt het volgende evenwicht van krachten:

$$(1.0) \quad \frac{D\vec{u}}{Dt} = \vec{K},$$

waarin $\vec{u}$ de horizontale snelheid van een waterdeeltje is en D/Dt de "meebewegingsafgeleide" voorstelt, dat wil zeggen $D\vec{u}/Dt$ is de variatie in $\vec{u}$ wanneer we hetzelfde deeltje in het oog houden. Men kan laten zien dat

$$\frac{D\vec{u}}{Dt} = \frac{\partial \vec{u}}{\partial t} + \left(u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} \right) \vec{u},$$

waarin u en v de componenten van $\vec{u}$ in x - respectievelijk y -richting voorstellen. De tweede term in het rechterlid wordt de *convectieve traagheids-term* genoemd.

Vergelijking (1.0) is de *bewegingsvergelijking* voor het water. In de volgende twee paragrafen zullen de bewegingsvergelijkingen voor de stroming in een rivier en in een zee nader onderzocht worden. Voor de afleiding van deze vergelijkingen zij verwezen naar e.g. DRONKERS [1964].

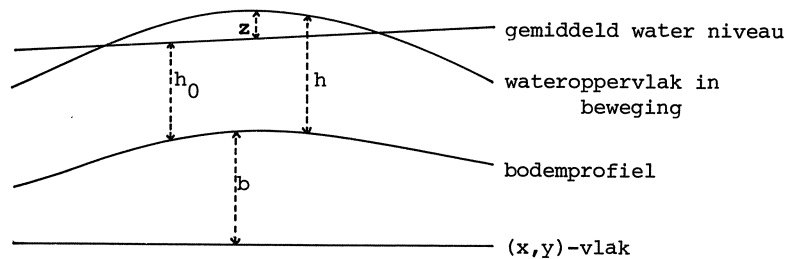


Fig. 1.1. Definitie van de grootheden b , h_0 , h en z

1.1. Bewegingsvergelijkingen voor stroming in een rivier

In het geval van stroming in een rivier gaat vergelijking (1.0) over in (zie DRONKERS [1964, p. 192])

$$(1.1) \quad \begin{aligned} \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} &= -\lambda u - g \frac{\partial}{\partial x} (h+b) + \frac{\omega_x}{h} \\ 0 &= -\omega u - g \frac{\partial}{\partial y} (h+b) + \frac{\omega_y}{h} . \end{aligned}$$

Hierin hebben de verschillende grootheden de volgende betekenis:

t	tijd
x,y	plaatscoördinaten langs de rivier (positief in stroomopwaartse richting) respectievelijk loodrecht op de rivier
u	de over de diepte gemiddelde snelheid eveneens positief in stroomopwaartse richting
λ	wrijvingscoëfficiënt tengevolge van de bodem
g	versnelling van de zwaartekracht
h	diepte beneden het wateroppervlak (zie figuur 1.1)
b	hoogte van het bodemprofiel ten opzichte van het (x,y)-vlak (zie figuur 1.1)
ω	Coriolis-parameter: $\omega = 2 \alpha \sin \phi$, waarin α de hoeksnelheid van de aarde is en ϕ de breedtegraad in het punt (x,y)
ω_x, ω_y	door wind en atmosferische drukverdeling bepaalde termen (zie DRONKERS [1964, p. 188]).

Merk op dat de stroming in een rivier als *eendimensionaal* beschouwd wordt ($v=0$). Verder kan voor relatief smalle rivieren de Coriolis-kracht en het windeffect in de y-richting verwaarloosd worden zodat de tweede bewegingsvergelijking in (1.1) dan buiten beschouwing blijft. De wrijvingscoëfficiënt λ is een gegeven functie van de onbekenden u, v and h, die in het algemeen zeer kleine waarden aanneemt (zie paragraaf 8.1).

1.2. Bewegingsvergelijkingen voor stroming in een ondiepe zee

In een zee is er uiteraard sprake van een *tweedimensionale* stroming. Er geldt (zie DRONKERS [1964, p.190]):

$$\begin{aligned}
 (1.2) \quad & \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} - \omega v = -\lambda u - g \frac{\partial}{\partial x} (h+b) + \frac{\omega_x}{h} \\
 & \frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + \omega v = -\lambda v - g \frac{\partial}{\partial y} (h+b) + \frac{\omega_y}{h} .
 \end{aligned}$$

Hierin zijn u en v de componenten van de over de diepte gemiddelde snelheid $\vec{u}$ in respectievelijk de x- en y-richting (x en y zijn nu rechthoekig).

In de literatuur over ondiep-waterberekeningen worden de bewegingsvergelijkingen dikwijls uitgedrukt in de waterverhoging z en de diepte h_0 ten opzichte van het gemiddelde waterniveau. In (1.2) substitueert men dan (zie

figuur 1.1)

$$(1.3) \quad h = h_0 + z.$$

Bovendien wordt voor het gemiddelde waterniveau vrijwel altijd een horizontale vlak genomen, iets wat voor de meeste zeeën gerechtvaardigd is. Men kan dan het (x,y)-vlak hiermee laten samenvallen, zodat

$$(1.4) \quad b = -h_0,$$

waarmee b uit (1.2) geëlimineerd wordt.

1.3. De continuïteitsvergelijking

Behalve aan de bewegingsvergelijkingen moeten de variabelen u, v en h (of z) ook voldoen aan de behoudswet van materie, de zogenaamde *continuïteitsvergelijking*. Deze vergelijking brengt in rekening dat er in een eenheid van volume evenveel water in komt als uitgaat. Dus effecten als verdamping, regen en samendrukking worden *niet* in rekening gebracht. Een eenvoudige beschouwing leidt tot de vergelijking (zie DRONKERS [1964, p.179])

$$(1.3) \quad \frac{\partial}{\partial x} (hu) + \frac{\partial}{\partial y} (hv) + \frac{\partial h}{\partial t} = 0.$$

1.4. Het volledige stelsel differentiaalvergelijkingen

De vergelijkingen (1.1) en (1.3) bepalen samen de stroming in een rivier. Wanneer we de tweede vergelijking in (1.1) weglaten, dus ons beperken tot smalle rivieren, kunnen we schrijven

$$(1.6) \quad \frac{\partial}{\partial t} \begin{pmatrix} u \\ h \end{pmatrix} = - \begin{pmatrix} \lambda & g \frac{\partial}{\partial x} \\ h \frac{\partial}{\partial x} & 0 \end{pmatrix} \begin{pmatrix} u \\ h \end{pmatrix} - u \frac{\partial}{\partial x} \begin{pmatrix} u \\ h \end{pmatrix} + \begin{pmatrix} f_x \\ 0 \end{pmatrix},$$

waarin $f_x = w_x/h - g\partial b/\partial x$.

Voor de waterbeweging in zeeën vinden we uit (1.2) en (1.5) iets dergelijks:

$$(1.7) \quad \frac{\partial}{\partial t} \begin{pmatrix} u \\ v \\ h \end{pmatrix} = - \begin{pmatrix} \lambda & -\omega & g \frac{\partial}{\partial x} \\ \omega & \lambda & g \frac{\partial}{\partial y} \\ h \frac{\partial}{\partial x} & h \frac{\partial}{\partial y} & 0 \end{pmatrix} \begin{pmatrix} u \\ v \\ h \end{pmatrix} - \left(u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} \right) \begin{pmatrix} u \\ v \\ h \end{pmatrix} + \begin{pmatrix} f_x \\ f_y \\ 0 \end{pmatrix},$$

waarin $f_x = \omega_x/h - g\partial b/\partial x$ en $f_y = \omega_y/h - g\partial b/\partial y$.

We kunnen (1.6) en (1.7) compact schrijven als

$$(1.8) \quad \frac{\partial}{\partial t} \vec{s} = \mathcal{D}^{(1)} \vec{s} + \mathcal{D}^{(2)} \vec{s} + \vec{f},$$

waarin $\vec{s}$, $\mathcal{D}^{(1)}$, $\mathcal{D}^{(2)}$, en $\vec{f}$ direct uit (1.6) respectievelijk (1.7) volgen. In deze vergelijking is de term $\mathcal{D}^{(2)} \vec{s}$ van tweede orde belang en wordt vaak weggelaten. Numerieke experimenten hebben aangetoond dat de oplossing in geringe mate verandert. Verder wordt in $\mathcal{D}^{(1)}$ en $\vec{f}$ de diepte h dikwijls vervangen door de diepte h_0 ten opzichte van het gemiddelde waterniveau. Dit is gerechtvaardigd wanneer de waterverhoging z klein is ten opzichte van de diepte. Door tenslotte nog de wrijvingscoëfficiënt λ constant te veronderstellen, gaat (1.8) over in een *lineair* stelsel differentiaalvergelijkingen. In de literatuur vormt dit lineaire stelsel het uitgangspunt voor vele analyses. Bij daadwerkelijke berekeningen dient men echter uit te gaan van de oorspronkelijke niet-lineaire stelsels (1.6) en (1.7).

1.5. Ruwe analyse van het gelineariseerde stelsel

We zullen de gelineariseerde vorm van (1.6) en (1.7) hier nog wat nader bekijken. In het geval van (1.6) geldt in componenten uitgeschreven

$$(1.6') \quad \begin{aligned} \frac{\partial}{\partial t} u &= -\lambda u - g \frac{\partial}{\partial x} h + f_x, \\ \frac{\partial}{\partial t} h &= -h_0 \frac{\partial}{\partial x} u. \end{aligned}$$

Differentiatie van de tweede vergelijking naar t geeft

$$(1.9) \quad \frac{\partial^2}{\partial t^2} h = -h_0 \frac{\partial^2}{\partial x \partial t} u,$$

en differentiatie van de eerste vergelijking naar x geeft

$$\frac{\partial^2}{\partial x \partial t} u = -\lambda \frac{\partial}{\partial x} u - g \frac{\partial^2}{\partial x^2} h + \frac{\partial}{\partial x} f_x = \frac{\lambda}{h_0} \frac{\partial}{\partial t} h - g \frac{\partial^2}{\partial x^2} h + \frac{\partial}{\partial x} f_x.$$

Substitutie hiervan in (1.9) levert een vergelijking voor h waarin u niet meer voorkomt:

$$(1.10) \quad \frac{\partial^2}{\partial t^2} h = -\lambda \frac{\partial}{\partial t} h + gh_0 \frac{\partial^2}{\partial x^2} h - h_0 \frac{\partial}{\partial x} f_x.$$

Stellen we vervolgens

$$(1.11) \quad h(x,t) = e^{-\frac{1}{2}\lambda t} \psi(x,t)$$

dan vinden we voor ψ de vergelijking

$$(1.10') \quad \frac{\partial^2 \psi}{\partial t^2} = gh_0 \frac{\partial^2 \psi}{\partial x^2} + \frac{1}{4} \lambda^2 \psi - e^{-\frac{1}{2}\lambda t} h_0 \frac{\partial}{\partial x} f_x.$$

In het algemeen is λ^2 verwaarloosbaar klein zodat we te maken hebben met de inhomogene golfvergelijking. Wanneer f_x een constante is dan volgt uit (1.10') en (1.11) dat de waterverhoging een gedempte golfbeweging uitvoert die zich met een snelheid $\sqrt{gh_0}$ voortplant; voor constante h_0 geldt namelijk

$$(1.12) \quad h(x,t) = e^{-\frac{1}{2}\lambda t} [a(x - \sqrt{gh_0}t) + b(x + \sqrt{gh_0}t)],$$

waarin a en b willekeurige functies zijn (dit is eenvoudig door substitutie in (1.10) te verifiëren). De functies a en b worden bepaald door de beginvoorwaarden, bijvoorbeeld

$$(1.13) \quad \begin{aligned} h(x,0) &= G_1(x) \\ \frac{\partial h}{\partial t}(x,0) &= G_2(x), \end{aligned}$$

waarin G_1 en G_2 gegeven functies zijn op een interval $[x_b, x_e]$ (x_b en x_e stellen begin en eindpunt van de rivier voor). We vinden dan

$$(1.14) \quad \begin{aligned} a(x) &= G_1(x) - \frac{G_2(x)}{2\sqrt{gh_0}} - \frac{\lambda}{4\sqrt{gh_0}} \int_{x_b}^x G_1(\xi) d\xi, \\ b(x) &= G_1(x) - a(x). \end{aligned}$$

Een belangrijke conclusie die men uit (1.12) en (1.13) kan trekken, is dat de beginvoorwaarden de oplossing slechts in een *beperkt* gebied van het (x,t) -vlak bepalen. Immers als G_1 en G_2 en daarmee a en b slechts op het

interval $[x_b, x_e]$ gegeven zijn dan moeten de argumenten van a en b in (1.12) ook binnen dit interval blijven. Dit leidt tot

$$(1.15) \quad x - \sqrt{gh_0} t \geq x_b, \quad x + \sqrt{gh_0} t \leq x_e.$$

Deze ongelijkheden bepalen een driehoek in het (x,t) -vlak (zie figuur 1.2).

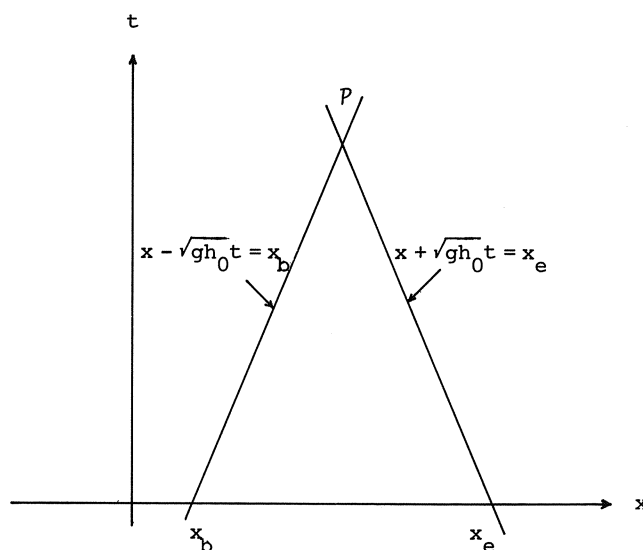


Fig. 1.2. Beïnvloedings- en afhankelijkheidsgebied

Deze driehoek wordt het *beïnvloedingsgebied* van de beginvoorwaarden genoemd, en het interval (x_b, x_e) het *afhankelijkheidsgebied* van het punt P ; de lijnen $x \pm \sqrt{gh_0} t = 0$ zijn de *karakteristieken* van het probleem. Wanneer men buiten het door de karakteristieken bepaalde gebied de oplossing wil weten, dan moeten meer gegevens over de stroming bekend zijn. Bijvoorbeeld, wanneer de rivier twee zeeën met elkaar verbindt (Panama-kanaal), zullen de waterhoogten in begin en eindpunt als constanten h_b en h_e beschouwd kunnen worden. Toevoeging van deze randvoorwaarden aan de differentiaalvergelijking (1.10) en de beginvoorwaarden (1.13) zorgt er voor dat de oplossing voor alle $t \geq 0$ in het interval $[x_b, x_e]$ bepaald is.

In het geval van een tweedimensionale stroming kan men uit de gelineariseerde vorm van (1.7) afleiden

$$(1.16) \quad \frac{\partial^2}{\partial t^2} h = -\lambda \frac{\partial}{\partial t} h + h_0 g \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) h - h_0 \left(\frac{\partial}{\partial x} f_x + \frac{\partial}{\partial y} f_y \right) + \omega \left(\frac{\partial}{\partial y} u - \frac{\partial}{\partial x} v \right).$$

De stromings componenten u en v zijn nog steeds aanwezig maar verdwijnen wanneer we de Corioliskracht zouden verwaarlozen. Op dezelfde wijze als voor de eendimensionale stroming vinden we voor λ^2 verwaarloosbaar klein

$$(1.16') \quad h(x,y,t) = e^{-\frac{1}{2}\lambda t} \psi(x,y,t)$$

$$\frac{\partial^2 \psi}{\partial t^2} = gh_0 \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} \right) - e^{\frac{1}{2}\lambda t} h_0 \left(\frac{\partial}{\partial x} f_x + \frac{\partial}{\partial y} f_y \right).$$

Het is eenvoudig te verifiëren dat voor constante f_x , f_y en h_0 een klasse van oplossingen gegeven wordt door

$$(1.17) \quad h(x,y,t) = e^{-\frac{1}{2}\lambda t} [a(lx + my - \sqrt{gh_0} t) + b(lx + my + \sqrt{gh_0} t)],$$

waarin $l^2 + m^2 = 1$ en a en b weer willekeurige functies zijn. De waterverhoging voert dus een golfbeweging uit die zich met een snelheid $\sqrt{gh_0}$ voortplant en uitdempt in de tijd evenredig met $\exp(-\frac{1}{2}\lambda t)$. Evenals bij het eendimensionale geval is ook hier sprake van een beïnvloedingsgebied en een afhankelijkheidsgebied. Het beïnvloedingsgebied van een punt P op de t -as is het gebied dat omhult wordt door alle vlakken $lx + my + \sqrt{gh_0} t = \text{const}$ die door het punt P gaan. Dit blijkt precies omsloten te worden door de kegel

$$(1.18) \quad x^2 + y^2 - gh_0 (t - t_p)^2 = 0,$$

ook wel de *karakteristiekekegel* genoemd. Het afhankelijkheidsgebied van P wordt omsloten door de cirkel

$$(1.18') \quad x^2 + y^2 = gh_0 t_p^2.$$

Wanneer de beginvoorwaarden in dit afhankelijkheidsgebied gegeven zijn, is de oplossing daarmee voor alle punten *binnen* de kegel (1.18) bepaald.

Wil men de waterverhoging voor alle t in een gegeven punt (x,y) van de beschouwde zee weten, dan moeten weer randvoorwaarden aan de differentiaalvergelijking toegevoegd worden. We onderscheiden oceaانبegrenzungen en kustbegrenzungen. Langs oceaانبegrenzungen is het gebruikelijk de water-

verhoging z constant te nemen in de tijd. We zullen steeds aannemen dat $z = 0$ ofwel

$$(1.19) \quad h(x, t) = h_0(x) \text{ langs de oceaانبegrenzing.}$$

In de globale waterbeweging die we hier beschouwen, kunnen de kusten als verticale bassin-randen beschouwd worden. Een dergelijke begrenzing dwingt het water tot een beweging evenwijdig aan de kusten, dus de kustvoorwaarde wordt

$$(1.20) \quad \vec{u} \text{ evenwijdig aan de kustbegrenzing.}$$

We zullen deze voorwaarde in formulevorm gieten; daartoe gaan we terug naar de oorspronkelijke bewegingsvergelijkingen (1.2) voor tweedimensionale stroming. We schrijven (1.2) in de vectorvorm

$$(1.21) \quad \frac{\partial}{\partial t} \vec{u} = -\lambda \vec{u} + R\vec{u} - g\nabla h - (\vec{u} \cdot \nabla) \vec{u} + \vec{f},$$

waarin

$$(1.22) \quad \nabla = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix}, \quad R = \begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix}, \quad \vec{f} = \begin{pmatrix} f_x \\ f_y \end{pmatrix},$$

en $(\vec{u} \cdot \nabla)$ het inwendig product tussen $\vec{u}$ en ∇ voorstelt. De vector $\vec{f}$ is gelijk aan de in (1.7) gedefinieerde vector waarin de laatste component weggelaten is. Laat $\vec{c}$ de eenheidsvector evenwijdig aan de kust zijn in positieve zin (zie figuur 1.3). Vormen we nu het inwendig product tussen $\vec{c}$ en $\frac{\partial}{\partial t} \vec{u}$, dan volgt uit (1.21)

$$(1.21') \quad \frac{\partial}{\partial t} (\vec{c} \cdot \vec{u}) = -\lambda (\vec{c} \cdot \vec{u}) + (\vec{c} \cdot R\vec{u}) - g(\vec{c} \cdot \nabla h) - (\vec{c} \cdot (\vec{u} \cdot \nabla) \vec{u}) + (\vec{c} \cdot \vec{f}).$$

Aangezien $R\vec{u}$ loodrecht op $\vec{u}$ staat is $(\vec{c} \cdot R\vec{u}) = 0$. Verder volgt uit (1.20) dat $(\vec{c} \cdot \vec{u}) \vec{c} = \vec{u}$; vermenigvuldigen we (1.21') dus met $\vec{c}$ dan vinden we langs de kusten de vergelijking

$$(1.21'') \quad \frac{\partial}{\partial t} \vec{u} = -\lambda \vec{u} - g(\vec{c} \cdot \nabla h) \vec{c} - (\vec{c} \cdot (\vec{u} \cdot \nabla) \vec{u}) \vec{c} + (\vec{c} \cdot \vec{f}) \vec{c}.$$

Samen met de continuïteitsvergelijking (1.3) bepaalt deze vergelijking de stroming langs de kust.

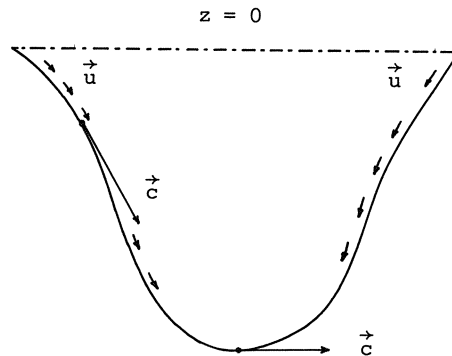


Fig. 1.3. Stroming evenwijdig aan de kust volgens
 $\vec{u} = (\vec{c} \cdot \vec{u}) \vec{c}$

II. NUMERIEKE ASPECTEN VAN DE METHODE VAN EULER

In het eerste hoofdstuk hebben we iets over de oplossing van de rivier- en de zee-vergelijkingen kunnen zeggen door een reeks vereenvoudigingen te veronderstellen zoals geen Corioliskracht, constante bodemwrijving, diepte en waterverhoging relatief klein, geen convectieve traagheidstermen, enz.

In een numerieke aanpak hoeven wat dit betreft geen concessies gedaan te worden; de problemen in numerieke oplossings-methoden liggen in de eerste plaats op het gebied van de *stabiliteit*, dat wil zeggen de gevoeligheid van het rekenproces voor verstoringen van de gegevens, en in de tweede plaats op het gebied van de *nauwkeurigheid*. Weliswaar, kan de nauwkeurigheid bij stabiele rekenprocessen in het algemeen opgevoerd worden door het discrete model waar het rekenproces vanuit gaat, te verfijnen, maar de hieraan verbonden rekentijden kunnen onaanvaardbaar lang worden. Men dient dan niet het *model* maar de *methode* te verfijnen. In de derde plaats speelt de benodigde geheugenruimte bij implementatie op een rekenmachine een rol. Men zal de waterbeweging in een zee al gauw in een duizendtal punten van het (x,y)-vlak willen representeren, zodat al direct met het u, v en h-veld 3000 getallen gemoeid zijn. Het zuinigste rekenproces heeft al een veelvoud van dit aantal nodig, zodat voor kleinere rekenmachines het geheugengebruik een rol kan gaan spelen. In dit hoofdstuk zullen we aan de hand van de methode van Euler een aantal van deze numerieke aspecten illustreren en in het bijzonder

nagaan wat de consequenties zijn voor de integratie van de "ondiepwatervergelijkingen". We zullen daarbij de randvoorwaarden buiten beschouwing laten, dat wil zeggen we beschouwen een oneindig grote zee.

In het vervolg zullen numerieke benaderingen met hoofdletters aangegeven worden; zo zullen U , V , H , $\vec{U}$ en $\vec{S}$ de numerieke benaderingen van u , v , h , $\vec{u}$ en $\vec{s}$ voorstellen.

2.1. De methode van Euler

In de numerieke analyse van gewone differentiaalvergelijkingen is de meest eenvoudige numerieke oplossingsmethode die van Euler. Voor de scalaire vergelijking

$$(2.1) \quad \frac{ds}{dt} = r(t, s)$$

wordt de methode van Euler gedefinieerd door het volgen van het richtingsveld van de differentiaalvergelijking (zie figuur 2.1). Stel dat in het beginpunt

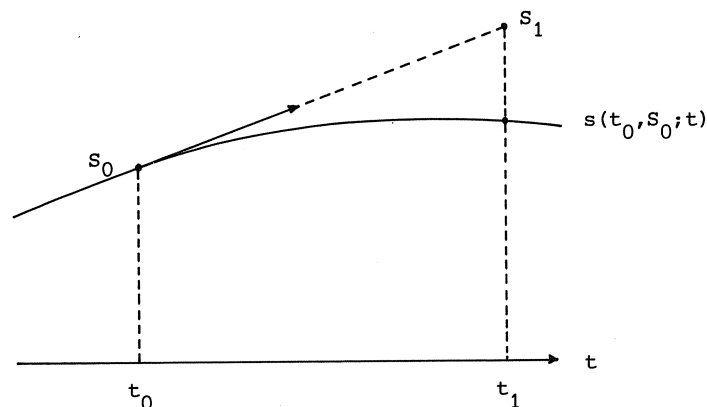


Fig. 2.1. Methode van Euler

t_0 de onbekende s gegeven is door S_0 ; door het punt (t_0, S_0) gaat een oplossing van de differentiaalvergelijking (we nemen aan dat het rechterlid r voldoende "glad" is), welke we met $s(t_0, S_0; t)$ aangeven. In de methode van Euler wordt nu in het punt (t_0, S_0) een raaklijn aan deze integraalkromme getrokken en de waarde S_1 die in t_1 bereikt wordt, geldt als benadering voor

$s(t_0, s_0; t_1)$. Dus

$$s_1 = s_0 + (t_1 - t_0) \left[\frac{d}{dt} s(t_0, s_0; t) \right]_{t_1} = s_0 + (t_1 - t_0) r(t_0, s_0).$$

Zo kunnen we op dezelfde wijze uit s_1 een benadering s_2 in t_2 berekenen, enz. We krijgen dan de algoritme

$$(2.2) \quad s_{n+1} = s_n + \tau_n r(t_n, s_n), \quad n = 0, 1, \dots,$$

waarin τ_n de integratiestap is, gedefinieerd door

$$(2.3) \quad \tau_n = t_{n+1} - t_n.$$

We kunnen (2.2) direct generaliseren voor *stelsels* differentiaalvergelijkingen van de vorm

$$(2.1') \quad \frac{d\vec{s}}{dt} = \vec{r}(t, \vec{s}),$$

namelijk

$$(2.2') \quad \vec{s}_{n+1} = \vec{s}_n + \tau_n \vec{r}(t_n, \vec{s}_n), \quad n = 0, 1, \dots$$

In het voorgaande zijn we er stilzwijgend vanuit gegaan dat r in (2.1) en $\vec{r}$ in (2.1') niet meer dan een scalaire functie respectievelijk een vectorfunctie is. Formeel kunnen we de Eulermethode ook toepassen op partiële differentiaalvergelijkingen, waarin $\vec{r}$ differentiaties naar andere variabelen bevat. Voor de partiële differentiaalvergelijking (1.8) krijgen we op deze wijze

$$(2.4) \quad \vec{s}_{n+1} = \vec{s}_n + \tau_n (\mathcal{D}_n^{(1)} + \mathcal{D}_n^{(2)}) \vec{s}_n + \tau_n \vec{F}_n, \quad n = 0, 1, \dots,$$

waarin $\vec{F}_n$ de in paragraaf 1.4 gedefinieerde vector $\vec{f}$ op het tijdstip t_n voorstelt. Wanneer $\vec{s}_0$ en $\vec{F}_n$, $n = 0, 1, \dots$, voldoende vaak differentieerbaar zijn, kunnen de $\vec{s}_n$, $n = 1, 2, \dots$, door herhaald toepassen van de differentiaaloperatoren $\mathcal{D}_n^{(1)}$ en $\mathcal{D}_n^{(2)}$ met behulp van (2.4) in principe gevonden worden. In de praktijk worden de differentiaaloperatoren $\mathcal{D}_n^{(1)}$ en $\mathcal{D}_n^{(2)}$ door *differentie-operatoren* vervangen, waarmee de *analytische* operatie differentieren in *algebraïsche* operaties (optellen en vermenigvuldigen) overgaat. Deze zogenaamde *discretisatie* zal in hoofdstuk VII ter sprake komen.

2.2. Stabiliteit van de methode van Euler

Stel dat we door afrondingsfouten met de Eulerformule (2.2) in plaats van S_n de waarde $\tilde{S}_n$ verkregen hebben. Het verschil tussen S_{n+1} en $\tilde{S}_{n+1}$ wordt dan gegeven door

$$(2.5) \quad S_{n+1} - \tilde{S}_{n+1} = S_n - \tilde{S}_n + \tau_n [r(t_n, s_n) - r(t_n, \tilde{s}_n)].$$

Voor voldoende kleine verstoringen, dus $|S_n - \tilde{S}_n| \ll |S_n|$, kunnen we schrijven

$$(2.5') \quad S_{n+1} - \tilde{S}_{n+1} \approx [1 + \tau_n \frac{\partial r}{\partial s}(t_n, s_n)](S_n - \tilde{S}_n),$$

waarin $\frac{\partial r}{\partial s}(t_n, s_n)$ de partiele afgeleide van de functie r naar s in het punt (t_n, s_n) voorstelt. Het ligt nu voor de hand om van stabiliteit te spreken wanneer de *amplificatiefactor* $(S_{n+1} - \tilde{S}_{n+1}) / (S_n - \tilde{S}_n)$ in modulus kleiner dan 1 is, dus

$$(2.6) \quad |1 + \tau_n \frac{\partial r}{\partial s}(t_n, s_n)| < 1.$$

Immers dan worden gemaakte fouten direct weer uitgedempt en kan geen opeenhoping van numerieke fouten optreden. We zullen stabiliteit in deze zin *sterke stabiliteit* noemen en we spreken van *zwakke stabiliteit* wanneer de amplificatiefactor *gelijk* aan 1 is. Uit (2.6) volgen de stabiliteitsvoorwaarden

$$(2.7) \quad \frac{\partial r}{\partial s}(t_n, s_n) < 0, \quad \tau_n < \frac{2}{-\frac{\partial r}{\partial s}(t_n, s_n)}.$$

De eerste voorwaarde houdt in dat de differentiaalvergelijking zélf stabiel moet zijn. Dit is als volgt in te zien: laten s en $\tilde{s}$ twee naburige oplossingen van (2.1) zijn, dan geldt

$$\frac{d}{dt}(s - \tilde{s}) = r(t, s) - r(t, \tilde{s}) \approx \frac{\partial r}{\partial s}(t, s)(s - \tilde{s}),$$

waar uit volgt dat de oplossingen s en $\tilde{s}$ de neiging hebben naar elkaar toe te gaan wanneer $\frac{\partial r}{\partial s}(t, x)$ negatief is.

De tweede voorwaarde in (2.7) betekent dat de integratiestap voldoende klein moet zijn.

Voor stelsels differentiaalvergelijkingen kunnen we iets analoogs doen. Vergelijking (2.5') gaat dan over in

$$(2.8) \quad \vec{S}_{n+1} - \vec{S}_{n+1} \cong [I + \tau_n J(t_n, \vec{S}_n)](\vec{S}_n - \vec{S}_n),$$

waarin $J(t_n, \vec{S}_n)$ een matrix van partiele afgeleiden van de componenten van $\vec{r}$ naar de componenten van $\vec{S}_n$ in het punt $(t_n, \vec{S}_n)$ is; daarbij bevat de i -de rij van J alle partiele afgeleiden van de i -de component van $\vec{r}$. J wordt de *Jacobiaan* van de differentiaalvergelijking genoemd. De matrix $I + \tau_n J(t_n, \vec{S}_n)$ zullen we de *amplificatiematrix* A_n noemen. Het analogon van (2.6) wordt

$$(2.9) \quad \|A_n\| = \|I + \tau_n J(t_n, \vec{S}_n)\| < 1,$$

waarin $\|\cdot\|$ de een of andere matrixnorm voorstelt. Men spreekt dan van stabiliteit in de norm $\|\cdot\|$. In deze syllabus zullen we steeds de *spectrale norm* $\|\cdot\|_2$ gebruiken; deze is als volgt gedefinieerd: zij A een willekeurige matrix en $\|\vec{S}\|_2$ de Euclidische norm van de vector $\vec{S}$, dan is

$$(2.10) \quad \|A\|_2 = \sup_{\|\vec{S}\|_2=1} \|A\vec{S}\|_2.$$

STELLING 2.1. Zij A^* de toegevoegd complexe getransponeerde matrix bij A en $\sigma(AA^*)$ de spectrale radius (in modulus grootste eigenwaarde) van AA^* , dan geldt

$$(2.11) \quad \|A\|_2 = \sqrt{\sigma(AA^*)}.$$

Wanneer A een normale matrix is ($AA^* = A^*A$), dan geldt

$$(2.12) \quad \|A\|_2 = \sigma(A).$$

BEWIJS.

In de eerste plaats geldt

$$\begin{aligned} \|A\|_2 &= \sup_E \|A\vec{S}\|_2 = \sup_E \sqrt{(A\vec{S}, A\vec{S})} = \sup_E \sqrt{(\vec{S}, A^*A\vec{S})} \leq \sup_E \sqrt{\|AA^*\|_2 \|\vec{S}\|_2^2} = \\ &= \sqrt{\|AA^*\|_2}, \end{aligned}$$

waarin E de eenheidsbol in de $\vec{S}$ -ruimte voorstelt en $(\cdot, \cdot)$ het met de Euclidische norm consistente inwendige product is. Omdat AA^* hermitisch is ($AA^* = (AA^*)^*$), dus zeker normaal is, vinden we door toepassing van (2.12)

$$\|A\|_2 \leq \sqrt{\|AA^*\|_2} = \sqrt{\sigma(AA^*)}.$$

Rest nog te bewijzen dat $\|A\|_2$ precies gelijk is aan $\sqrt{\sigma(AA^*)}$; dit volgt uit

$$\begin{aligned} \|AA^*\|_2 &\leq \|A\|_2 \|A^*\|_2 = \|A\|_2 \sup_E \sqrt{(A^*S, A^*S)} = \|A\|_2 \sup_E \sqrt{(\vec{S}, AA^* \vec{S})} \leq \\ &\leq \|A\|_2 \sqrt{\|AA^*\|_2}, \end{aligned}$$

zodat

$$\|A\|_2 \geq \sqrt{\|AA^*\|_2},$$

waarmee

$$\|A\|_2 = \sqrt{\|AA^*\|_2}.$$

Om relatie (2.12) te kunnen bewijzen doen we een beroep op de fundamentele stelling van Schur en Toeplitz welke zegt dat elke (vierkante) matrix A te schrijven is als UTU^* , waarin U een unitaire matrix is ($U^*=U^{-1}$) en T een boven-driehoekmatrix is. Uit de relatie $AA^* = A^*A$ volgt direct dat T ook normaal is dus

$$TT^* = T^*T.$$

Schrijft men dit expliciet uit dan is direct te zien dat T een diagonaal-matrix is. Omdat T gelijkvormig is met A zijn de diagonaalelementen van T juist de eigenwaarden van A . Laten $\vec{e}_1$ en $\vec{e}_2$ nu twee orthogonale eigenvectoren zijn van T met eigenwaarden α_1 en α_2 , dan zijn $U\vec{e}_1$ en $U\vec{e}_2$ eigenvectoren van A die ook orthogonaal zijn, immers

$$A(U\vec{e}_1) = UD^*(U\vec{e}_1) = UD\vec{e}_1 = \alpha_1(U\vec{e}_1)$$

en

$$(U\vec{e}_1, U\vec{e}_2) = (\vec{e}_1, U^*U\vec{e}_2) = (\vec{e}_1, \vec{e}_2) = 0.$$

Een normale matrix bezit dus een orthogonaal stelsel eigenvectoren. Hieruit volgt:

$$\begin{aligned} \|A\|_2 &= \sup_E \sqrt{(A\vec{S}, A\vec{S})} = \sup_E \sqrt{\left(A \sum_j c_j U\vec{e}_j, A \sum_j c_j U\vec{e}_j \right)} = \\ &= \sup_E \sqrt{\left(\sum_j \alpha_j c_j U\vec{e}_j, \sum_j \alpha_j c_j U\vec{e}_j \right)} = \sup_E \sqrt{\sum_j (\alpha_j c_j U\vec{e}_j, \alpha_j c_j U\vec{e}_j)} \leq \\ &\leq \sigma(A). \end{aligned}$$

Aangezien een eigenwaarde nooit groter kan zijn dan de norm van de matrix geldt ook $\|A\|_2 \geq \sigma(A)$ waarmee relatie (2.12) bewezen is.

Wanneer de amplificatiematrix een *niet-normale* matrix is en het spectrum van het product $A_n A_n^*$, dus van

$$(2.13) \quad [I + \tau_n J(t_n, \vec{S}_n)] [I + \tau_n J^*(t_n, \vec{S}_n)] ,$$

te moeilijk te onderzoeken valt, kan men proberen $J(t_n, \vec{S}_n)$ te schrijven als de som van een normale matrix J_0 en een kleine verstoring Q_n , dus

$$A_n = I + \tau_n J_0 + \tau_n Q_n.$$

Volgens stelling 2.1 geldt

$$(2.14) \quad \|A_n\|_2 = \|I + \tau_n J(t_n, \vec{S}_n)\|_2 \leq \sigma(I + \tau_n J_0) + \tau_n \|Q_n\|_2 ,$$

en de stabiliteitsvoorwaarde (2.9) gaat over in

$$(2.9') \quad \sigma(I + \tau_n J_0) < 1 - \tau_n \|Q_n\|_2 .$$

In de volgende paragraaf zullen we bovenstaande resultaten toepassen op de bewegingsvergelijking (1.2) van een tweedimensionale stroming waarin we de vervalkracht en het windveld weglaten.

2.3. Stabiliteit t.o.v. bodemwrijving, Corioliskracht en convectieve traagheidstermen

Beschouw de vergelijking (1.2) met weglating van de vervalkracht en het windveld:

$$(1.2') \quad \frac{\partial}{\partial t} \vec{u} = \begin{pmatrix} -\lambda & \omega \\ -\omega & -\lambda \end{pmatrix} \vec{u} - \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix} \vec{u} .$$

Deze vergelijking beschrijft het effect van de bodemwrijving, de Corioliskracht en de convectieve traagheidstermen op de snelheid $\vec{u}$.

Toepassing van het Eulerschema (2.2') geeft de algoritme

$$(2.15) \quad \vec{U}_{n+1} = \left[I + \tau_n \begin{pmatrix} -\lambda & \omega \\ -\omega & -\lambda \end{pmatrix} - \tau_n \begin{pmatrix} \frac{\partial U_n}{\partial x} & \frac{\partial U_n}{\partial y} \\ \frac{\partial V_n}{\partial x} & \frac{\partial V_n}{\partial y} \end{pmatrix} \right] \vec{U}_n.$$

Wanneer we de afhankelijkheid van de operator tussen de vierkante haken van de stroming $\vec{U}$ verwaarlozen, dan is de amplificatiematrix van de vorm

$$(2.16) \quad A_n = I + \tau_n J_0 + \tau_n Q_n$$

met

$$J_0 = \begin{pmatrix} -\lambda & \omega \\ -\omega & -\lambda \end{pmatrix}, \quad Q_n = - \begin{pmatrix} \frac{\partial U_n}{\partial x} & \frac{\partial U_n}{\partial y} \\ \frac{\partial V_n}{\partial x} & \frac{\partial V_n}{\partial y} \end{pmatrix},$$

waarin λ en ω onafhankelijk van x en y verondersteld worden. Het is eenvoudig na te gaan dat J_0 een normale matrix is zodat (2.13) en derhalve (2.9') van toepassing is. Rest nog de spectrale radius van $I + \tau_n J_0$ en de spectrale norm van Q_n te berekenen. De eigenwaarden van J_0 voldoen aan de vergelijking

$$\delta^2 + 2\lambda\delta + \lambda^2 + \omega^2 = 0,$$

waaruit volgt

$$(2.17) \quad \delta_{\pm} = -\lambda \pm i\omega.$$

Dus

$$(2.18) \quad \sigma(I + \tau_n J_0) = \sqrt{(1 - \lambda\tau_n)^2 + \tau_n^2 \omega^2}.$$

De waarde van $\|Q_n\|$ kan berekend worden met behulp van formule (2.11); na enig rekenwerk vinden we

$$(2.19) \quad \|Q_n\|_2 = \sqrt{\sigma(Q_n Q_n^*)} = \sqrt{\frac{1}{2} \left\{ \left(\frac{\partial U_n}{\partial x} \right)^2 + \left(\frac{\partial U_n}{\partial y} \right)^2 + \left(\frac{\partial V_n}{\partial x} \right)^2 + \left(\frac{\partial V_n}{\partial y} \right)^2 \right\} + \sqrt{\left\{ \left(\frac{\partial U_n}{\partial x} - \frac{\partial V_n}{\partial y} \right)^2 + \left(\frac{\partial U_n}{\partial y} + \frac{\partial V_n}{\partial x} \right)^2 \right\} \left\{ \left(\frac{\partial U_n}{\partial x} + \frac{\partial V_n}{\partial y} \right)^2 + \left(\frac{\partial U_n}{\partial y} - \frac{\partial V_n}{\partial x} \right)^2 \right\}}}.$$

Voorwaarde (2.9') wordt nu

$$(2.19) \quad \tau_n < 2 \frac{\lambda - \|Q_n\|_2}{\lambda^2 + \omega^2 - \|Q_n\|_2^2}, \quad \lambda > \|Q_n\|_2.$$

Hieruit volgt dat we alleen zeker van stabiliteit zijn wanneer de bodemwrijving voldoende groot is ten opzichte van de veranderingen van de stroming $\vec{U}$ in het (x,y) -vlak. De terechte vrees ontstaat dat het Euler-schema niet zo geschikt is om de "ondiep-water" vergelijkingen te integreren. Hierbij moet wel opgemerkt worden dat we een iets optimistischer beeld krijgen wanneer we in plaats van (2.9') eisen dat de spectraal radius van het product (2.13) kleiner dan 1 is. Deze afleiding laten we aan de lezer over.

2.4. De voortplanting van fouten in de methode van Euler

In de vorige paragrafen hebben we steeds de voorwaarde gesteld dat de fout opgebouwd na n integratiestappen in de volgende integratiestap in norm kleiner moet worden. In deze paragraaf zullen we nagaan hoe het de fout $\vec{S}_n - \vec{\tilde{S}}_n$ in de navolgende integratiestappen vergaat. Dit gedrag van de fout drukken we uit in de spectrale radius van de amplificatiematrix

$$A_n = I + \tau_n J(t_n, \vec{S}_n)$$

en in de spectrale norm van de matrix die de relatieve verandering van de amplificatiematrix voor $t > t_n$ weergeeft.

Fundamenteel voor beschouwingen over voortplanting van fouten zijn de volgende twee lemma's.

LEMMA 2.1. *Zij $q(A)$ de hoogste orde van alle diagonale submatrices J_q van de Jordan-voorstelling van de matrix A waarvoor de spectrale radius gelijk aan die van A is, dan geldt*

$$\|A^v\|_2 \cong \text{const} \left[\frac{v}{\sigma(A)} \right]^{q(A)-1} [\sigma(A)]^v \quad \text{als } v \rightarrow \infty.$$

BEWIJS.

Hiervoor zij verwezen naar bijvoorbeeld VARGA [1962, p.64]. We merken hier alleen op dat $q(A)$ altijd kleiner dan of gelijk aan de multiplicititeit van die eigenwaarden α van A waarvoor geldt $|\alpha| = \sigma(A)$. Geven we de maximale multiplicititeit van deze eigenwaarden met $\mu(A)$ aan, dan geldt dus

$$q(A) \leq \mu(A).$$

LEMMA 2.2. Wanneer de matrix A_j te schrijven is als $B_j + C_j$, waarbij B_j voldoet aan

$$\left\| \prod_{j=1+k}^1 B_j \right\| \leq K e^{-\alpha(k+1)}, \quad \begin{aligned} l &= 0, 1, \dots, v-k \\ k &= 0, 1, \dots, v \end{aligned}$$

en C_j voldoet aan

$$\|C_j\| \leq L_j, \quad j = 0, 1, \dots, v,$$

waarin $K \geq 1$, α en L_j zekere reële getallen zijn en v een geheel getal is, dan geldt

$$\left\| \prod_{j=v}^0 A_j \right\| \leq K e^{-\alpha(v+1)} \prod_{j=v}^0 (1 + K e^{\alpha} L_j).$$

BEWIJS.

Dit lemma is een uitbreiding van een lemma van Kreiss en Strang ($\alpha=0$, $L_j=L$, STRANG [1964]) en werd bewezen door ENGQUIST [1969].

Deze lemma's kunnen gebruikt worden om inzicht te krijgen hoe de fout $\vec{S}_n - \vec{\tilde{S}}_n$ zich ontwikkelt in de volgende integratiestappen. Uit (2.8) leiden we af dat

$$(2.20) \quad \vec{S}_{n+v+1} - \vec{\tilde{S}}_{n+v+1} = A_{n+v} (\vec{S}_{n+v} - \vec{\tilde{S}}_{n+v}) = \dots = \prod_{j=v}^0 A_{n+j} (\vec{S}_n - \vec{\tilde{S}}_n).$$

De amplificatiematrices A_{n+j} schrijven we als

$$A_{n+j} = \bar{A}_n + (A_{n+j} - \bar{A}_n)$$

en we passen lemma 2.2 toe met $B_j = \bar{A}_n$ en $C_j = A_{n+j} - \bar{A}_n$. Volgens lemma 2.1 vinden we voor het product van de matrices B_j de uitdrukking

$$\left\| \prod_{j=1+k}^1 B_j \right\|_2 = \|\bar{A}_n^{k+1}\|_2 \cong \text{const} \left[\frac{k+1}{\sigma(\bar{A}_n)} \right]^{q(\bar{A}_n)-1} [\sigma(\bar{A}_n)]^{k+1} \quad \text{als } k \rightarrow \infty.$$

Laten we nu de matrix $\bar{A}_n$ zodanig kiezen dat $q(\bar{A}_n) = 1$ en $\sigma(\bar{A}_n) = \sigma(A_n)$. In de gevallen waarin voor A_n al geldt dat $q(A_n) = 1$ (bijvoorbeeld wanneer $\mu(A_n) = 1$), kiezen we eenvoudig $\bar{A}_n = A_n$. Indien A_n hieraan niet voldoet, verstoren we A_n zodanig dat $q(\bar{A}_n) = 1$ wordt terwijl de spectrale radius constant blijft.

Voor voldoende grote waarde van K geldt nu

$$\left\| \prod_{j=1+k}^1 B_j \right\|_2 = \|\bar{A}_n^{k+1}\|_2 \leq K e^{(k+1) \ln \sigma(A_n)}, \quad k = 0, 1, \dots, \nu.$$

Stel verder dat

$$\|C_j\|_2 = \|A_{n+j} - \bar{A}_n\|_2 \leq L, \quad j = 0, 1, \dots, \nu;$$

dan volgt uit lemma 2.2

$$(2.21) \quad \left\| \prod_{j=\nu}^0 A_{n+j} \right\|_2 \leq K e^{(\nu+1) \ln \sigma(A_n)} \prod_{j=\nu}^0 (1 + K L e^{-\ln \sigma(A_n)}) = K [\sigma(A_n) + K L]^{\nu+1}.$$

Voor $\bar{A}_n \neq A_n$ is L een maat voor de "veranderlijkheid" van de amplificatiematrix A_{n+j} . Uit (2.20) en (2.21) volgt nu

$$(2.22) \quad \|\vec{S}_{n+\nu+1} - \vec{S}_{n+\nu+1}\|_2 \leq K [\sigma(A_n) + K L]^{\nu+1} \|\vec{S}_n - \vec{S}_n\|_2.$$

In het bijzondere geval van *lineaire* differentiaalvergelijkingen en *constante* integratiestappen geldt $A_n = I + \tau J = A$, zodat direct uit lemma 2.1 volgt

$$(2.22') \quad \|\vec{S}_{n+\nu+1} - \vec{S}_{n+\nu+1}\|_2 \leq K \left(\frac{\nu+1}{\sigma(A)} \right)^{q(A)-1} [\sigma(A)]^{\nu+1} \|\vec{S}_n - \vec{S}_n\|_2,$$

waarin $q(A) \leq \mu(A)$.

Uit deze relaties volgt direct

STELLING 2.2.

(a) De tentijde t_n opgebouwde fout wordt in volgende integratiestappen uitgedempt wanneer

$$\sigma(A_n) < 1 - K L.$$

(b) Voor *lineaire* stelsels en *constante* integratiestappen wordt de tentijde t_n opgebouwde fout in volgende integratiestappen uitgedempt wanneer

$$\sigma(A) < 1$$

en versterkt volgens ν^{q-1} , $q \leq \mu(A)$, wanneer

$$\sigma(A) = 1.$$

Voor *lineaire* stelsels die met *constante* integratiestappen geïntegreerd worden, geeft deze stelling concrete en ook realistische informatie over de verdere ontwikkeling van de op $t = t_n$ aanwezige fout. Voor *niet-lineaire* stelsels zijn de voorwaarden van de stelling moeilijk te verifiëren in praktische toepassingen en bovendien zijn ze in het algemeen te pessimistisch. We leren er echter in elk geval uit dat het raadzaam is om er voor te zorgen dat

$$(2.23) \quad \sigma(A_n) = \sigma(I + \tau_n J(t_n, \vec{S}_n)) < 1$$

en meer marge te nemen naar mate het stelsel differentiaalvergelijkingen sterker niet-lineair is of wanneer men de integratiestap sterk wil variëren.

In het voorgaande hebben we de integratiestappen $\{\tau_{n+j}\}_1^V$ steeds als een vooraf *gegeven* rij getallen beschouwd. Men kan zich echter ook afvragen wat er met de fout in een vooraf gegeven punt T van de t-as gebeurt wanneer we de integratiestappen steeds kleiner kiezen. Stabiliteitsbeschouwingen die de ontwikkeling van de fout in een vast punt T bij verfijning van de integratiestappen betreffen worden *uniforme* stabiliteitsbeschouwingen genoemd. In zo'n uniforme stabiliteitsanalyse wil men voorwaarden vinden die garanderen dat de grotere nauwkeurigheid tengevolge van kleinere integratiestappen niet weer verloren gaat door accumulatie van bijvoorbeeld afrondfouten. Op het eerste gezicht zou men denken dat de uitspraken in bovenstaande stelling zonder meer van toepassing zijn op de fouten-opbouw in een vooraf gegeven punt T. Dit is echter niet het geval. Het aantal stappen V tussen het punt t_n en T neemt namelijk op een bepaalde manier toe. Stel, bijvoorbeeld, dat de stappen $\{\tau_{n+j}\}_1^V$ steeds gehalveerd worden, dan neemt V toe als $O(\tau^{-1})$ waarin τ de grootste stap in de rij $\{\tau_{n+j}\}_1^V$ voorstelt. Door hiervan gebruik te maken kunnen veel scherpere uitspraken gedaan worden.

STELLING 2.3.

(a) *Bij verfijning blijft de fout uniform begrensd wanneer*

$$\sigma(A_n) \leq 1 + O(\tau) \quad \text{als } \tau \rightarrow 0,$$

$$\mu(A_n) \geq q(A_n) = 1.$$

(b) *Voor lineaire stelsels met constante integratiestappen gedraagt de fout zich bij verfijning tenminste als $\tau^{1-q(A)}$ wanneer*

$$\sigma(A) \leq 1 + O(\tau) \quad \text{als } \tau \rightarrow 0.$$

(c) Bij verfijning blijft de fout uniform begrensd wanneer

$$\|A_n\| \leq 1 + O(\tau) \quad \text{als } \tau \rightarrow 0.$$

BEWIJS.

(a) Kiezen we in (2.22) voor $\bar{A}_n$ de matrix A_n , dan geldt

$$\begin{aligned} L &= \max_{0 \leq j \leq v} \|C_j\|_2 = \max_{0 \leq j \leq v} \|A_{n+j} - A_n\|_2 = \\ &= \max_{0 \leq j \leq v} \|\tau_{n+j} J_{n+j} - \tau_n J_n\|_2 \leq \tau \max_{0 \leq j \leq v} \|J_{n+j}\|_2 = O(\tau) \end{aligned}$$

als $\tau \rightarrow 0$.

Hieruit, en uit de voorwaarden van de stelling vinden volgens (2.22)

$$\|\vec{S}_{n+v+1} - \vec{\tilde{S}}_{n+v+1}\|_2 \leq \text{const.} [1 + O(\tau)]^{\frac{1}{\tau}} \|\vec{S}_n - \vec{\tilde{S}}_n\|_2 = O(\|\vec{S}_n - \vec{\tilde{S}}_n\|)$$

als $\tau \rightarrow 0$.

(b) Substitutie van $v = O(\tau^{-1})$ en $\sigma(A) \leq 1 + O(\tau)$ in (2.22') geeft

$$\begin{aligned} \|\vec{S}_{n+v+1} - \vec{\tilde{S}}_{n+v+1}\|_2 &\leq \text{const.} \tau^{1-q(A)} [1 + O(\tau)]^{\frac{1}{\tau}} \|\vec{S}_n - \vec{\tilde{S}}_n\| = \\ &= O(\tau^{1-q(A)}) \|\vec{S}_n - \vec{\tilde{S}}_n\|_2 \end{aligned}$$

als $\tau \rightarrow 0$.

Aangezien de orde constante in vele gevallen naar nul gaat voor $\tau \rightarrow 0$, volgt uit deze relatie dat de fout zich niet ongunstiger ontwikkelt dan $\tau^{1-q(A)}$ als $\tau \rightarrow 0$.

(c) Uit (2.20) leiden we af

$$\begin{aligned} \|\vec{S}_{n+v+1} - \vec{\tilde{S}}_{n+v+1}\| &\leq \prod_{j=v}^0 \|A_{n+j}\| \|\vec{S}_n - \vec{\tilde{S}}_n\| \leq (1+O(\tau))^{v+1} \|\vec{S}_n - \vec{\tilde{S}}_n\| \leq \\ &\leq \text{const.} (1+O(\tau))^{\frac{1}{\tau}} \|\vec{S}_n - \vec{\tilde{S}}_n\| = O(\|\vec{S}_n - \vec{\tilde{S}}_n\|). \end{aligned}$$

Vergelijkingen we de uitspraken in stelling 2.2 en 2.3, dan valt direct op dat voor stabiliteit bij verfijning minder zware eisen aan de spectraalradius van de amplificatiematrix gesteld worden, maar dat de fout "slechts"

uniform begrensd blijft, terwijl deze bij niet-uniforme stabiliteitsbeschouwingen steeds kleiner wordt voor voldoende kleine spectrale radius.

Aan de hand van het gedrag van een op een bepaalde tijd geïntroduceerde fout worden er in de literatuur verschillende stabiliteitsdefinities gegeven. We noemen hier de belangrijkste:

Stabiliteit in de zin van Forsythe and Wasow: fout gedraagt zich als een of andere "lage" macht van τ^{-1} als $\tau \rightarrow 0$ (vergelijk stelling 2.3(b)).

Stabiliteit in de zin van Rjabenki en Fillippow: fout is uniform begrensd als $\tau \rightarrow 0$ (vergelijk stelling 2.3(a) en (c)).

Stabiliteit in de zin van O'Brien, Hyman en Kaplan: fout dempt uit als het aantal stappen toeneemt en $\tau_{n+j} \geq \text{constante} > 0$ (vergelijk stelling 2.2(a)).

2.5. Het Routh-Hurwitz-criterium

In de voorgaande paragrafen hebben we gezien dat het afleiden van stabiliteitsvoorwaarden steeds weer eindigt met de eis dat het spectrum van de amplificatiematrix binnen of eventueel op de eenheidscirkel ligt. Dit betekent dat we voor het probleem staan voorwaarden te vinden waar onder de wortels van een polynoom (karakteristiek polynoom van de amplificatiematrix) binnen of op de eenheidscirkel liggen. Dit is een klassiek probleem en kan opgelost worden met behulp van het lemma van Routh en Hurwitz.

LEMMA 2.2. *Zij gegeven het polynoom*

$$P_r(w) = a_0 w^r + a_1 w^{r-1} + \dots + a_r$$

en de $(n \times n)$ -matrices (de zogenaamde Hurwitz-matrices)

$$Q_n = \begin{pmatrix} a_1 & a_3 & a_5 & \dots & a_{2n-1} \\ a_0 & a_2 & a_4 & \dots & a_{2n-2} \\ 0 & a_1 & a_3 & \dots & a_{2n-3} \\ 0 & a_0 & a_2 & \dots & a_{2n-4} \\ & & \dots & & \\ 0 & 0 & \dots & & a_n \end{pmatrix}.$$

De wortels van P liggen allen in het negatieve halfvlak dan en alleen dan wanneer

$$a_0 > 0, \det Q_n > 0, n = 1, 2, \dots, r.$$

BEWIJS.

Zie bijvoorbeeld LANCASTER [1969].

Stel nu dat de wortels van het karakteristieke polynoom

$$(2.24) \quad C_r(z) = z^r + c_1 z^{r-1} + \dots + c_r$$

binnen een cirkel liggen met middelpunt in $z = 0$ en straal ρ , dan liggen de wortels van het polynoom

$$(1-w)^r C_r\left(\rho \frac{1+w}{1-w}\right)$$

allen in het negatieve halfvlak, immers de transformatie

$$z = \rho \frac{1+w}{1-w}$$

beeldt het gebied $|z| < \rho$ in het z -vlak af op het negatieve halfvlak van het w -vlak. Passen we vervolgens lemma 2.2 toe op het polynoom

$$(2.25) \quad P_r(w) = (1-w)^r C_r\left(\rho \frac{1+w}{1-w}\right),$$

dan vinden we de voorwaarden waaraan de coëfficiënten b_j van C_r voldoen. Omgekeerd garanderen deze voorwaarden dat de wortels van C_r binnen een cirkel met straal ρ liggen. We zullen deze voorwaarden nu afleiden voor $r = 2$ en $r = 3$.

$r = 2$

Voor $r = 2$ wordt het polynoom P_r volgens (3.25)

$$\begin{aligned} P_2(w) &\equiv (1-w)^2 \left[\left(\rho \frac{1+w}{1-w} \right)^2 + c_1 \rho \frac{1+w}{1-w} + c_2 \right] = \\ &= (\rho^2 - c_1 \rho + c_2) w^2 + 2(\rho^2 - c_2) w + (\rho^2 + c_1 \rho + c_2). \end{aligned}$$

Hieruit volgt direct

$$a_0 = \rho^2 - c_1\rho + c_2, \det Q_1 = 2(\rho^2 - c_2), \det Q_2 = 2(\rho^2 - c_2)(\rho^2 + c_1 + c_2).$$

Toepassing van dit lemma met $z = 2$ geeft de voorwaarden

$$(2.26) \quad \begin{aligned} |c_1| &< \frac{c_2}{\rho} + \rho \\ c_2 &< \rho^2 \end{aligned}.$$

$r = 3$

Na enig rekenwerk vinden we uit (2.25)

$$\begin{aligned} P_3(w) \equiv & [\rho^2(\rho - c_1) + \rho c_2 - c_3]w^3 + [\rho^2(3\rho - c_1) - \rho c_2 + 3c_3]w^2 + \\ & + [\rho^2(3\rho + c_1) - \rho c_2 - 3c_3]w + [\rho^2(\rho + c_1) + \rho c_2 + c_3], \end{aligned}$$

waarmee de coëfficiënten $a_0, \dots, a_3$ bepaald zijn. Toepassing van het lemma met $r = 3$ geeft

$$a_0 > 0, \det Q_1 = a_1 > 0, \det Q_2 = a_1 a_2 - a_0 a_3 > 0,$$

$$\det Q_3 = a_3 \det Q_2 > 0,$$

ofwel

$$a_0 > 0, a_1 > 0, a_2 > 0, a_3 > 0, a_1 a_2 - a_0 a_3 > 0.$$

Substitutie van de coëfficiënten a_j leidt tot

$$(2.27) \quad \begin{aligned} |\rho^2 c_1 + c_3| &< \rho(c_2 + \rho^2) \\ -(\rho^2 c_1 - 3c_3) &> \rho(c_2 - 3\rho^2) \\ c_3(\rho^2 c_1 - c_3) &> \rho^4(c_2 - \rho^2) \end{aligned}.$$

2.6. De constante coëfficiënten methode

In paragraaf 2.3 hebben we de amplificatie-matrix van de bewegingsvergelijkingen (zonder vervalkrachtterm) beschouwd en een bovengrens voor de

spectrale norm afgeleid. We willen nu echter met behulp van het Routh-Hurwitz-criterium iets over het spectrum van de amplificatie-matrix A_n zeggen. Dan worden we echter geconfronteerd met het feit dat de matrix-elementen van A_n van x en y afhangen omdat deze de stromingscomponenten U en V bevatten. In het algemeen zal de amplificatiematrix bij partiele differentiaalvergelijkingen van de plaatscoördinaatvector $\vec{x}$ afhangen, of, wanneer de $\vec{x}$ -ruimte al gediscretiseerd is tot een rooster met roosterpunten $\vec{x}_j$, zullen de elementen van A_n functies van $\vec{x}_j$ zijn. In het laatste geval kunnen we de elementen d_{kl} van A_n door diagonaal-matrices vervangen waarbij de diagonaal-elementen de achtereenvolgende waarden $d_{kl}(\vec{x}_j)$, $j = 1, 2, \dots, r$ hebben (r is hier het aantal gebruikte roosterpunten). In beide gevallen zal men in de praktijk op grote moeilijkheden stuiten bij een analyse van het spectrum. Daarom stelt men zich veelal tevreden met een analyse van de zogenaamde *locale stabiliteit*. Hierbij denkt men zich de operator A_n , gedefinieerd in een zeker punt $\vec{x}$, toegepast in *alle* punten $\vec{x}$ van het beschouwde $\vec{x}$ -gebied en stelt de stabiliteitsvoorwaarden voor deze operator met constante coëfficiënten op. We spreken dan van *locale stabiliteit in het punt $\vec{x}$* . Vervolgens eist men dat A_n in elk punt $\vec{x}$ lokaal stabiel is. In de literatuur staat deze benadering van het stabiliteitsprobleem bekend als de *constante coëfficiëntenmethode* (zie RICHTMYER en MORTON [1967]).

Als rechtvaardiging van het uitgangspunt dat lokale stabiliteit in alle relevante punten $\vec{x}$, de "overall" stabiliteit impliceert, voert men wel aan dat uit praktische ervaring blijkt dat instabiliteit altijd lokaal begint en dan zich in de $\vec{x}$ -ruimte uitbreidt. In RICHTMYER en MORTON [1967] wordt voor enkele belangrijke klassen van differentiaalvergelijkingen aangetoond dat een lichte verzwaring van de locale stabiliteitsvoorwaarden inderdaad overall-stabiliteit garandeert.

In deze syllabus zullen we de lokale amplificatie-matrix behorend bij een punt $\vec{x}$ met $A_n(\vec{x})$ aangeven en diens eigenwaarden met $\alpha_n(\vec{x})$ of kortweg α .

2.7. Stabiliteit t.o.v. bodemwrijving, Corioliskracht en convectieve traagheidstermen (vervolg)

Laten we nog eens de algoritme (2.15) beschouwen. De amplificatie-matrix schrijven we in de vorm

$$A_n = \begin{pmatrix} 1 - \tau_n \lambda_x & \tau_n \omega_x \\ -\tau_n \omega_y & 1 - \tau_n \lambda_y \end{pmatrix}$$

waarin we eenvoudigheidshalve

$$(2.28) \quad \begin{aligned} \lambda_x &= \lambda + \frac{\partial U_n}{\partial x}, & \lambda_y &= \lambda + \frac{\partial V_n}{\partial y}, \\ \omega_x &= \omega - \frac{\partial U_n}{\partial y}, & \omega_y &= \omega + \frac{\partial V_n}{\partial x}, \end{aligned}$$

gesteld hebben. De eigenwaarden α van de locale amplificatie-matrix $A_n(\vec{x})$ voldoen aan de vergelijking

$$C_2(\alpha) \equiv \alpha^2 - (2 - \tau_n(\lambda_x + \lambda_y))\alpha + 1 - \tau_n(\lambda_x + \lambda_y) + \tau_n^2(\lambda_x \lambda_y + \omega_x \omega_y) = 0.$$

De wortels van deze vergelijking liggen volgens criterium (3.26) binnen de cirkel $|z| = \rho$ wanneer

$$\begin{aligned} |2 - \tau_n(\lambda_x + \lambda_y)| &< \frac{1 + \rho^2 - \tau_n(\lambda_x + \lambda_y) + \tau_n^2(\lambda_x \lambda_y + \omega_x \omega_y)}{\rho}, \\ 1 - \tau_n(\lambda_x + \lambda_y) + \tau_n^2(\lambda_x \lambda_y + \omega_x \omega_y) &< \rho^2. \end{aligned}$$

Uitwerking van deze ongelijkheden leidt tot de voorwaarden

$$(2.29) \quad \lambda_x + \lambda_y > 0, \quad \lambda_x \lambda_y + \omega_x \omega_y > 0$$

en voor τ_n tot de voorwaarden

$$(2.30) \quad \begin{aligned} \tau_n &< \frac{\lambda_x + \lambda_y - \frac{1 - \rho^2}{\tau_n}}{\lambda_x \lambda_y + \omega_x \omega_y}, \\ \tau_n &< \frac{(1 + \rho)^2 + (\lambda_x \lambda_y + \omega_x \omega_y) \tau_n^2}{(1 + \rho)(\lambda_x + \lambda_y)}, \\ \tau_n &> \sqrt{\frac{1 - \rho^2}{\lambda_x \lambda_y + \omega_x \omega_y}}. \end{aligned}$$

Voorwaarden (2.29) kunnen niet beïnvloed worden door de integratiestap kleiner te kiezen of ρ dichter bij 1 te nemen. Deze voorwaarden worden door het wiskundig model zelf bepaald en zijn juist de stabiliteitsvoorwaarden van de bewegingsvergelijkingen (1.2'). Dit is als volgt in te zien: schrijf (1.2') in de vorm

$$\frac{\partial}{\partial t} \vec{u} = \begin{pmatrix} -\lambda_x & \omega_x \\ -\omega_y & -\lambda_y \end{pmatrix} \vec{u}.$$

Het verschil van twee oplossingen $\vec{u}$ en $\tilde{u}$ voldoet aan de vergelijking

$$(2.31) \quad \frac{\partial}{\partial t} (\vec{u} - \tilde{u}) = \begin{pmatrix} -\lambda_x & \omega_x \\ -\omega_y & -\lambda_y \end{pmatrix} (\vec{u} - \tilde{u}) + \begin{pmatrix} \tilde{\lambda}_x - \lambda_x & \omega_x - \tilde{\omega}_x \\ \tilde{\omega}_y - \omega_y & \tilde{\lambda}_y - \lambda_y \end{pmatrix} \tilde{u}.$$

Evenals in de amplificatiematrix van het Eulerschema (2.15), verwaarlozen we verstoringen in λ_x , λ_y , ω_x en ω_y , zodat de tweede term in (2.31) vervalt. De voorwaarden voor convergentie van $\vec{u}$ en $\tilde{u}$ als t toeneemt zijn nu dat de Jacobiaan van de variatievergelijking zijn spectrum in het negatieve halfvlak heeft, dus de wortels van de vergelijking

$$\omega^2 + (\lambda_x + \lambda_y)\omega + \lambda_x \lambda_y + \omega_x \omega_y = 0$$

moeten in het negatieve halfvlak liggen. Volgens lemma 2.2 moet dan voldaan zijn aan

$$\det Q_1 = \lambda_x + \lambda_y > 0, \quad \det Q_2 = \lambda_x \lambda_y + \omega_x + \omega_y > 0,$$

dus juist aan de voorwaarden (2.29).

Voorwaarden (2.30) geven een bovengrens voor de integratiestap; het is eenvoudig in te zien dat deze bovengrens altijd positief gemaakt kan worden door ρ voldoende dicht bij 1 te kiezen, voorop gesteld natuurlijk dat aan de "intrinsieke" stabiliteitsvoorwaarden (2.29) voldaan is. Om enig idee van deze bovengrens te krijgen werken we (2.30) uit voor het geval waarin de convectieve traagheidstermen verwaarloosd zijn, dus

$$\lambda_x = \lambda_y = \lambda, \quad \omega_x = \omega_y = \omega.$$

We vinden dan (merk op dat aan (2.29) automatisch voldaan is als $\lambda > 0$)

$$(2.30') \quad \tau_n < \frac{2\lambda - \frac{1-\rho^2}{2}}{\lambda^2 + \omega^2}, \quad \tau_n < \frac{(1+\rho)^2 + (\lambda^2 + \omega^2)\tau_n^2}{2(1+\rho)\lambda},$$

$$\tau_n > \sqrt{\frac{1-\rho^2}{\lambda^2 + \omega^2}}.$$

Voor de Noordzee zijn de gemiddelde waarden van λ en ω

$$\lambda \cong 25_{10}^{-6} \text{ sec}^{-1}, \quad \omega \cong 125_{10}^{-6} \text{ sec}^{-1}.$$

Substitutie leidt na enige herleiding tot (alleen de eerste en derde voorwaarde zijn relevant)

$$(2.30'') \quad 8000\sqrt{1-\rho^2} < \tau_n < 3000 - 6_{10}^8 \frac{1-\rho^2}{\tau_n}.$$

In figuur 2.2 is het gebied aangegeven van stabiele (ρ, τ_n) -punten. Hieruit zien we dat de maximale marge voor stabiliteit, dat wil zeggen de kleinste

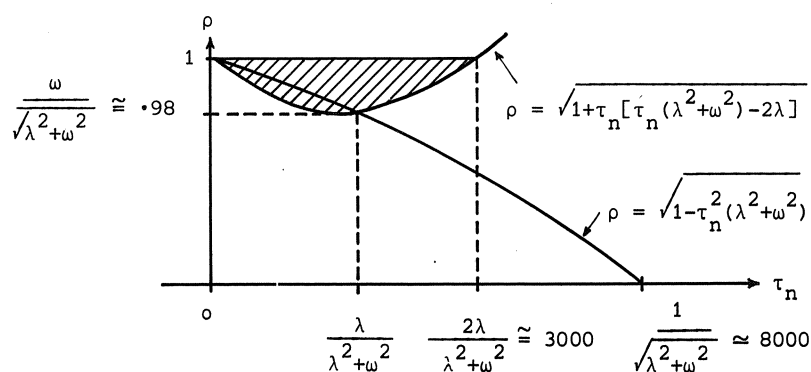


Fig. 2.2. Stabiele (ρ, τ_n) -punten

waarde voor de spectraal radius van de amplificatie matrix gegeven wordt door

$$\min_{\tau_n} \sigma(A_n) = \min_{0 \leq \tau_n \leq 3000} \sqrt{1 + \frac{\tau_n (\tau_n - 3000)}{6_{10}^8}} \cong .98;$$

deze waarde wordt bereikt voor een integratiestap van 1500 sec. Deze marge is niet erg groot en de al eerder gemaakte opmerking dat het Eulerschema niet geschikt is voor de integratie van de ondiep-watervergelijkingen wordt hier bevestigd.

2.8. Stabiliteit voor het volledige stelsel ondiep-watervergelijkingen

In deze paragraaf analyseren we de stabiliteit van het stelsel (1.8) en het Euler-schema (2.4). We zullen daarbij niet de splitsing in lineaire termen ($\mathcal{D}^{(1)}$ s) en niet-lineaire termen ($\mathcal{D}^{(2)}$ s) hanteren zoals bij analytische oplossingsmethoden wenselijk is. We schrijven (1.8) daarentegen in de vorm

$$(2.32) \quad \frac{\partial}{\partial t} \vec{s} = \mathcal{D} \vec{s} + \vec{f},$$

$$\mathcal{D} = - \begin{pmatrix} \lambda_x & -\omega_x & g \frac{\partial}{\partial x} \\ \omega_y & \lambda_y & g \frac{\partial}{\partial y} \\ h_x + h \frac{\partial}{\partial x} & h_y + h \frac{\partial}{\partial y} & 0 \end{pmatrix}, \quad h_x = \frac{\partial h}{\partial x}, \quad h_y = \frac{\partial h}{\partial y},$$

waarin de parameters λ_x , λ_y , ω_x en ω_y overeenkomstig (2.28) gedefinieerd zijn.

We gaan eerst de stabiliteit van (2.32) na. Evenals in de voorgaande paragrafen verwaarlozen we verstoringen van de matrix $\mathcal{D}$, zodat voor twee naburige oplossingen $\vec{s}$ en $\vec{\tilde{s}}$ geldt

$$(2.33) \quad \frac{\partial}{\partial t} (\vec{s} - \vec{\tilde{s}}) = \mathcal{D}(\vec{s} - \vec{\tilde{s}}).$$

We beperken ons weer tot lokale stabiliteitsbeschouwingen en gaan na wanneer de eigenwaarden van de operator $\mathcal{D}(\vec{x})$ met coëfficiënten gelijk aan die in het punt $\vec{x}$, in het negatieve halfvlak liggen. Er treedt nu echter een nieuwe moeilijkheid op: omdat $\mathcal{D}(\vec{x})$ de differentiaaloperatoren $\partial/\partial x$ en $\partial/\partial y$ bevat, kan het Routh-Hurwitz-criterium niet zonder meer toegepast worden. Von Neumann heeft in verband met soortgelijke problemen voorgesteld de verstoringen $\vec{s} - \vec{\tilde{s}}$ in een Fourierreeks te ontwikkelen en de amplificatie-matrices voor de componenten van deze Fourierreeks op te stellen. De waarde van deze methode ligt hierin dat in de meeste gevallen de eigenfuncties van de amplificatie-operator van de partiele differentiaalvergelijking van de vorm

$$(2.34) \quad \vec{a}(\vec{\gamma}) \exp[i(\vec{\gamma}, \vec{x})], \quad \vec{\gamma} \text{ reëel}$$

zijn, waarin $\vec{a}$ niet van $\vec{x}$ afhangt en $\vec{\gamma}$ de "index" van de eigenfunctie voorstelt. Laten we de eigenfuncties van de operator $\mathcal{D}(\vec{x})$ in de vorm (2.34) zoeken. Aangezien de componenten $i\gamma_1$ en $i\gamma_2$ van $i\vec{\gamma}$ juist de eigenwaarden

van respectievelijk $\partial/\partial x$ en $\partial/\partial y$ behorende bij de functie (2.34) zijn, geldt de relatie

$$(2.35) \quad \mathcal{D}(\vec{x}) \vec{a} e^{i(\vec{\gamma}, \vec{x})} = \tilde{\mathcal{D}}(\vec{x}, \vec{\gamma}) \vec{a} e^{i(\vec{\gamma}, \vec{x})},$$

waarin $\tilde{\mathcal{D}}_n(\vec{x}, \vec{\gamma})$ de amplificatiematrix voor de functies (2.34) is en gedefinieerd wordt door

$$(2.36) \quad \tilde{\mathcal{D}}(\vec{x}, \vec{\gamma}) = - \begin{pmatrix} \lambda_x(\vec{x}) & -\omega_x(\vec{x}) & i\gamma_1 \\ \omega_y(\vec{x}) & \lambda_y(\vec{x}) & i\gamma_2 \\ h_x(\vec{x}) + ih(\vec{x})\gamma_1 & h_y(\vec{x}) + ih(\vec{x})\gamma_2 & 0 \end{pmatrix}.$$

De matrix $\tilde{\mathcal{D}}(\vec{x}, \vec{\gamma})$ heeft scalaire matrix-elementen; kiest men voor $\vec{a}$ een eigenvector van $\tilde{\mathcal{D}}$ dan zijn de eigenfuncties van de operator $\mathcal{D}(\vec{x})$ inderdaad van de vorm (2.34) en zijn de eigenwaarden δ van $\mathcal{D}(\vec{x})$ juist die van $\tilde{\mathcal{D}}(\vec{x}, \vec{\gamma})$, dus de wortels van de karakteristieke vergelijking

$$(2.37) \quad \det[\tilde{\mathcal{D}}(\vec{x}, \vec{\gamma}) - \delta I] = 0.$$

Enig rekenwerk levert de derdegraads vergelijking (het argument $\vec{x}$ laten we verder achterwege)

$$(2.37') \quad \begin{aligned} \delta^3 + a_1 \delta^2 + a_2 \delta + a_3 &= 0, \\ a_1 &= \lambda_x + \lambda_y, \\ a_2 &= \lambda_x \lambda_y + \omega_x \omega_y + gh \left[\gamma_1^2 + \gamma_2^2 - i \left(\frac{h_x}{h} \gamma_1 + \frac{h_y}{h} \gamma_2 \right) \right], \\ a_3 &= gh \left[\lambda_x \gamma_2^2 + \lambda_y \gamma_1^2 + \gamma_1 \gamma_2 (\omega_x - \omega_y) - i \left(\frac{\lambda_x h_y + \omega_x h_x}{h} \gamma_2 + \frac{\lambda_y h_x - \omega_y h_y}{h} \gamma_1 \right) \right]. \end{aligned}$$

Omdat γ_1 en γ_2 reëel verondersteld zijn, zijn de coëfficiënten a_2 en a_3 complex-waardig zodat het Routh-Hurwitz-criterium niet toegepast kan worden. Een nadere beschouwing leert echter dat de imaginaire gedeelten in a_2 en a_3 een factor van de orde $(|h_x| + |h_y|)/h$ verschillen van de reële gedeelten. Dus in punten $\vec{x}$ waar

$$(2.38) \quad \frac{|h_x| + |h_y|}{h} \ll 1$$

zijn de coëfficiënten a_2 en a_3 nagenoeg reëel; het Routh-Hurwitz-criterium levert dan de stabiliteitsvoorwaarden

$$\begin{aligned}
 & \lambda_x + \lambda_y > 0, \\
 (2.39) \quad & \lambda_x \gamma_2^2 + (\omega_x - \omega_y) \gamma_1 \gamma_2 + \lambda_y \gamma_1^2 > 0, \\
 & (\lambda_x + \lambda_y) (\lambda_x \lambda_y + \omega_x \omega_y) > -gh (\lambda_x \gamma_1^2 - (\omega_x - \omega_y) \gamma_1 \gamma_2 + \lambda_y \gamma_2^2).
 \end{aligned}$$

Aangezien γ_1 en γ_2 willekeurige reële getallen zijn is aan deze ongelijkheden alleen dan voldaan wanneer $\gamma_1 \neq 0$ of $\gamma_2 \neq 0$ en

$$\begin{aligned}
 (2.39') \quad & \lambda_x > 0, \quad \lambda_y > 0, \\
 & \lambda_x \lambda_y + \omega_x \omega_y > \frac{1}{4} (\omega_x + \omega_y)^2.
 \end{aligned}$$

Indien γ_1 en γ_2 gelijktijdig nul zijn, kan aan (2.39) niet voldaan worden zodat $\mathcal{D}(\vec{x})$ minstens één eigenwaarde gelijk nul heeft.

Merk op dat aan (2.39') automatisch voldaan is wanneer de convectieve traagheidstermen in het wiskundig model weggelaten worden ($\lambda_x = \lambda_y = \lambda$ en $\omega_x = \omega_y = \omega$).

De volgende stap is de stabiliteit van het Euler-schema

$$(2.40) \quad \vec{S}_{n+1} = [I + \tau_n \mathcal{D}_n] \vec{S}_n + \tau_n \vec{F}_n.$$

We beschouwen weer de lokale amplificatie-operatoren; voor het Euler-schema worden deze gegeven door

$$(2.41) \quad A_n(\vec{x}) = I + \tau_n \mathcal{D}_n(\vec{x}).$$

Toepassing op de functies (2.34) geeft de bijbehorende amplificatiematrices

$$(2.41') \quad \tilde{A}_n(\vec{x}, \vec{\gamma}) = I + \tau_n \tilde{\mathcal{D}}_n(\vec{x}, \vec{\gamma}),$$

waarin $\tilde{\mathcal{D}}_n$ door (2.36) gedefinieerd wordt. We zullen nu nagaan wanneer

$$(2.42) \quad \sigma(\tilde{A}_n(\vec{x}, \vec{\gamma})) \leq \rho < 1.$$

Daartoe moeten we aan de ongelijkheden (2.27) voldoen, waarin de coëfficiënten

c_1, c_2 en c_3 nu de coëfficiënten van de karakteristieke vergelijking van $\tilde{A}_n(\vec{x}, \vec{\gamma})$ voorstellen. Eenvoudig rekenwerk geeft

$$(2.43) \quad \det[(1-\alpha)I + \tau_n \tilde{D}_n(\vec{x}, \vec{\gamma})] \equiv \alpha^3 + c_1 \alpha^2 + c_2 \alpha + c_3 = 0,$$

waarin

$$\begin{aligned} c_1 &= -3 + \tau_n (\lambda_x + \lambda_y), \\ c_2 &= 3 - 2\tau_n (\lambda_x + \lambda_y) + \tau_n^2 [(\lambda_x \lambda_y + \omega_x \omega_y) + \text{gh}(\gamma_1^2 + \gamma_2^2)], \\ c_3 &= -1 + \tau_n (\lambda_x + \lambda_y) - \tau_n^2 [\lambda_x \lambda_y + \omega_x \omega_y] + \text{gh} \tau_n^2 [(\lambda_y - 1) \gamma_1^2 + (\lambda_x - 1) \gamma_2^2 + \\ &\quad + \gamma_1 \gamma_2 (\omega_x - \omega_y)]. \end{aligned}$$

Hierin zijn de argumenten $\vec{x}$ weggelaten en is voorwaarde (2.38) al in rekening gebracht. Substitutie van deze uitdrukkingen in de voorwaarden (2.27) levert de uiteindelijke stabiliteitsvoorwaarden voor de methode van Euler. Het speciale geval $\lambda_x = \lambda_y = \lambda$, $\omega_x = \omega_y = \omega$ en $\rho = 1$ is onderzocht door LAUWERIER en DAMSTÉ [1961]; in de terminologie van deze syllabus zouden hun resultaten leiden tot

$$(2.44) \quad \tau_n < \min_{\vec{\gamma}} \frac{\lambda}{\text{gh}(\gamma_1^2 + \gamma_2^2)}, \quad \tau_n < \min \left\{ \frac{2}{3\lambda}, \frac{2\lambda}{\lambda^2 + \omega^2} \right\}.$$

Bij de interpretatie van deze voorwaarden moeten we wel bedenken dat in het Euler-schema (2.40) de differentiaaloperatoren $\partial/\partial x$ en $\partial/\partial y$ in de matrix nog niet door differentiequotienten vervangen zijn. We zien dus dat het Euler-schema voor de *partiele* differentiaalvergelijking (2.32) alleen dan stabiel is wanneer in de eerste plaats $\lambda > 0$ is en ten tweede wanneer men zich beperkt tot verstoringen die uit de "lagere harmonischen" zijn opgebouwd, dat wil zeggen

$$(2.45) \quad \vec{s} - \tilde{\vec{s}} = \sum_{\gamma_1^2 + \gamma_2^2 < \Gamma} \vec{a}(\gamma) e^{i(\vec{\gamma}, \vec{x})},$$

waarin Γ een of ander reëel getal is. In de praktijk worden de differentiaaloperatoren uiteraard door differentieoperatoren vervangen en moeten overeenkomstig γ_1 en γ_2 door de eigenwaarden van deze differentie-operatoren vervangen worden. Laten we de discretisering van $\partial/\partial x$ en $\partial/\partial y$ met $[\partial/\partial x]$ en $[\partial/\partial y]$ aangeven; beschouw in het bijzonder differentie-operatoren gedefinieerd op een uniform, rechthoekig rooster. Het eenvoudigste geval ontstaat als

$$(2.46) \quad \left[\frac{\partial}{\partial x} \right] \vec{S}_n(x, y) = \frac{\vec{S}_n(x + \xi, y) - \vec{S}_n(x - \xi, y)}{2\xi},$$

en een analoge uitdrukking voor $[\partial/\partial y]$. Hierin stelt ξ de afstand van de roosterpunten in de x -richting voor (zie figuur 2.3). Passen we deze centrale differentie-operator toe op de functies (2.34) dan vinden we

$$\begin{aligned} \left[\frac{\partial}{\partial x} \right]_{ae}^{\vec{\gamma}} i(\vec{\gamma}, \vec{x}) &= \frac{\vec{a}}{2\xi} e^{i(\gamma_1(x+\xi)+\gamma_2 y)} - e^{i(\gamma_1(x-\xi)+\gamma_2 y)} \\ &= \frac{i \sin \gamma_1 \xi}{\xi} \vec{a} e^{i(\vec{\gamma}, \vec{x})}, \\ \left[\frac{\partial}{\partial y} \right]_{ae}^{\vec{\gamma}} i(\vec{\gamma}, \vec{x}) &= \frac{i \sin \gamma_2 \eta}{\eta} \vec{a} e^{i(\vec{\gamma}, \vec{x})}. \end{aligned}$$

De differentie-operatoren $[\partial/\partial x]$ en $[\partial/\partial y]$ hebben dus de functies (2.34) als eigenfuncties met eigenwaarden δ_x en δ_y waarin

$$(2.47') \quad \delta_x = i \frac{\sin \gamma_1 \xi}{\xi}, \quad \delta_y = i \frac{\sin \gamma_2 \eta}{\eta}.$$

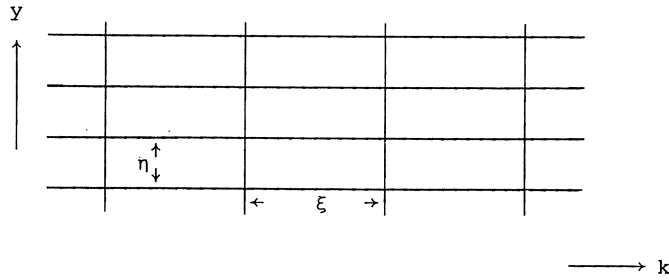


Fig. 2.3. Rooster in het (x, y) -vlak

Vervangen we in (2.36) $i\gamma_1$ en $i\gamma_2$ door δ_x en δ_y , en wijzigen we alle volgende formules op overeenkomstige wijze dan gaan de voorwaarden (2.44) over in

$$(2.44') \quad \tau_n < \min_{\vec{\gamma}} \frac{\lambda}{\text{gh}\left(\frac{\sin^2 \gamma_1 \xi}{\xi^2} + \frac{\sin^2 \gamma_2 \eta}{\eta^2}\right)} = \frac{\lambda \xi^2 \eta^2}{\text{gh}(\xi^2 + \eta^2)},$$

$$\tau_n < \min \left\{ \frac{2}{3\lambda}, \frac{2\lambda}{\lambda^2 + \omega^2} \right\}.$$

In tegenstelling tot de voorwaarden (2.44) behorend bij de "ongediscretiseerde" differentiaaloperatoren $\partial/\partial x$ en $\partial/\partial y$ krijgen we nu automatisch een van nul verschillende stabiele waarde voor τ_n zonder de verstoringen te hoeven beperken in de zin van (2.45). Dit hangt samen met het feit dat hogere harmonischen op een rooster met roosterparameters ξ en η niet meer voorstelbaar zijn, zodat verstoringen automatisch uit een eindig aantal componenten $\vec{a} \exp[i(\vec{\gamma}, \vec{x})]$ opgebouwd zijn. In figuur 2.4 zijn de nog juist voorstelbare harmonischen geschetst (ononderbroken lijn) en de net niet meer voorstelbare harmonische (onderbroken lijn). Door een fijner rooster te kiezen, dus de

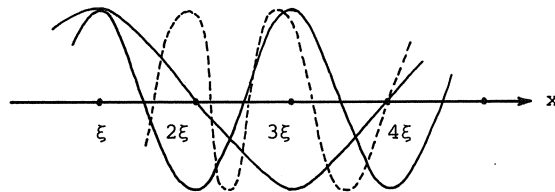


Fig. 2.4. Wel- en niet-voorstelbare harmonischen op een rooster met roosterafstand ξ

operatoren $\partial/\partial x$ en $\partial/\partial y$ nauwkeuriger te benaderen, kunnen steeds hogere harmonischen gerepresenteerd worden en moet de tijdstap τ_n steeds kleiner gekozen worden. Dit gedrag is geheel in overeenstemming met het limietgeval waarin het Eulerschema direct op de partiele differentiaalvergelijking toegepast wordt.

Tenslotte berekenen we de maximale stabiele integratiestap voor de Noordzee-waarden gegeven in de vorige paragraaf en een rooster met $\xi = \eta = 20$ km. Uit (2.44') volgt dan

$$(2.44) \quad \tau_n < \frac{25 \cdot 10^{-6} (210^4)^2}{20h} = \frac{500}{h} \text{ sec.}$$

Aangezien de Noordzee in het noordelijk gedeelte enkele honderden meters diep is, schrijft deze voorwaarde integratiestappen van enkele seconden voor. Voor daadwerkelijke berekeningen is dit een onaanvaardbare eis zodat we naar andere integratiemethoden moeten uitzien. In het bijzonder zullen we trachten algoritmen te construeren die aan de volgende kwalificaties voldoen:

- (1) zo mogelijk expliciet.
- (2) stabiel ook wanneer $\lambda = 0$
- (3) 2^e orde nauwkeurig in de tijd-variabele
- (4) vrij te kiezen roosterpunten.

III. GEGENERALISEERDE EULER-METHODEN

In dit hoofdstuk gaan we in op de eerste twee wensen die we aan het eind van het vorige hoofdstuk hebben geformuleerd. We zullen laten zien dat de stabiliteit zelfs voor $\lambda = 0$ aanzienlijk verbeterd kan worden door de Euler-methode zo te wijzigen dat de algoritme "impliciet" wordt. Hoe impliciet des te beter het stabiliteitsgedrag. Deze gemodificeerde Euler-methoden zullen we *gegeneraliseerde Eulermethoden* noemen en als volgt beschrijven:

$$(3.0) \quad [I - \tau_n E_n] \vec{S}_{n+1} = [I + \tau_n (D_n - E_n)] \vec{S}_n + \tau_n \vec{F}_n,$$

waarin E_n een operator is die "varieert" van 0 tot D_n . Voor $E_n = 0$ ontstaat het reeds beproven, volledig expliciete Euler-schema (zie schema (2.40)) en voor $E_n = D_n$ krijgen we een schema dat grote gelijkenis vertoont met de volledig impliciete, terugwaartse Euler-formule:

$$(3.1) \quad E_n = D_n: \quad [I - \tau_n D_n] \vec{S}_{n+1} = \vec{S}_n + \tau_n \vec{F}_n,$$

$$\text{terugwaarts Euler : } [I - \tau_n D_{n+1}] \vec{S}_{n+1} = \vec{S}_n + \tau_n \vec{F}_{n+1}.$$

Het belangrijkste verschil is dat formule (3.1) *lineair* impliciet is, dat wil zeggen om $\vec{S}_{n+1}$ te vinden moet een *lineair* stelsel opgelost worden, en dat de terugwaartse Euler-formule tot een *niet-lineair* stelsel aanleiding geeft.

3.1. Vergelijking van de stabiliteit van de meest expliciete en de meest impliciete vorm

Om de Euler-formula (3.1) te kunnen vergelijken drukken we de eigenwaarden van de respectieve amplificatiematrices uit in de eigenwaarden van de matrix D_n . De amplificatie-matrices worden gegeven door:

$$(3.2) \quad I + \tau_n \mathcal{D}_n \quad \text{en} \quad [I - \tau_n \mathcal{D}_n]^{-1},$$

waarin weer variaties in $\mathcal{D}_n$ verwaarloosd zijn. Het spectrum van deze matrices wordt gegeven door respectievelijk

$$(3.3) \quad 1 + \tau_n \delta \quad \text{en} \quad [1 - \tau_n \delta]^{-1}$$

waarin δ de eigenwaarden van $\mathcal{D}_n$ doorloopt. In de figuren 3.1 en 3.2 zijn in het $\tau_n \delta$ -vlak de gebieden aangegeven waar de eigenwaarden α van de amplificatie-matrices binnen een cirkel met straal ρ liggen. (Zogenaamde *stabiliteitsgebieden*).

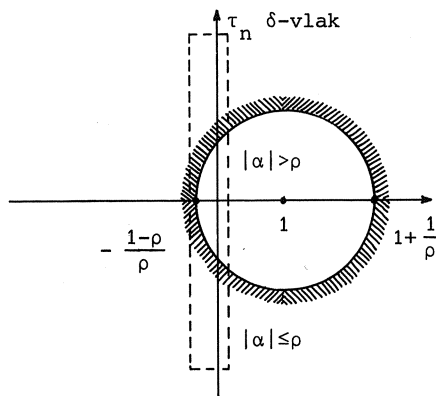


Fig. 3.1. Gebied van $\tau_n \delta$ -punten waarvoor $|\alpha| \leq \rho$ in geval van formule (3.1).

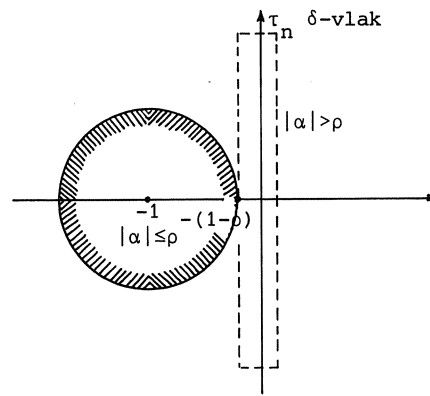


Fig. 3.2. Gebied van $\tau_n \delta$ -punten waarvoor $|\alpha| \leq \rho$ in geval van de Euler-formule.

Hieruit concluderen we dat formule (3.1) stabiel is wanneer de eigenwaarden δ *buiten* een cirkel met middelpunt in $(1/\tau_n, 0)$ en straal $1/(\rho\tau_n)$ liggen, terwijl de Euler-formule eist dat de eigenwaarden van $\mathcal{D}_n$ binnen een cirkel met middelpunt in $(-1/\tau_n, 0)$ en straal ρ liggen. Voor $\rho \leq 1$ omvat het stabiliteitsgebied van formule (3.1) volledig het stabiliteitsgebied van de Euler-formule.

Zoals in het vorige hoofdstuk uiteengezet is, gaat men tot lokale beschouwingen over wanneer men de eigenwaarden van de amplificatiematrix echt

wil uitrekenen. Wanneer we dus na willen gaan waar de eigenwaarden δ van $\mathcal{D}_n$ liggen om hieruit het stabiliteitsgedrag van het Euler-schema en formule (3.1) te kunnen vergelijken, dan gaan we over op de locale matrices $\mathcal{D}_n(\vec{x})$. De eigenwaarden van $\mathcal{D}_n(\vec{x})$ voldoen aan vergelijking (2.37') wanneer in de coëfficiënten a_j de toestandsvector S_n in het punt $\vec{x}$ gesubstitueerd wordt. Verder nemen we aan dat de operatoren $\partial/\partial x$ en $\partial/\partial y$ in $\mathcal{D}_n(\vec{x})$ door differentie-operatoren $[\partial/\partial x]$ en $[\partial/\partial y]$ met eigenwaarden δ_x en δ_y vervangen zijn. Dit betekent dat γ_1 en γ_2 in (2.37') door $-i\delta_x$ en $-i\delta_y$ vervangen dienen te worden. We zullen twee bijzondere gevallen beschouwen, het geval

$$(3.4) \quad \omega_x = \omega_y \cong 0, \quad \lambda_x = \lambda_y \cong \lambda, \quad \frac{|H_x| + |H_y|}{H} \ll 1$$

dus zonder Coriolis-kracht en convectieve traagheidstermen, en het geval

$$(3.5) \quad \omega_x = \omega_y \cong \omega, \quad \lambda_x = \lambda_y \cong 0, \quad \frac{|H_x| + |H_y|}{H} \ll 1$$

dus zonder bodemwrijving en convectieve traagheidstermen.

In het geval (3.4) vinden we

$$\delta_1 \cong -\lambda, \quad \delta_{2,3} \cong -\frac{1}{2} \lambda \pm \sqrt{\frac{1}{4} \lambda^2 + gH(\delta_x^2 + \delta_y^2)}$$

en in het geval (3.5)

$$\delta_1 \cong 0, \quad \delta_{2,3} \cong \pm \sqrt{-\omega^2 + gH(\delta_x^2 + \delta_y^2)}.$$

Aangezien zowel λ en ω klein verondersteld mogen worden, en δ_x en δ_y zuiver imaginair zijn vinden we in beide gevallen een spectrum van nagenoeg imaginaire eigenwaarden. Het lijkt redelijk om aan te nemen dat in een concreet geval het spectrum van $\mathcal{D}_n(\vec{x})$ in een langgerekte strook langs de imaginaire as zal liggen (in de figuren 3.1 en 3.2 met stippellijnen aangegeven). Hieruit volgt dat formule (3.1) voor ρ voldoende dicht bij 1 stabiel is ten opzichte van *alle* eigenwaarden δ in het negatieve halfvlak, terwijl het Euler-schema alleen maar ten opzichte van deze eigenwaarden stabiel gemaakt kan worden door én p dicht bij 1 én τ_n voldoende klein te kiezen. Formule (3.1) heeft dus zeker wat de grootte van de toegestane integratiestappen betreft

aantrekkelijke eigenschappen, alhoewel de damping van het schema, zeker voor $\lambda = 0$, nihil zal zijn (zwakke stabiliteit). Een voor de praktijk echter zeer onaantrekkelijke eigenschap is de volledige impliciteit van formule (3.1). In de volgende paragraaf illustreren we dit door formule (3.1) uit te schrijven met in acht neming van de randvoorwaarden.

3.2. Het volledige differentieschema voor de meest impliciete vorm

Om de numerieke bewerkelijkheid van schema (3.1) op waarde te kunnen schatten zullen we eerst het differentieschema zo ver mogelijk uitschrijven zonder echter ons al vast te leggen op een speciaal rooster of een bepaalde keuze van de differentie-operatoren $[\partial/\partial x]$ en $[\partial/\partial y]$. Voor een goede beschrijving moeten we echter wel de randvoorwaarden (1.19) en (1.20) in de beschouwingen betrekken. Daartoe gaan we na hoe de operator $\mathcal{D}$ gedefinieerd is langs de oceaan- respectievelijk de kustbegrenzing.

Langs de oceaan volgt de vorm van de operator direct uit (1.19), namelijk

$$(3.6) \quad \mathcal{D} = - \begin{pmatrix} \lambda_x & -\omega_x & g \frac{\partial}{\partial x} \\ \omega_y & \lambda_y & g \frac{\partial}{\partial y} \\ 0 & 0 & 0 \end{pmatrix}$$

Langs de kusten gaan de bewegingsvergelijkingen in (1.21") over, of in componenten uitgeschreven (c_x en c_y stellen de componenten van $\vec{c}$ voor)

$$\frac{\partial}{\partial t} u = - (\lambda + c_x (c_x \frac{\partial u}{\partial x} + c_y \frac{\partial v}{\partial x})) u - c_x (c_x \frac{\partial u}{\partial y} + c_y \frac{\partial v}{\partial y}) v +$$

$$-g c_x (c_x \frac{\partial}{\partial x} + c_y \frac{\partial}{\partial y}) h + c_x (c_x f_x + c_y f_y),$$

$$\frac{\partial}{\partial t} v = -c_y (c_x \frac{\partial u}{\partial x} + c_y \frac{\partial v}{\partial x}) u - (\lambda + c_y (c_x \frac{\partial u}{\partial y} + c_y \frac{\partial v}{\partial y})) v +$$

$$-g c_y (c_x \frac{\partial}{\partial x} + c_y \frac{\partial}{\partial y}) h + c_y (c_x f_x + c_y f_y).$$

Definiëren we langs de kusten λ_x , λ_y , ω_x en ω_y volgens (vergelijk (2.28))

$$\begin{aligned}
 \lambda_x &= \lambda + c_x(c_x \frac{\partial u}{\partial x} + c_y \frac{\partial v}{\partial x}), \\
 \lambda_y &= \lambda + c_y(c_x \frac{\partial u}{\partial y} + c_y \frac{\partial v}{\partial y}), \\
 \omega_x &= -c_x(c_x \frac{\partial u}{\partial x} + c_y \frac{\partial v}{\partial x}), \\
 \omega_y &= c_y(c_x \frac{\partial u}{\partial x} + c_y \frac{\partial v}{\partial x}),
 \end{aligned}
 \tag{3.7}$$

dan kan $\mathcal{D}$ langs de kusten geschreven worden als

$$\mathcal{D} = - \begin{pmatrix} \lambda_x & -\omega_x & gc_x(c_x \frac{\partial}{\partial x} + c_y \frac{\partial}{\partial y}) \\ \omega_y & \lambda_y & gc_y(c_x \frac{\partial}{\partial x} + c_y \frac{\partial}{\partial y}) \\ h_x + h \frac{\partial}{\partial x} & h_y + h \frac{\partial}{\partial y} & 0 \end{pmatrix}.
 \tag{3.8}$$

Verder gaat de inhomogene term $\vec{f}$ langs de kusten over in (vergelijk (1.7))

$$\vec{f} = \begin{pmatrix} c_x(c_x f_x + c_y f_y) \\ c_y(c_x f_x + c_y f_y) \\ 0 \end{pmatrix}.
 \tag{3.9}$$

Om tot één voorstelling voor het totale differentie-schema te komen voeren we indicatoren a_x , a_y , a_{xy} en b in die volgens tabel 3-1 gedefinieerd zijn

Tabel 3.1. Indicatoren voor het differentieschema

roosterpunten $\vec{x}_j$	$a_x(\vec{x}_j)$	$a_y(\vec{x}_j)$	$a_{xy}(\vec{x}_j)$	$b(\vec{x}_j)$
interne punten	1	1	0	1
kust-punten	$c_x^2(\vec{x}_j)$	$c_y^2(\vec{x}_j)$	$c_x(\vec{x}_j)c_y(\vec{x}_j)$	1
ocean-punten	1	1	0	0

Vervolgens definiëren we de diagonaalmatrices

$$\begin{aligned}
\Lambda_x &= (\lambda_x(\vec{x}_j) \delta_{ij}), & \Lambda_y &= (\lambda_y(\vec{x}_j) \delta_{ij}), \\
\Omega_x &= (\omega_x(\vec{x}_j) \delta_{ij}), & \Omega_y &= (\omega_y(\vec{x}_j) \delta_{ij}), \\
(3.10) \quad A_x &= (a_x(\vec{x}_j) \delta_{ij}), & A_y &= (a_y(\vec{x}_j) \delta_{ij}), & A_{xy} &= (a_{xy}(\vec{x}_j) \delta_{ij}), \\
B &= (b(\vec{x}_j) H(\vec{x}_j) \delta_{ij}), & B_x &= (b(\vec{x}_j) H_x(\vec{x}_j) \delta_{ij}), & B_y &= (b(\vec{x}_j) H_y(\vec{x}_j) \delta_{ij});
\end{aligned}$$

hierin stelt δ_{ij} het Kronecker-symbool voor en doorloopt $\vec{x}_j$ alle roosterpunten zowel interne punten als kust- en oceaantpunten (zie figuur 3.3). Behalve deze diagonaalmatrices definiëren we nog de matrices $\mathcal{D}_x$ en $\mathcal{D}_y$ die de werking van de differentie-operatoren $[\partial/\partial x]$ in alle roosterpunten te samen beschrijven. Deze matrices zijn in het algemeen bandmatrices (bijvoorbeeld de door (2.4.6) gedefinieerde differentie-operator $[\partial/\partial x]$ geeft aanleiding tot een matrix $\mathcal{D}_x$ met 2 banden). Stellen tenslotte U_n , V_n en H_n vectoren voor waarvan de componenten $U_n(\vec{x}_j)$, $V_n(\vec{x}_j)$ en $H_n(\vec{x}_j)$ zijn en waarin $\vec{x}_j$ alle roosterpunten doorloopt (in dezelfde volgorde als in (3.10)) dan kan formula (3.1) als volgt geschreven worden

$$\begin{aligned}
(3.11) \quad & \begin{pmatrix} I + \tau_n \Lambda_x & -\tau_n \Omega_x & \tau_n g(A_x \mathcal{D}_x + A_{xy} \mathcal{D}_y) \\ \tau_n \Omega_y & I + \tau_n \Lambda_y & \tau_n g(A_{xy} \mathcal{D}_x + A_y \mathcal{D}_y) \\ \tau_n (B_x + B \mathcal{D}_x) & \tau_n (B_y + B \mathcal{D}_y) & I \end{pmatrix} \begin{pmatrix} U_{n+1} \\ V_{n+1} \\ H_{n+1} \end{pmatrix} = \\
& \begin{pmatrix} U_n \\ V_n \\ H_n \end{pmatrix} + \tau_n \vec{F}_n, \quad \vec{F}_n = \begin{pmatrix} A_x^F x + A_{xy}^F g \\ A_{xy}^F x + A_y^F y \\ 0 \end{pmatrix},
\end{aligned}$$

waarbij in de coëfficiëntenmatrix en in de inhomogene term alle grootheden in t_n geëvalueerd worden.

Bij implementatie van de algoritme (3.11) moet op elk tijdsniveau t_n het lineaire stelsel (3.11) opgelost worden. Weliswaar is de coëfficiëntenmatrix betrekkelijk "ijl", dat wil zeggen het grootste deel van de elementen is nul, maar de afmetingen van deze matrix zijn formidabel: overdekt men een zee met 50×50 roosterpunten dan dient men steeds een stelsel van 7500 vergelijkingen op te lossen. Het ziet er naar uit dat we met schema (3.11)

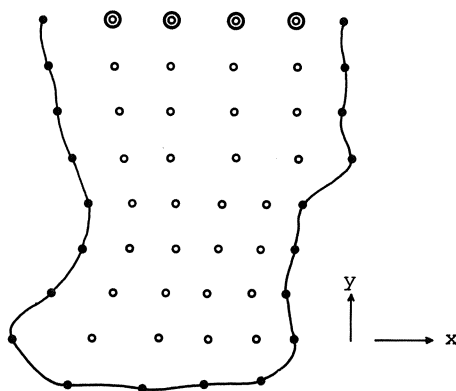


Fig. 3.3. Roosterpunten $\vec{x}_j$ onderscheiden in interne punten $\circ$, kustpunten $\bullet$ en oceaantpunten $\odot$.

stabiliteit gekocht hebben voor een grote bewerkelijkheid, tenzij men efficiënt gebruik weet te maken van het specifieke ijlheidspatroon van de coëfficiënten matrix in (3.11). Dit zou nader onderzocht moeten worden, maar valt buiten het kader van deze syllabus.

In de volgende paragrafen zullen we een aantal in de literatuur voorgestelde algoritmen bespreken; ze blijken allen een tussenvorm van de volledig expliciete Euler-methode en de volledig impliciete formule (3.1) te zijn en geschreven te kunnen worden in de vorm (3.0). We beperken ons voornamelijk tot de "tijdsdiscretisatie" en zullen in een volgend hoofdstuk dieper ingaan op de "plaatsdiscretisatie" dat wil zeggen de keuze van de roosterpunten en de matrices $\mathcal{D}_x$ en $\mathcal{D}_y$.

3.3. Het Fischer-Hansen-schema

Alhoewel FISCHER [1959] zich beperkte tot de gelineariseerde differentiaalvergelijkingen (dus λ en ω constant, en geen convectieve traagheidstermen) is zijn wijze van tijdsdiscretisatie direct over te dragen op de niet-lineaire vergelijkingen; daarom zullen we zijn schema geven voor het algemenere niet-lineaire geval. In de notaties van (2.28), (3.7) en (3.10) ziet het Fischer-schema er als volgt uit:

$$(3.12) \quad \begin{pmatrix} I & O & O \\ O & I & O \\ \tau_n(B_x + B\mathcal{D}_x) & \tau_n(B_y + B\mathcal{D}_y) & I \end{pmatrix} \begin{pmatrix} U_{n+1} \\ V_{n+1} \\ H_{n+1} \end{pmatrix} = \begin{pmatrix} I - \tau_n \Lambda_x & \tau_n \Omega_x & -\tau_n g(A_x \mathcal{D}_x + A_{xy} \mathcal{D}_y) \\ -\tau_n \Omega_y & I - \tau_n \Lambda_y & -\tau_n g(A_{xy} \mathcal{D}_x + A_y \mathcal{D}_y) \\ O & O & I \end{pmatrix} \begin{pmatrix} U_n \\ V_n \\ H_n \end{pmatrix} + \tau_n \vec{F}_n.$$

Aangezien in de coëfficiëntenmatrix in het linkerlid alle elementen boven de diagonaal nul zijn, is schema (3.12) direct naar U_{n+1} , V_{n+1} en H_{n+1} op te lossen en daarmee expliciet uit te schrijven. We krijgen

$$U_{n+1} = (I - \tau_n \Lambda_x) U_n + \tau_n \Omega_x V_n - \tau_n g(A_x \mathcal{D}_x + A_{xy} \mathcal{D}_y) H_n + \tau_n (A_x F_x + A_{xy} F_y)_n,$$

$$(3.12') \quad V_{n+1} = -\tau_n \Omega_y U_n + (I - \tau_n \Lambda_y) V_n - \tau_n g(A_{xy} \mathcal{D}_x + A_y \mathcal{D}_y) H_n + \tau_n (A_{xy} F_x + A_y F_y)_n,$$

$$H_{n+1} = H_n - \tau_n (B_x + B\mathcal{D}_x) U_{n+1} - \tau_n (B_y + B\mathcal{D}_y) V_{n+1}.$$

Het enige verschil met het Euler-schema is dat in de vergelijking voor H_{n+1} nu U_{n+1} en V_{n+1} inplaats van U_n en V_n staat. Deze eenvoudige modificatie die bij implementatie in een computer bovendien nog geheugenbesparend werkt, maakt het stabiliteitsgedrag aanzienlijk beter. Voor het lineaire model, een uniform vierkant rooster, en centrale differentie-operatoren $[\partial/\partial_y]$ en $[\partial/\partial_x]$ gedefinieerd volgens (2.46) heeft Fischer de voorwaarden opgesteld waaronder de eigenwaarden van de lokale amplificatiematrices behorende bij de *interne* roosterpunten binnen de eenheidscirkel liggen. Voor $\lambda \tau_n \ll 1$ en $\omega \tau_n \ll 1$ geeft hij de voorwaarden

$$(3.13) \quad \tau_n < \frac{\sqrt{2}}{gH} \xi, \quad \tau_n < \frac{\lambda}{2\omega},$$

waarin ξ de roosterpunt-afstand in x- en y-richting is. Vergelijken we deze voorwaarden met de overeenkomstige voorwaarden voor het Euler-schema die uit (2.44') volgen door $\eta = \xi$ te stellen, dan zien we dat de maximale stabiele integratiestap nu *lineair* met de maasgrootte samenhangt en niet kwadratisch. In het geval van de Noordzee met een rooster waarin $\xi \approx 20$ km is volgt uit (3.13)

$$(3.13') \quad \tau_n < \frac{9000}{\sqrt{H}};$$

dit laat $18\sqrt{h}$ maal grotere stappen toe dan het Eulerschema (vergelijk (2.44')). Vooral voor de relatief diepe Noordzee is dit een grote winst. Het Fischer-schema is echter onbruikbaar voor modellen zonder bodemwrijving.

Het door HANSEN [1961] gegeven schema ziet er in onze notatie als volgt uit:

$$\begin{aligned} U_{n+\frac{1}{2}} &= (I - \tau \Lambda_x) U_{n-\frac{1}{2}} + \tau \Omega_x V_{n-\frac{1}{2}} - \tau g (A_{xx} \mathcal{D}_x + A_{xy} \mathcal{D}_y) H_n + \tau (A_{xx} F_x + A_{xy} F_y)_n, \\ (3.12'') \quad V_{n+\frac{1}{2}} &= -\tau \Omega_y \bar{U}_{n-\frac{1}{2}} + (I - \tau \Lambda_y) V_{n-\frac{1}{2}} - \tau g (A_{xy} \mathcal{D}_x + A_{yy} \mathcal{D}_y) H_n + \\ &\quad + \tau (A_{xy} F_x + A_{yy} F_y)_n, \\ H_{n+1} &= H_n - \tau (B_x + B \mathcal{D}_x) U_{n+\frac{1}{2}} - \tau (B_y + B \mathcal{D}_y) V_{n+\frac{1}{2}}. \end{aligned}$$

Hierin worden Λ_x , Ω_x , Λ_y en Ω_y in $t_{n-\frac{1}{2}}$, en B_x , B_y en B in t_n geëvalueerd. Vergelijking met het Fischer-schema (3.12') leert dat het enige verschil de indicering van de vectoren U , V en H is (zie figuren 3.4). De motivering is het nagenoeg onafhankelijk van t zijn van de matrices Λ_x , Λ_y , Ω_x , Ω_y , B_x en B ; bovenstaande formulering van het differentieschema is hierdoor een nauwkeurige discretisering van de partiële differentiaalvergelijking. Dit is direct te zien wanneer we (3.12'') beschrijven als

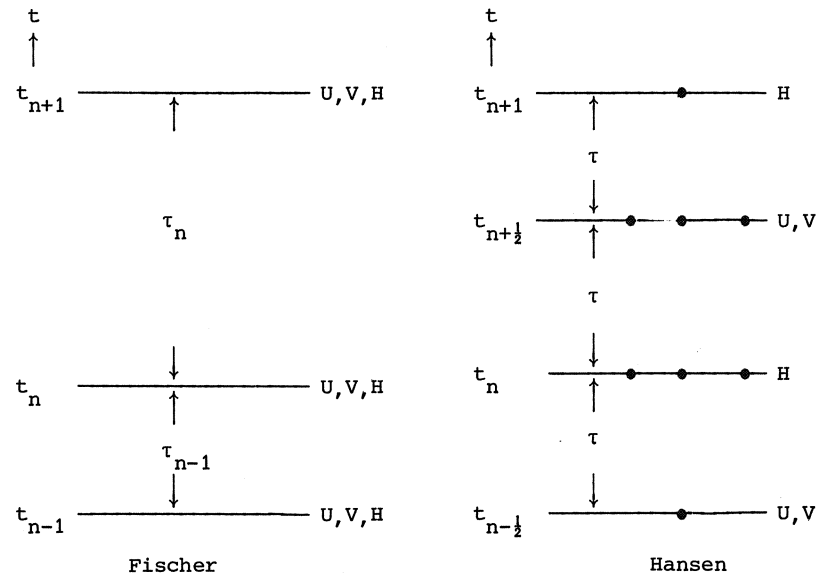


Fig. 3.4. Evaluatie-niveaux voor U, V en H in het Fischer- en Hansen-schema.

$$\frac{U_{n+1/2} - U_{n-1/2}}{\tau} = \Lambda_x U_{n-1/2} + \Omega_x V_{n-1/2} - g(A_x \mathcal{D}_x + A_{xy} \mathcal{D}_y) H_n + (A_x^F + A_{xy}^F) U_n$$

en iets analoogs voor de V- en H-vergelijking. Wanneer ook de bodemwrijvings-term en de Coriolis-term in t_n geëvalueerd zouden worden, zou deze vergelijking een 2^e orde nauwkeurige benadering van de partiële differentiaal-vergelijking voor u geven. Aangezien echter Λ_x en Ω_x relatief kleine bijdragen aan het rechterlid leveren ten opzichte van de vervalterm en de inhomogene term, is de Hansen-formulering inderdaad nauwkeuriger dan die van Fischer (merk op dat de continuïteitsvergelijking nog dichter bij een 2^e orde benadering zit). Het uit elkaar schuiven van de niveaus waarop de stromingen en de diepte berekend worden betekent wel dat met *uniforme* integratiestappen gerekend moet worden.

3.4. Het Sielecki-schema

Een variant op het Fischer-schema werd door SIELECKI [1968] voorgesteld; toegepast op het niet lineaire model zou haar schema van de volgende vorm zijn:

$$\begin{aligned}
 (3.14) \quad & \begin{pmatrix} I & 0 & 0 \\ \tau_n \Omega_Y & I & 0 \\ \tau_n (B_X + B D_X) & \tau_n (B_Y + B D_Y) & I \end{pmatrix} \begin{pmatrix} U_{n+1} \\ V_{n+1} \\ H_{n+1} \end{pmatrix} = \\
 & = \begin{pmatrix} I - \tau_n \Lambda_X & \tau_n \Omega_X & -\tau_n g (A_X D_X + A_{XY} D_Y) \\ 0 & I - \tau_n \Lambda_Y & -\tau_n g (A_{XY} D_X + A_Y D_Y) \\ 0 & 0 & I \end{pmatrix} \begin{pmatrix} U_n \\ V_n \\ H_n \end{pmatrix} + \tau_n \vec{F}_n.
 \end{aligned}$$

Alhoewel dit schema nog steeds expliciet uitgeschreven kan worden, is deze verdere opvoering van de impliciteit bijzonder gunstig voor het stabiliteitsgedrag. Voor het lineaire model met $\lambda = 0$, centrale differenties en een uniform vierkant rooster vindt men de lokale, interne stabiliteitsvoorwaarde

$$(3.15) \quad \tau_n < \sqrt{\frac{2}{gH + \omega^2 \xi^2}} \xi \cong \sqrt{\frac{2}{gH}} \xi.$$

Vergelijking met (3.13) laat zien dat het Sielecki-schema *zonder* bodemwrijving dezelfde stabiliteitsvoorwaarde heeft als het Fischer-schema *met* bodemwrijving.

3.5. Gesymmetriseerd schema

Voortbouwend op de ideeën van Fischer, Hansen en Sielecki komen we tot het volgende schema:

$$\begin{aligned}
 (3.16) \quad & \begin{pmatrix} I + \frac{1}{2} \tau \Lambda_X & -\frac{1}{2} \tau \Omega_X & 0 \\ \frac{1}{2} \tau \Omega_Y & I + \frac{1}{2} \tau \Lambda_Y & 0 \\ \tau (B_X + B D_X) & \tau (B_Y + B D_Y) & I \end{pmatrix} \begin{pmatrix} U_{n+\frac{1}{2}} \\ V_{n+\frac{1}{2}} \\ H_{n+1} \end{pmatrix} = \\
 & = \begin{pmatrix} I - \frac{1}{2} \tau \Lambda_X & \frac{1}{2} \tau \Omega_X & -\tau g (A_X D_X + A_{XY} D_Y) \\ -\frac{1}{2} \tau \Omega_Y & I - \frac{1}{2} \tau \Lambda_Y & -\tau g (A_{XY} D_X + A_Y D_Y) \\ 0 & 0 & I \end{pmatrix} \begin{pmatrix} U_{n-\frac{1}{2}} \\ V_{n-\frac{1}{2}} \\ H_n \end{pmatrix} + \tau \vec{F}_n.
 \end{aligned}$$

Dit schema is als het ware het "maximaal impliciete" schema dat nog expliciet uitgeschreven kan worden. Immers, laten de rechterleden van dit schema R_1 , R_2 en R_3 zijn, dan kunnen we schrijven:

$$U_{n+\frac{1}{2}} = [I + \frac{1}{2} \tau (\Lambda_x + \Lambda_y) + \frac{1}{4} \tau^2 (\Lambda_x \Lambda_y + \Omega_x \Omega_y)]^{-1} \cdot [(I + \frac{1}{2} \tau \Lambda_y) R_1 + \frac{1}{2} \tau \Omega_x R_2], \quad (3.16')$$

$$V_{n+\frac{1}{2}} = [I + \frac{1}{2} \tau \Lambda_y]^{-1} [R_2 - \frac{1}{2} \tau \Omega_y U_{n+\frac{1}{2}}],$$

$$H_{n+1} = H_n - \tau (B_x + B \mathcal{D}_x) U_{n+\frac{1}{2}} - \tau (B_y + B \mathcal{D}_y) V_{n+\frac{1}{2}}.$$

Aangezien de hier optredende inverse matrices steeds diagonaalmatrices betreffen, zijn de vectoren $U_{n+\frac{1}{2}}$, $V_{n+\frac{1}{2}}$ en H_{n+1} rechtstreeks te berekenen uit $U_{n-\frac{1}{2}}$, $V_{n-\frac{1}{2}}$ en H_n .

We hebben schema (3.16) *gesymmetriseerd schema* genoemd omdat het een vrijwel maximale symmetrie bezit binnen de klasse van expliciet voorstelbare schema's. Dit kunnen we zien door (3.16) te schrijven als

$$\begin{aligned} \frac{U_{n+\frac{1}{2}} - U_{n-\frac{1}{2}}}{\tau} &= -\frac{1}{2} \Lambda_x (U_{n+\frac{1}{2}} + U_{n-\frac{1}{2}}) + \frac{1}{2} \Omega_x (V_{n+\frac{1}{2}} + V_{n-\frac{1}{2}}) \\ &\quad - g(A_{xx} \mathcal{D}_x + A_{xy} \mathcal{D}_y) H_n + (A_{xx}^F + A_{xy}^F)_{n'}, \\ (3.16'') \quad \frac{V_{n+\frac{1}{2}} - V_{n-\frac{1}{2}}}{\tau} &= -\frac{1}{2} \Omega_y (U_{n+\frac{1}{2}} + U_{n-\frac{1}{2}}) - \frac{1}{2} \Lambda_y (V_{n+\frac{1}{2}} + V_{n-\frac{1}{2}}) \\ &\quad - g(A_{xy} \mathcal{D}_x + A_{yy} \mathcal{D}_y) H_n + (A_{xy}^F + A_{yy}^F)_n \\ \frac{H_{n+1} - H_n}{\tau} &= - (B_x + B \mathcal{D}_x) U_{n+\frac{1}{2}} - (B_y + B \mathcal{D}_y) V_{n+\frac{1}{2}}. \end{aligned}$$

Voor constante matrices $\Lambda_x, \Lambda_y, \Omega_x, \Omega_y, B_x, B_y$ en B zou dit schema volledig symmetrisch zijn en daarmee 2^e orde exact wat de tijdsdiscretisatie betreft. In niet-lineaire modellen zijn deze matrices echter van t afhankelijk en worden

analoog aan het Hansen-schema (3.12") de matrices $\Lambda_x, \Omega_x, \Lambda_y, \Omega_y$ in $t_{n-\frac{1}{2}}$ en B_x, B_y en B in t_n geëvalueerd. In dit verband moet genoemd worden dat K. Dekker voorstelt om schema (3.16") nog wat symmetrischer te maken door in de continuïteitsvergelijking de matrix B te vervangen door het gemiddelde van B_n en B_{n+1} . Definieren we in aanvulling op (3.10), de diagonaalmatrix

$$(3.10') \quad C = \frac{1}{2} (b(\vec{x}_j) (\frac{\partial U}{\partial x}(\vec{x}_j) + \frac{\partial V}{\partial y}(\vec{x}_j)) \delta_{ij}),$$

dan kan men schrijven

$$(3.17) \quad H_{n+1} = [I + \tau C]^{-1} [H_n - \tau (B_x + \frac{1}{2} B D_x) U_{n+\frac{1}{2}} - \tau (B_y + \frac{1}{2} B D_y) V_{n+\frac{1}{2}}],$$

waarin B_x, B_y en B in t_n en C in $t_{n+\frac{1}{2}}$ geëvalueerd worden.

3.6. De stabiliteit van het gesymmetriseerde schema

Tot dusver hebben we ons beperkt tot het geven van stabiliteitsvoorwaarden voor het *lineaire* model. In deze paragraaf geven we een afleiding van de (locale) stabiliteitsvoorwaarden van het gesymmetriseerde schema (3.16) voor het *niet-lineaire* model. Hierbij volgen we de analyse gegeven in VAN DER HOUWEN [1975]. Een aanpassing van deze analyse aan de discretisatie (3.17) van Dekker laten we aan de lezer over.

Volgens de methode van constante coëfficiënten vervangen we de matrices in schema (3.16) door "constante matrices", dat wil zeggen de rij-elementen van elke matrix worden vervangen door de rij-elementen welke met een vast roosterpunt $\vec{x}$ corresponderen. Vervolgens denken we ons de toestandsvector

$$(U_{n-\frac{1}{2}}, V_{n-\frac{1}{2}}, H_n)^T$$

verstoord door een vector opgebouwd uit componenten van het type (2.34); zo'n component ziet er dus uit als

$$(3.18) \quad \begin{pmatrix} \Delta \tilde{U}_{n-\frac{1}{2}}(\vec{\gamma}) & \begin{pmatrix} \exp[i(\vec{\gamma}, \vec{x}_1)] \\ \dots \\ \exp[i(\vec{\gamma}, \vec{x}_r)] \end{pmatrix} \\ \Delta \tilde{V}_{n-\frac{1}{2}}(\vec{\gamma}) & \begin{pmatrix} \exp[i(\vec{\gamma}, \vec{x}_1)] \\ \dots \\ \exp[i(\vec{\gamma}, \vec{x}_r)] \end{pmatrix} \\ \Delta \tilde{H}_n(\vec{\gamma}) & \begin{pmatrix} \exp[i(\vec{\gamma}, \vec{x}_1)] \\ \dots \\ \exp[i(\vec{\gamma}, \vec{x}_r)] \end{pmatrix} \end{pmatrix},$$

waarin $\vec{\gamma}$ de component-index is, r het aantal roosterpunten voorstelt en $\Delta \tilde{U}_{n-\frac{1}{2}}(\vec{\gamma})$, $\Delta \tilde{V}_{n-\frac{1}{2}}(\vec{\gamma})$, $\Delta \tilde{H}_n(\vec{\gamma})$ nader te bepalen scalaires zijn. Tenslotte definiëren we de eigenwaarden $\delta_{\vec{x}}$ en $\delta_{\vec{y}}$ van $\mathcal{D}_{\vec{x}}(\vec{x})$ en $\mathcal{D}_{\vec{y}}(\vec{x})$, dus

$$\mathcal{D}_{\vec{x}}(\vec{x}) \begin{pmatrix} \exp[i(\vec{\gamma}, \vec{x}_1)] \\ \dots \\ \exp[i(\vec{\gamma}, \vec{x}_r)] \end{pmatrix} = \delta_{\vec{x}}(\vec{x}, \vec{\gamma}) \begin{pmatrix} \exp[i(\vec{\gamma}, \vec{x}_1)] \\ \dots \\ \exp[i(\vec{\gamma}, \vec{x}_r)] \end{pmatrix}$$

en iets dergelijks voor $\delta_{\vec{y}}$. Merk op dat $\delta_{\vec{x}}$ en $\delta_{\vec{y}}$ zowel van de index $\vec{\gamma}$ als het gekozen roosterpunt $\vec{x}$ afhangen.

Substitueren we nu verstoringen van de vorm (3.18) in het "gelocaliseerde" schema (3.16) dan zien we dat deze verstoringen vermenigvuldigd worden met de amplificatiematrix

$$(3.20) \quad A(\vec{x}, \vec{\gamma}) = \begin{pmatrix} (1 + \frac{1}{2}\tau\lambda_{\vec{x}})I & -\frac{1}{2}\tau\omega_{\vec{x}} I & 0 \\ \frac{1}{2}\tau\omega_{\vec{y}} I & (1 + \frac{1}{2}\tau\lambda_{\vec{y}})I & 0 \\ \tau b(H_{\vec{x}} + H\delta_{\vec{x}})I & \tau b(H_{\vec{y}} + H\delta_{\vec{y}})I & I \end{pmatrix}^{-1} \cdot$$

$$\begin{pmatrix} (1 - \frac{1}{2}\tau\lambda_{\vec{x}})I & \frac{1}{2}\tau\omega_{\vec{x}} I & -\tau\lambda(a_{\vec{x}}\delta_{\vec{x}} + a_{\vec{xy}}\delta_{\vec{y}})I \\ -\frac{1}{2}\tau\omega_{\vec{y}} I & (1 - \frac{1}{2}\tau\lambda_{\vec{y}})I & -\tau g(a_{\vec{xy}}\delta_{\vec{y}} + a_{\vec{y}}\delta_{\vec{y}})I \\ 0 & 0 & I \end{pmatrix}$$

Hierin hebben we de argumenten $\vec{x}$ en $\vec{\gamma}$, en tijdsindices weggelaten. We schrijven A in de vorm

$$(3.20') \quad A(\vec{x}, \vec{y}) = A(\vec{x}, \vec{y}) \quad I = \tilde{M}^{-1} \tilde{N} I,$$

waarin $\tilde{M}$ en $\tilde{N}$ (3x3)-matrices zijn gedefinieerd overeenkomstig (3.20). Uit deze schrijfwijze volgt direct dat de eigenwaarden α van $A(\vec{x}, \vec{y})$ voldoen aan de vergelijking

$$\det [\tilde{M}^{-1} \tilde{N} - \alpha I] = 0,$$

ofwel

$$(3.21) \quad \det [\tilde{M}^{-1}] - \det [\tilde{N} - \alpha \tilde{M}] = 0.$$

Aannemende dat $\tilde{M}$ niet singulier is, vinden we voor α de vergelijking

$$(3.21') \quad \det \begin{bmatrix} (1-\alpha)(1+\frac{1}{2}\tau\lambda_x) - \tau\lambda_x & -\frac{1}{2}\tau\omega_x(1-\alpha) + \tau\omega_x & -\tau g(a_x\delta_x + a_{xy}\delta_y) \\ \frac{1}{2}\tau\omega_y(1-\alpha) - \tau\omega_y & (1-\alpha)(1+\frac{1}{2}\tau\lambda_y) - \tau\lambda_y & -\tau g(a_{xy}\delta_y + a_y\delta_y) \\ -\tau b(H_x + H\delta_x)\alpha & -\tau b(H_y + H\delta_y)\alpha & 1 - \alpha \end{bmatrix} = 0.$$

We zullen de wortels van deze vergelijking onderzoeken voor *interne* roosterpunten $\vec{x}$; volgens tabel 3.1 geldt dan

$$a_x = a_y = b = 1, \quad a_{xy} = 0.$$

Verder veronderstellen we, evenals in (2.38),

$$\frac{|H_x| + |H_y|}{H} \ll 1.$$

Elementair rekenwerk leidt dan tot de eigenwaardenvergelijking

$$(3.22) \quad b_0(\alpha-1)^3 + b_1(\alpha-1)^2 + b_2(\alpha-1) + b_3 = 0,$$

waarin

$$b_0 = (1+\frac{1}{2}\tau\lambda_x)(1+\frac{1}{2}\tau\lambda_y) + \frac{1}{4}\tau^2\omega_x\omega_y,$$

$$b_1 = \tau(\lambda_x + \lambda_y) + b_2 - b_3,$$

$$b_2 = \tau^2 (\lambda_x \lambda_y + \omega_x \omega_y) - \tau^2 gH(\delta_x^2 + \delta_y^2) + \frac{3}{2} b_3,$$

$$b_3 = -\tau^3 gN(\lambda_y \delta_x^2 + (\omega_x - \omega_y) \delta_x \delta_y + \lambda_x \delta_y^2).$$

Door vergelijking (3.22) uit te schrijven in machten van α en toepassing van het Hurwitz-criterium (2.27) vinden we dat $|\alpha| < 1$ wanneer

$$b_0 > 0,$$

$$b_3 > 0,$$

$$(3.23) \quad 8b_0 - 4b_1 + 2b_2 - b_3 > 0,$$

$$4(b_1 - b_2) + 3b_3 > 0,$$

$$(b_2 - b_3)(b_1 - b_2 + b_3) > b_0 b_3$$

waarin er van uitgegaan is dat de coëfficiënten b_j allen reëel zijn. We zullen in het vervolg aannemen dat δ_x en δ_y zuiver-imaginair zijn (vergelijk (2.47')); de coëfficiënten b_j zijn dan in elk geval reëel. Substitutie van b_j in (3.27) levert de voorwaarden

$$1 + \frac{1}{2}\tau(\lambda_x + \lambda_y) + \frac{1}{4}\tau^2(\lambda_x \lambda_y + \omega_x \omega_y) > 0,$$

$$\lambda_y \delta_x^2 + (\omega_x - \omega_y) \delta_x \delta_y + \lambda_x \delta_y^2 < 0,$$

(3.23')

$$4 + \tau^2 gH(\delta_x^2 + \delta_y^2) > 0,$$

$$4(\lambda_x + \lambda_y) + \tau^2 gH(\lambda_y \delta_x^2 + (\tau^2 gH(\lambda_y \delta_x^2 + (\omega_x - \omega_y) \delta_x \delta_y + \lambda_x \delta_y^2) > 0,$$

$$(\lambda_x + \lambda_y)\epsilon > gH[(\lambda_x - \frac{1}{4}\tau^2 \lambda_y \epsilon) \delta_x^2 - (\omega_x - \omega_y)(1 + \frac{1}{4}\tau^2 \epsilon) \delta_x \delta_y + (\lambda_y - \frac{1}{4}\tau^2 \lambda_x \epsilon) \delta_y^2],$$

waarin $\epsilon = \lambda_x \lambda_y + \omega_x \omega_y$.

Voor $\delta_x \neq 0$ of $\delta_y \neq 0$ en $\tau^2 \epsilon$ verwaarloosbaar klein ten op zichte van $\lambda_x/\lambda_y, \lambda_y/\lambda_x$ en 1 volgen hieruit de lokale stabiliteitsvoorwaarden:

$$\tau < \frac{\text{Min}}{\gamma} \frac{2}{\sqrt{gH|\delta_x^2 + \delta_y^2|}},$$

$$(3.24) \quad \lambda_x > 0, \lambda_y > 0,$$

$$\lambda_x \lambda_y + \omega_x \omega_y > \frac{1}{4}(\omega_x + \omega_y)^2.$$

Voor centrale differentie-operatoren $\mathcal{D}_x$ en $\mathcal{D}_y$ gaat de voorwaarde voor τ over in de eerste voorwaarde van Fischer (zie formules (3.13)). De condities voor λ_x , λ_y , ω_x en ω_y zijn juist de inherente stabiliteitsvoorwaarden van het wiskundige model zelf (vergelijk de voorwaarden (2.39)). Indien δ_x en δ_y gelijktijdig nul zijn kan niet aan de tweede ongelijkheid in (3.23') voldaan worden. Dit betekent dat er minstens één eigenwaarde van $A(\vec{x})$ op de eenheidscirkel ligt; dit is echter onvermijdbaar en volgt direct uit het feit dat de matrix $\mathcal{D}(\vec{x})$ altijd een eigenwaarde in nul heeft hoe de discretisaties van $\partial/\partial x$ en $\partial/\partial y$ ook gekozen worden (zie vergelijking (2.37')).

Vergelijken we het gesymmetriseerde schema met het schema van Sielecki, dan zien we dat het stabiliteitsgedrag gelijk is en dat de voordelen in een grotere nauwkeurigheid gezocht moeten worden.

Tenslotte merken we op dat voorwaarden (3.24) in de interne roosterpunten gelden. Hoe de stabiliteitsvoorwaarden in de kust- en oceaantpunten zijn, laten we aan de lezer over.

3.7. Impliciete schema's

Het ziet er naar uit dat de stabiliteitsvoorwaarden (3.2.4) het uiterst bereikbare is voor expliciet uitschrijfbare gegeneraliseerde Euler-schema's. Om wezenlijk grotere integratiestappen te kunnen nemen moeten we tot echt impliciete schema's overgaan. We zullen nu schema's van de vorm (zonder $\vec{F}_n$ -term)

$$(3.25) \quad \begin{pmatrix} I + \frac{1}{2}\tau\Lambda_x & -\theta_3\tau\Omega_x & \theta_1\tau g(A_x\mathcal{D}_x + A_{xy}\mathcal{D}_y) \\ \theta_4\tau\Omega_y & I + \frac{1}{2}\tau\Lambda_y & \theta_2\tau g(A_{xy}\mathcal{D}_x + A_y\mathcal{D}_y) \\ \theta_5\tau(B_x + B\mathcal{D}_x) & \theta_6\tau(B_y + B\mathcal{D}_y) & I \end{pmatrix} \begin{pmatrix} U_{n+1} \\ V_{n+1} \\ H_{n+1} \end{pmatrix} =$$

$$= \begin{pmatrix} I - \frac{1}{2}\tau\Lambda_x & (I - \theta_3)\tau\Omega_x & (\theta_1 - I)\tau g(A_x\mathcal{D}_x + A_{xy}\mathcal{D}_y) \\ (\theta_4 - I)\tau\Omega_y & I - \frac{1}{2}\tau\Lambda_y & (\theta_2 - I)\tau g(A_{xy}\mathcal{D}_x + A_y\mathcal{D}_y) \\ (\theta_5 - I)\tau(B_x + B\mathcal{D}_x) & (\theta_6 - I)\tau(B_y + B\mathcal{D}_y) & I \end{pmatrix} \begin{pmatrix} U_n \\ V_n \\ H_n \end{pmatrix}.$$

beschouwen. Hierin worden alle coëfficiëntenmatrices in t_n geëvalueerd en zijn de θ_i vrij te kiezen roosterfuncties. In het bijzonder zullen we nader het geval

$$(3.26) \quad \begin{aligned} \theta_1 = \theta_2 = \theta_4 = \theta_5 = \frac{1}{2}, \theta_3 = \theta_6 = 0 & \quad \text{als } n = \text{even,} \\ \theta_1 = \theta_2 = \theta_3 = \theta_6 = \frac{1}{2}, \theta_4 = \theta_5 = 0 & \quad \text{als } n = \text{oneven} \end{aligned}$$

bekijken. Een schema dat grote gelijkenis vertoont met bovenstaand schema (3.25) - (3.26) is onderzocht door LEENDERTSE [1967]. We zullen echter niet ingaan op zijn schema omdat de verschillende modificaties die Leendertse op grond van experimenten aangebracht heeft, buiten het kader van deze syllabus vallen en bovendien niet essentieel zijn voor de grondgedachte.

Allereerst gaan we de vergelijkingen in stelsel (3.25) herordenen en wel volgens

$$(3.27) \quad \dots, U(\vec{x}_j), H(\vec{x}_j), U(\vec{x}_{j+1}), H(\vec{x}_{j+1}), \dots; \dots, V(\vec{x}_j), V(\vec{x}_{j+1}), \dots$$

$\vec{x}_j$ doorloopt de roosterpunten "evenwijdig aan de x-as"

als $n = \text{even}$ en volgens

$$(3.27') \quad \dots, V(\vec{x}_j), H(\vec{x}_j), V(\vec{x}_{j+1}), H(\vec{x}_{j+1}), \dots; \dots, U(\vec{x}_j), U(\vec{x}_{j+1}), \dots$$

$\vec{x}_j$ doorloopt de roosterpunten "evenwijdig aan de y-as"

als $n = \text{oneven}$ (zie figuur 3.4)

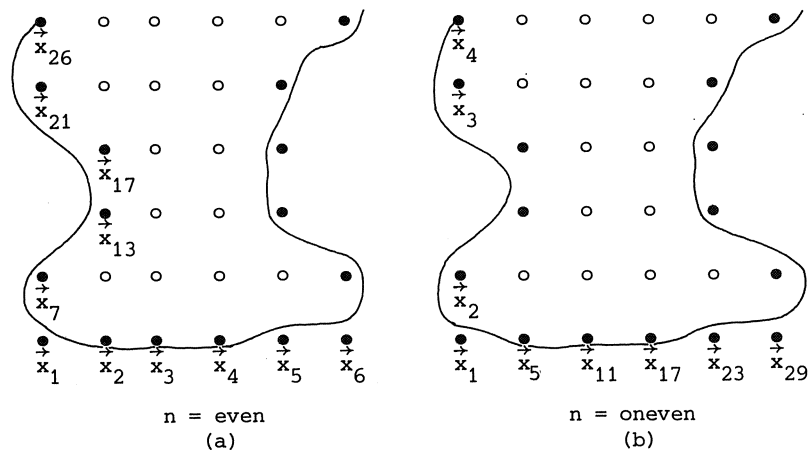


Fig. 3.4. Ordening van de roosterpunten "evenwijdig" aan de x-as respectievelijk y-as

Verder kiezen we evenals Leendertse de stroomvector in de kustpunten of evenwijdig aan de x-as of evenwijdig aan de y-as. Dit betekent dat we een type kuststroming krijgen zoals in figuur 3.5 is aangegeven.

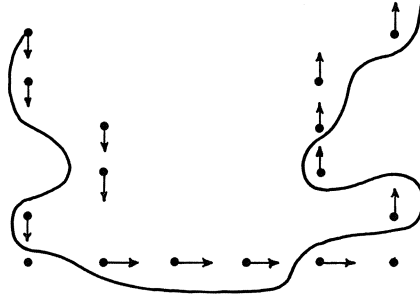


Fig. 3.5. Kuststroming volgens Leendertse.

Deze presentatie van de randvoorwaarden heeft het voordeel dat de matrix A_{xy} in (3.25) de nul-matrix wordt; een nadeel is dat de "sturende" werking van de kust erg ruw in rekening gebracht wordt.

We zullen nu het patroon van de coëfficiëntenmatrix van het lineaire stelsel (3.25) onderzoeken voor het geval dat $\mathcal{D}_x$ en $\mathcal{D}_y$ door middel van *centrale differenties* gedefinieerd zijn. Zij n even en het rooster van de vorm zoals in figuur 3.4 weergegeven; stel verder dat we asymmetrische differenties gebruiken in randpunten als $\vec{x}_1$ en $\vec{x}_6$. Voor U_{n+1} en H_{n+1} krijgen we dan een stelsel van de volgende vorm

$$(3.28) \quad \begin{pmatrix} x & & & & & & & & & & \\ & x & x & & & & & & & & \\ & x & x & 0 & 0 & x & & & & & \\ & & x & x & x & & & & & & \\ & & & x & x & 0 & 0 & x & & & \\ & & x & 0 & x & x & x & & & & \\ & & & & x & x & 0 & 0 & x & & \\ & & & & & x & 0 & x & x & x & \\ & & & & & & . & . & . & & \end{pmatrix} \begin{pmatrix} U(\vec{x}_1) \\ H(\vec{x}_1) \\ U(\vec{x}_2) \\ H(\vec{x}_2) \\ U(\vec{x}_3) \\ H(\vec{x}_3) \\ U(\vec{x}_4) \\ H(\vec{x}_4) \\ \dots \\ n+1 \end{pmatrix} = \begin{pmatrix} 0 \\ R_3(\vec{x}_1) \\ R_1(\vec{x}_2) \\ R_3(\vec{x}_2) \\ . \\ . \\ . \\ . \\ . \\ n+1 \end{pmatrix},$$

waarin R_1 en R_3 de rechterleden in stelsel (3.25) voorstellen. De vergelijking voor V_{n+1} is expliciet te schrijven als

$$(3.29) \quad V_{n+1} = [I + \frac{1}{2}\tau\Lambda_y]^{-1} [R_2 - \frac{1}{2}\tau\Omega_y U_{n+1} - \frac{1}{2}\tau g A_y \mathcal{D}_y H_{n+1}].$$

Stelsel (3.28) heeft een coëfficiëntenmatrix met bandbreedte 7, dus een aanzienlijk smallere band dan die van schema (3.11). Men kan deze bandbreedte echter nog kleiner maken; daartoe vervangen we de matrix B_x door een matrix C_x die gedefinieerd is volgens

$$(3.30) \quad C_x U(\vec{x}_j) = \frac{1}{2} B_x (U(\vec{x}_{j+1}) + U(\vec{x}_{j-1})).$$

in alle inwendige punten en onveranderd blijft in de randpunten.

De coëfficiëntenmatrix van (3.28) gaat door deze modificatie over in

$$(3.28') \quad \begin{pmatrix} x & & & & & & & & \\ & x & x & & & & & & \\ & x & x & 0 & 0 & x & & & \\ x & 0 & 0 & x & x & & & & \\ & & x & x & 0 & 0 & x & & \\ & x & 0 & 0 & x & x & & & \\ & & & x & x & 0 & 0 & x & \\ & & & x & 0 & 0 & x & x & \\ & & & & & & & \dots & \end{pmatrix}.$$

We zien nu dat de $2^e, 3^e, 6^e, 7^e, 10^e, 11^e, \dots$ vergelijking ontkoppeld zijn van de $1^e, 4^e, 5^e, 8^e, 9^e, 12^e, 13^e, \dots$ vergelijking, of wel de vergelijkingen voor $H(\vec{x}_1), U(\vec{x}_2), \dots, H(\vec{x}_{2j+1}), U(\vec{x}_{2j+2}), \dots$ zijn onafhankelijk van de vergelijkingen voor $U(\vec{x}_1), \dots, H(\vec{x}_{2j}), U(\vec{x}_{2j+1}), \dots$.

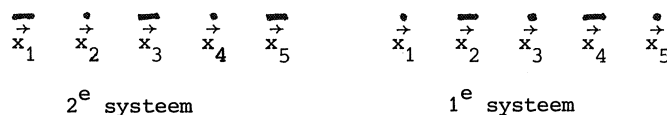


Fig. 3.6. Stroomingspunten (→) en verhogingspunten (•) voor de ontkoppelde systemen

Hiermee zijn de vergelijkingen bij elke horizontale rij roosterpunten van figuur 3.4a opgesplitst in twee onafhankelijke systemen. In het eerste systeem worden de x-component van de stroming "even" roosterpunten en de diepte in "oneven" roosterpunten berekend, in het tweede systeem vice versa (zie figuur 3.6).

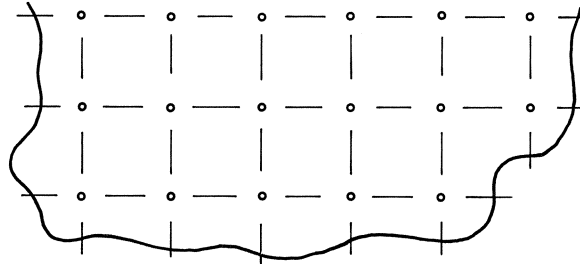


Fig. 3.7. "Space-staggered grid" van Hansen, Platzman en Leendertse.

o : Verhogingspunten
 - : U-stromingspunten
 | : V-stromingspunten

Beide systemen hebben een *tridiagonale* matrix waarvoor bijzonder efficiënte algoritmen beschikbaar zijn.

We leggen er de nadruk op dat de ontkoppeling van de twee systemen betekent dat volstaan kan worden met het oplossen van één der systemen en wel datgene waarin de randvoorwaarden beide verrekend zijn, dus in figuur 3.6 het 2^e systeem met de nevenvoorwaarde dat elke rij uit een *oneven* aantal roosterpunten bestaat. Door weglating van het tweede systeem wordt de *orde van nauwkeurigheid niet minder*; het enige effect is dat men in *minder* punten de oplossing uitrekent.

Bovenstaande beschouwingen gelden voor de *even* tijd niveaus. Voor *oneven* n kan men een soortgelijke analyse geven die ook leidt tot een tri-diagonaal systeem voor de y -component van de stroming en de diepte. Weglating van het systeem dat roosterpunten betreft die *niet* in het systeem van de even tijd niveaus voorkomen, geeft dan aanleiding tot een rooster zoals in figuur 3.7 aangegeven is (in de Engelstalige literatuur een "space-staggered grid" genoemd). Een dergelijk rooster werd al door PLATZMAN [1959] en HANSEN [1961] gebruikt, en is ook door LEENDERTSE [1967] overgenomen.

We staan nog even stil bij de evaluatie van de rechterleden R_1 , R_2 , R_3 en de evaluatie van de matrices Λ_x , Λ_y , Ω_x , Ω_y , B , B_x en B_y op het "uit elkaar getrokken" rooster van figuur 3.7. Omdat in een roosterpunt nu niet alle drie componenten U , V en H beschikbaar zijn, spreken we af dat niet-

beschikbare componenten waar nodig door *middeling* berekend worden. Bijvoorbeeld bij de berekening van U_{n+1} wordt in de U-punten gevraagd naar

$$\begin{aligned}\Omega_x V_n &= \left(\left(\omega - \frac{\partial U_n}{\partial y}(\vec{x}_j) \right) V_n(\vec{x}_j) \delta_{ij} \right) \\ &\approx \left(\left(\omega - \left[\frac{\partial}{\partial y} \right] U_n(\vec{x}_j) \right) V_n(\vec{x}_j) \delta_{ij} \right) \\ &= (\omega I - \mathcal{D}_y U_n) V_n.\end{aligned}$$

We zouden dus in de U-punten V_n moeten kennen en in de middelpunten van de "verhogingsmazen" (zie figuur 3.8) de waarden van U_n .

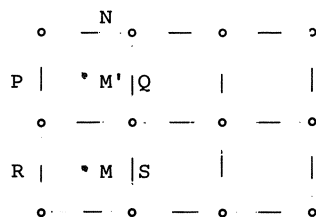


Fig. 3.8. Berekening van $V_n(M)$ en $U_n(M)$

Deze kunnen met behulp van de formules (zie figuur 3.8)

$$V_n(M) = \frac{1}{4}(V_n(P) + V_n(Q) + V_n(R) + V_n(S)), U_n(M') = \frac{1}{2}(U_n(N) + U_n(M)),$$

berekend worden.

Alleen dicht bij de kusten kan dit moeilijkheden opleveren. In de praktijk brengt men de convectieve traagheidstermen aan de kust meestal niet in rekening (zie HANSEN [1961] en LEENDERTSE [1967]), zodat daarmee ook dit probleem opgelost is.

De stabiliteit van schema (3.25) kan op de inmiddels bekende manier geanalyseerd worden.

IV. GESTABILISEERDE RUNGE-KUTTAMETHODEN

De in het voorgaande hoofdstuk behandelde gegeneraliseerde Euler-methoden zijn voor *niet-lineaire modellen* allen van de *eerste orde* in de variabele t . In dit hoofdstuk bespreken we een klasse van 2e orde methoden waarmee stelsels gewone differentiaalvergelijkingen van de vorm

$$(4.1) \quad \frac{d\vec{s}}{dt} = \vec{r}(t, \vec{s})$$

geïntegreerd kunnen worden. In paragraaf 2.8 hebben we gezien dat de waterbeweging in ondiepe zeeën beschreven kan worden door de vergelijking

$$(4.2) \quad \frac{\partial}{\partial t} \vec{s} = \mathcal{D} \vec{s} + \vec{f},$$

waarin $\mathcal{D}$ een operator is die van $\vec{s}$ afhangt en partiele differentiaties naar x en y bevat. Door deze differentiaties door differentie-quotienten te vervangen op een vooraf gekozen rooster in het (x,y) -vlak, gaat (4.2) over in een (zeer groot) stelsel van gewone differentiaalvergelijkingen van de vorm (4.1). Deze "partiele discretisering" wordt de *methode der lijnen* of ook wel *semi-discretisatie* genoemd.

4.1. Algemene Runge-Kuttamethoden

De algemene m -punts expliciete Runge-Kuttaformule wordt gedefinieerd door

$$(4.3) \quad \begin{aligned} \vec{S}_{n+1}^{(0)} &= \vec{S}_n, \\ \vec{S}_{n+1}^{(j)} &= \vec{S}_n + \tau_n \sum_{\ell=0}^{j-1} \lambda_{j,\ell} \vec{r}(t_n + \mu_{\ell} \tau_n, \vec{S}_{n+1}^{(\ell)}), \\ \vec{S}_{n+1} &= \vec{S}_{n+1}^{(m)} \end{aligned}$$

Hierin zijn $\vec{S}_n$ en $\vec{S}_{n+1}$ numerieke benaderingen voor $\vec{s}(t_n)$ en $\vec{s}(t_n + \tau_n)$; de parameters $\lambda_{j,\ell}$ en μ_j zijn nog vrij te kiezen en kunnen gebruikt worden om de orde van *nauwkeurigheid* en de *stabiliteit* op te voeren. In de praktijk kiest men altijd (en in deze syllabus zullen we dat ook doen)

$$(4.4) \quad \mu_0 = 0, \quad \mu_j = \sum_{\ell=0}^{j-1} \lambda_{j,\ell}, \quad j = 1, 2, \dots, m,$$

hetgeen geen beperking voor de algemeenheid blijkt te zijn.

Het Runge-Kuttaschema (4.3) is in feite een gecompliceerd Euler-schema. Dit illustreren we met behulp van figuur 4.1.

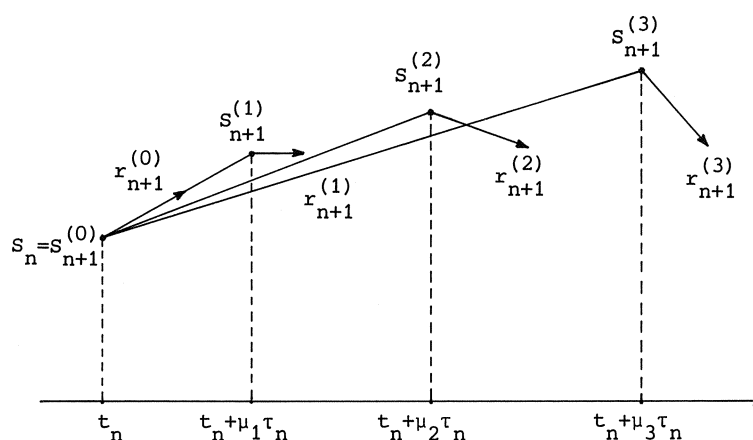


Fig. 4.1. Meetkundige illustratie van de tussenpunten $S_{n+1}^{(j)}$ in een scalar geval ($r_{n+1}^{(j)} = r(t_n + \mu_j \tau_n, S_{n+1}^{(j)})$).

Het eerste tussenresultaat $\vec{S}_{n+1}^{(1)}$ verkrijgen we door een Euler-stap met stapgrootte $\lambda_{10} \tau_n = \mu_1 \tau_n$ in t_n toe te passen in de richting $\vec{r}_{n+1}^{(0)} = \vec{r}(t_n + \mu_0 \tau_n, \vec{S}_{n+1}^{(0)}) = \vec{r}(t_n, \vec{S}_n)$. In figuur 4.1 is dit voor een scalaire vergelijking getekend. In het punt $(t_n + \mu_1 \tau_n, \vec{S}_{n+1}^{(1)})$ bepalen we de richting van het richtingsveld $\vec{r}_{n+1}^{(1)} = \vec{r}(t_n + \mu_1 \tau_n, \vec{S}_{n+1}^{(1)})$. We stellen nu uit de richtingen $\vec{r}_{n+1}^{(0)}$ en $\vec{r}_{n+1}^{(1)}$ een nieuwe richting

$$\frac{\lambda_{20}}{\mu_2} \vec{r}_{n+1}^{(0)} + \frac{\lambda_{21}}{\mu_2} \vec{r}_{n+1}^{(1)}$$

samen en doen in deze richting een Euler-stap over een afstand $\mu_2 \tau_n$, enz. Het richtingsveld wordt als het ware afgetast tot dat binnen het aantal toegestane stappen de meest "gunstige" richting gevonden is. De criteria voor een "goede" richting zijn *nauwkeurigheid* en *stabiliteit*.

4.2. Runge-Kuttamethoden van de 2^e orde

Een Runge-Kuttamethode wordt van de orde p genoemd wanneer de eerste $p + 2$ termen van de Taylor-ontwikkeling van $\vec{S}_{n+1} = \vec{S}_{n+1}(\tau_n)$ in machten van τ_n overeenstemmen met de eerste $p + 2$ termen van de Taylor-ontwikkeling van de oplossing $\vec{s}(t_n + \tau_n)$ van vergelijking (4.1) door het punt $(t_n, \vec{S}_n)$ (zie figuur 4.2).

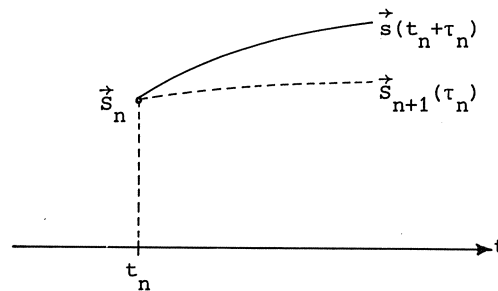


Fig. 4.2. Numerieke en lokaal analytische oplossing.

We zullen dit illustreren voor een 2-punts Runge-Kuttaformule toegepast op een scalaire differentiaalvergelijking:

$$S_{n+1}^{(0)} = S_n,$$

$$(4.3') \quad S_{n+1}^{(1)} = S_n + \lambda_{10} \tau_n r(t_n, S_n),$$

$$S_{n+1} = S_n + \lambda_{20} \tau_n r(t_n, S_n) + \lambda_{21} \tau_n r(t_n + \mu_1 \tau_n, S_{n+1}^{(1)}).$$

Ontwikkeling van S_{n+1} in een Taylorreeks geeft

$$\begin{aligned}
 S_{n+1} &= S_n + \lambda_{20} \tau_n (t_n, S_n) + \lambda_{21} \tau_n [r(t_n, S_n) + \\
 &\quad + \mu_1 \tau_n \frac{\partial r}{\partial t}(t_n, S_n) + \lambda_{10} \tau_n r(t_n, S_n) \frac{\partial r}{\partial s}(t_n, S_n)] \\
 &\quad + O(\tau_n^3) = \\
 &= S_n + (\lambda_{20} + \lambda_{21}) \tau_n r(t_n, S_n) + \\
 &\quad + \lambda_{21} \lambda_{20} \tau_n^2 \left[\frac{\partial r}{\partial t}(t_n, S_n) + \left(\frac{\partial r}{\partial s} r \right)(t_n, S_n) \right] + O(\tau_n^3),
 \end{aligned}$$

Uit de differentiaalvergelijking volgt nu dat

$$\begin{aligned}
 S_{n+1} &= s(t_n) + (\lambda_{20} + \lambda_{21}) \tau_n \frac{\partial s}{\partial t}(t_n) + \\
 &\quad + \lambda_{21} \lambda_{10} \tau_n^2 \frac{\partial^2 s}{\partial t^2}(t_n) + O(\tau_n^3),
 \end{aligned}$$

waarin $s(t)$ de oplossing van de differentiaalvergelijking is met $s(t_n) = S_n$. Voorwaarde voor 2^e orde nauwkeurigheid is blijkbaar

$$(4.5) \quad \lambda_{20} + \lambda_{21} = 1, \quad \lambda_{21} \lambda_{10} = \frac{1}{2}.$$

Op deze wijze vindt men voor een m-punts formule en een stelsel differentiaalvergelijkingen

$$(4.6) \quad \sum_{j=0}^{m-1} \lambda_{m,j} = 1, \quad \sum_{j=1}^{m-1} \lambda_{m,j} \sum_{\ell=0}^{j-1} \lambda_{j,\ell} = \frac{1}{2}.$$

4.3. Het stabiliteitspolynoom

Stel dat we het 2-puntsschema (4.3') ook toepassen op een naburige oplossing $\tilde{S}_n$; er geldt dan

$$\begin{aligned}
S_{n+1} - \tilde{S}_{n+1} &\cong S_n - \tilde{S}_n + \lambda_{20} \tau_n \frac{\partial r}{\partial s}(t_n, S_n) (S_n - \tilde{S}_n) \\
&\quad + \lambda_{21} \tau_n \frac{\partial r}{\partial s}(t_n + \mu_1 \tau_n, S_{n+1}^{(1)}) (S_{n+1}^{(1)} - \tilde{S}_{n+1}^{(1)}) \\
&\cong \{1 + \lambda_{20} \tau_n \frac{\partial r}{\partial s}(t_n, S_n) + \lambda_{21} \tau_n \frac{\partial r}{\partial s}(t_n + \mu_1 \tau_n, S_{n+1}^{(1)}) \\
&\quad + [1 + \lambda_{10} \tau_n \frac{\partial r}{\partial s}(t_n, S_n)]\} (S_n - \tilde{S}_n).
\end{aligned}$$

Wanneer nu $\partial r / \partial s$ een langzaam variërende functie is, kunnen we schrijven

$$(4.7) \quad S_{n+1} - \tilde{S}_{n+1} \cong R(\tau_n J_n) (S_n - \tilde{S}_n),$$

waarin

$$J_n = \frac{\partial r}{\partial s}(t_n, S_n),$$

$$R(z) = 1 + (\lambda_{20} + \lambda_{21})z + \lambda_{21}\lambda_{10}z^2.$$

Relatie (4.7) geldt algemeen voor m-punts Runge-Kutta formules en stelsels differentiaalvergelijkingen; de enige wijziging is dat R dan een polynoom van de graad m in $\tau_n J_n$ is, waarbij J_n de Jacobiaan van de rechterlidfunctie is in het punt (τ_n, S_n) . Het polynoom $R(z)$ wordt het *stabiliteitspolynoom* genoemd.

We geven zonder afleiding (zie hiervoor de syllabus MC 24.1) het stabiliteitspolynoom van het voor grote stelsels geschikte schema (beperkt geheugengebruik):

$$\begin{aligned}
\vec{S}_{n+1}^{(0)} &= \vec{S}_n, \\
(4.8) \quad \vec{S}_{n+1}^{(j)} &= \vec{S}_n + \lambda_{j,j-1} \tau_n \vec{r}(t_n + \mu_{j-1} \tau_n, \vec{S}_{n+1}^{(j-1)}), \quad j = 1, 2, \dots, m, \\
\vec{S}_{n+1} &= \vec{S}_{n+1}^{(m)};
\end{aligned}$$

er geldt

$$(4.9) \quad P(z) = 1 + \beta_1 z + \beta_2 z^2 + \dots + \beta_m z^m,$$

waarin de coëfficiënten functies zijn van de Runge-Kuttaparameters $\lambda_{j,l}$:

$$\begin{aligned} \beta_1 &= \lambda_{m,m-1}, \quad \beta_2 = \lambda_{m,m-1} \lambda_{m-1,m-2}, \quad \dots, \\ (4.10) \quad \beta_j &= \lambda_{m,m-1} \lambda_{m-1,m-2} \dots \lambda_{m-j+1,m-j}, \quad j = 3, \dots, m. \end{aligned}$$

Omgekeerd kunnen we de $\lambda_{j,j-1}$ in de β_j 's uitdrukken ($s_0=1$):

$$(4.10') \quad \lambda_{j,j-1} = \frac{\beta_{m+1-j}}{\beta_{m-j}}, \quad j = 1, \dots, m.$$

Uit (4.6) volgt dat voor 2^e orde nauwkeurigheid de coëfficiënten β_1 en β_2 de waarden 1 resp. $\frac{1}{2}$ moeten hebben; de overige β_j 's zijn vrij en kunnen gekozen worden om de stabiliteit zo groot mogelijk te maken.

4.4. Optimale formules voor differentiaalvergelijkingen met imaginair spectrum

We keren nu terug tot het stelsel differentiaalvergelijkingen (4.2). In paragraaf 3.1 hebben we de Jacobiaan, of juist een benadering voor de Jacobiaan, van dit stelsel onderzocht en laten zien dat deze een nagenoeg imaginair eigenwaarden-spectrum heeft. Uit (4.7) volgt dan dat we het stabiliteitspolynoom $R(z)$ over een zo groot mogelijk imaginair interval binnen de eenheidscirkel moeten proberen te houden. Immers, wanneer $|R(z)| < 1$ voor $z \in (-i\beta, i\beta)$, waarin β de *imaginaire stabiliteitsgrens* genoemd wordt, dan wordt de stabiliteitsvoorwaarde $|\tau_n \delta| < \beta$ voor alle eigenwaarden δ van de Jacobiaan, of wel

$$(4.11) \quad \tau_n < \frac{\beta}{\sigma(J_n)},$$

waarin $\sigma(J_n)$ de spectrale radius van J_n is. Voor stelsel (4.2) leidt dit tot (vergelijk (3.24)):

$$(4.11') \quad \tau_n < \min_{\vec{\gamma}} \frac{\beta}{\sqrt{g_H |\delta_x^2 + \delta_y^2|}}.$$

Het probleem om een polynoom van de vorm

$$R(z) = 1 + z + \frac{1}{2}z^2 + \beta_3 z^3 + \dots + \beta_m z^m$$

over een zo groot mogelijk imaginair interval binnen de eenheidscirkel te houden, is onderwerp van uitvoerig onderzoek geweest op het Mathematisch Centrum. Aanvankelijk lukte het slechts de lagere graadspolynomen (3 en 4) te vinden door een ad hoc constructie toe te passen met als resultaat

$$(4.12) \quad m = 3 : R(z) = 1 + z + \frac{1}{2}z^2 + \frac{1}{4}z^3 \quad \beta = 2$$

$$m = 4 : R(z) = 1 + z + \frac{1}{2}z^2 + \frac{1}{6}z^3 + \frac{1}{24}z^4, \quad \beta = 2\sqrt{2}.$$

De hiermee gegenereerde Runge-Kuttaformulas zien er als volgt uit:

$$(4.13) \quad m = 2 : \vec{S}_{n+1} = \vec{S}_n + \tau_n \vec{r}(t_n + \frac{1}{2}\tau_n, \vec{S}_n + \frac{1}{2}\tau_n \vec{r}(t_n, \vec{S}_n)) + \frac{1}{2}\tau_n \vec{r}(t_n + \frac{1}{2}\tau_n, \vec{S}_n + \frac{1}{2}\tau_n \vec{r}(t_n, \vec{S}_n)),$$

$$\tau_n < \frac{2}{\sigma(J_n)} \approx \min_{\gamma} \frac{2}{\sqrt{gH} |\delta_x^2 + \delta_y^2|};$$

$$(4.14) \quad m = 3 : \vec{S}_{n+1} = \vec{S}_n + \tau_n \vec{r}(t_n + \frac{1}{2}\tau_n, \vec{S}_n + \frac{1}{2}\tau_n \vec{r}(t_n + \frac{1}{3}\tau_n, \vec{S}_n + \frac{1}{3}\tau_n \vec{r}(t_n + \frac{1}{4}\tau_n, \vec{S}_n + \frac{1}{4}\tau_n \vec{r}(t_n, \vec{S}_n)))) + \frac{1}{3}\tau_n \vec{r}(t_n + \frac{1}{2}\tau_n, \vec{S}_n + \frac{1}{2}\tau_n \vec{r}(t_n + \frac{1}{3}\tau_n, \vec{S}_n + \frac{1}{3}\tau_n \vec{r}(t_n + \frac{1}{4}\tau_n, \vec{S}_n + \frac{1}{4}\tau_n \vec{r}(t_n, \vec{S}_n))))),$$

$$\tau_n < \frac{2\sqrt{2}}{\sigma(J_n)} \approx \min_{\gamma} \frac{2\sqrt{2}}{\sqrt{gH} |\delta_x^2 + \delta_y^2|}.$$

In de stabiliteitsvoorwaarden zijn ook de door (4.11') bepaalde maximale integratiestappen voor stelsel (4.2) toegevoegd. Vergelijken we deze maximale stappen met de voor het gesymmetriseerde schema afgeleide integratiestap in (3.24), dan zien we dat het drie-puntsschema dezelfde voorwaarde voor τ_n geeft en het vierpuntsschema een $\sqrt{2}$ grotere stap. Het gesymmetriseerde schema vraagt echter maar één rechterlideovaluatie per stap en is dus aanzienlijk goedkoper. Om de bewerkelijkheid te meten voeren we de *effectieve* integratiestap

$$(4.15) \quad \tau_{\text{eff}} = \frac{\tau_{\text{max}}}{m}$$

in, waarin m het aantal rechterlidevaluaties voorstelt; τ_{eff} is dus de maximale stap die men per rechterlidevaluatie mag nemen. We vinden nu het tabelletje

Schema	$\tau_{\text{eff}} \max_{\gamma} \sqrt{gH \delta_x^2 + \delta_y^2 }$	orde p
Gesymmetriseerd	2	1
3-punts Runge-Kutta	2/3	2
4-punts Runge-Kutta	$\sqrt{2}/2$	2

Hieruit valt in de eerste plaats af te leiden dat de hogere orde van nauwkeurigheid duur betaald moet worden en in de tweede plaats dat de 4-punts formule iets effectiever is dan de driepuntsformule. Dit werpt de vraag op of men de effectiviteit wezenlijk kan opvoeren door m groter te kiezen. We zullen een schets geven van de analyses die verricht zijn aan dit polynoomprobleem (voor details verwijzen we naar het MC-rapport MR 105).

De eerste stap is om het polynoom

$$(4.16) \quad Q_m(-z^2) \approx |R(z)|^2$$

op te stellen. Wanneer R van de graad m in z is, dan is Q_m van de graad m in $-z^2$; verder is het argument van Q_m positief als het argument van R zuiver imaginair is. Het op deze wijze gedefinieerde polynoom Q_m kan geschreven worden als

$$Q_m(z) = c_k^2(z) + z s_\ell^2(z),$$

$$(4.16') \quad c_k(z) = 1 - \frac{1}{2} z + \beta_4 z^2 - \dots + \beta_{2k} (-1)^k z^k,$$

$$s_\ell(z) = 1 - \beta_3 + \beta_5 z^2 - \dots + \beta_{2\ell+1} (-1)^\ell z^\ell,$$

waarin $k = \ell = (m-1)/2$ als m oneven is en $k = \ell + 1 = m/2$ als m even is.

Het probleem voor $Q_m(z)$ is nu om de coëfficiënten β_j zo te kiezen dat Q_m langs de positieve z -as zolang mogelijk kleiner dan 1 is. Vinden we voor de lengte van dit interval de waarde b dan is $\beta = \sqrt{b}$.

De tweede stap is de nog onbekende lengte van het interval $(0, b)$ in de nog onbekende coëfficiënten β_j te stoppen. Daartoe gaan we over op de variabele

$$\zeta = \frac{z}{b}.$$

Hiermee gaat (4.16') over in

$$\begin{aligned} Q_m(z) &= C_k^2(\zeta) + b \zeta S_\ell^2(\zeta), \\ (4.16'') \quad C_k(\zeta) &= 1 - \alpha_2 \zeta + \alpha_4 \zeta^2 - \dots + (-1)^k \alpha_{2k} \zeta^k, \\ S_\ell(\zeta) &= 1 - \alpha_3 \zeta + \alpha_5 \zeta^2 - \dots + (-1)^\ell \alpha_{2\ell+1} \zeta^\ell, \end{aligned}$$

waarin $\alpha_{2j} = \beta_{2j} b^j$ en $\alpha_{2j+1} = \beta_{2j+1} b^j$.

Uit $Q_m(z) \leq 1$ voor $0 \leq b$ volgt

$$b \leq \frac{1 - C_k^2(\zeta)}{\zeta S_\ell^2(\zeta)} \text{ voor } 0 \leq \zeta \leq 1$$

zodat

$$b \leq \max_{\alpha} \min_{0 \leq \zeta \leq 1} \frac{1 - C_k^2(\zeta)}{\zeta S_\ell^2(\zeta)}.$$

Stelling 4.1. Zijn $C_k(\zeta)$ en $S_\ell(\zeta)$ gegeven door (4.16'') dan is

$$\max_{\alpha} \min_{0 \leq \zeta \leq 1} \frac{1 - C_k^2(\zeta)}{\zeta S_\ell^2(\zeta)} \leq 4 k^2.$$

Bewijs. Zie MC-rapport MR 105.

Uit deze stelling volgt dat b nooit groter kan zijn dan $(m-1)^2$ als m oneven is en m^2 als m even is. Derhalve geldt

$$(4.17) \quad \beta \leq \begin{matrix} m-1 & \text{als } m \text{ oneven} \\ m & \text{als } m \text{ even} \end{matrix},$$

zodat

$$(4.18) \quad \tau_{\text{eff}} \sigma(J_n) \leq \begin{matrix} \frac{m-1}{m} & \text{als } m \text{ oneven} \\ 1 & \text{als } m \text{ even} \end{matrix}.$$

De vraag of τ_{eff} wezenlijk groter gemaakt kan worden door meer punten mee te nemen in het Runge-Kuttaschema moet met "enigszins" beantwoord worden.

Overigens is het nog de vraag of er polynomen te vinden zijn die de door stelling 4.1 gegeven bovengrens voor b halen. Het is inderdaad gelukt om een gesloten uitdrukking voor $Q_m(z)$, en dus voor $R(z)$, te vinden met $b =$ bovengrens voor alle oneven waarden van m . Zonder afleiding geven we $R(z)$ in de vorm

$$(4.19) \quad R(z) = T_k\left(1 + \frac{z^2}{2k}\right) + \frac{z}{k}\left(1 + \frac{z^2}{4k}\right) U_{k-1}\left(1 + \frac{z^2}{2k}\right),$$

waarin $k = \frac{m-1}{2}$ en waarin T_k en U_{k-1} Chebyshev polynomen van de eerste en tweede soort zijn.

Voor even waarden van m is de bovengrens niet haalbaar.

4.5. Gegeneraliseerde Runge-Kuttamethoden

We zullen formules van de norm (4.3) waarin de *scalaire* parameters $\lambda_{j,l}$ vervangen zijn door *matrices* $\Lambda_{j,l}$ van de orde m *gegeneraliseerde Runge-Kuttaformules* noemen. Het bijzondere geval waarin de $\Lambda_{j,l}$ polynomen of rationale functies zijn van $\tau_n J_n$ is uitvoerig op het Mathematisch Centrum onderzocht (zie e.g. MC. Syllabus 24.1). Onlangs is een onderzoek begonnen naar het effect van algemenere matrix-operatoren in het bijzonder voor het geval waarin het spectrum van Jacobiaan van de differentiaalvergelijking zuiver imaginair is. We volstaan hier met een eenvoudig voorbeeld om te laten zien dat het loont dergelijke algemene Runge-Kutta formules te onderzoeken. Beschouw het stelsel differentiaalvergelijkingen waarin $\vec{r}(\vec{s})$ van de volgende norm is:

$$(4.20) \quad \vec{s} = \begin{pmatrix} \vec{u} \\ \vec{z} \end{pmatrix}, \quad \vec{r}(\vec{s}) = \begin{pmatrix} \vec{z} \\ \vec{f}(\vec{u}) \end{pmatrix},$$

en waarin de Jacobiaan J_f van de functie $\vec{f}$ negatieve eigenwaarden δ_f heeft. Dit voorbeeld ontstaat uit de reductie tot eerste orde vorm van de differentiaalvergelijking $d^2\vec{u}/dt^2 = \vec{f}(\vec{u})$. De Jacobiaan van de functie $\vec{r}$ is van de form

$$J = \begin{pmatrix} 0 & I \\ J_f & 0 \end{pmatrix}$$

met de eigenwaarden δ die voldoen aan

$$\delta^2 - \delta_f = 0;$$

dus J heeft imaginaire eigenwaarden

$$\delta = \pm i\sqrt{-\delta_f}.$$

Zou men hier de 2-puntsformule (4.13) op toepassen dan vinden we de stabiliteitsvoorwaarde

$$(4.21) \quad \tau_n < \frac{2}{\sigma(J_n)} = \frac{2}{\sqrt{\sigma((J_f)_n)}}.$$

Past men echter een gegeneraliseerde methode met

$$(4.22) \quad \Lambda_{10} = \begin{pmatrix} \frac{1}{2}I & 0 \\ 0 & 0 \end{pmatrix}, \quad \Lambda_{2,1} = \begin{pmatrix} \frac{1}{2}I & \frac{1}{16}\tau I \\ 0 & 0 \end{pmatrix}, \quad \Lambda_{3,2} = \begin{pmatrix} I & \frac{1}{2}\tau I \\ 0 & I \end{pmatrix}$$

toe, ofwel uitgeschreven in $\vec{u}$, $\vec{z}$ en $\vec{f}(u)$ (substitutie van (4.22) voor $\lambda_{j,l}$ en (4.20) voor $\vec{r}$ in formule (4.8))

$$(4.23) \quad \begin{aligned} \vec{u}_{n+1} &= \vec{u}_n + \tau_n \vec{z}_n + \frac{1}{2} \tau_n^2 \vec{f}(\vec{u}_n + \frac{1}{2} \tau_n \vec{z}_n + \frac{1}{16} \tau_n^2 \vec{f}(\vec{u}_n + \frac{1}{2} \tau_n \vec{z}_n)), \\ \vec{z}_{n+1} &= 2 \frac{\vec{u}_{n+1} - \vec{u}_n}{\tau_n} - \vec{z}_n, \end{aligned}$$

dan levert dit een per integratiestap *minder* bewerkelijke formule (slechts 2 functie-evaluaties), terwijl de stabiliteitsvoorwaarde

$$(4.24) \quad \tau_n \leq \frac{4}{\sqrt{\sigma((J_f)_n)}}$$

tweemaal grotere stappen toestaat (vergelijk (4.22)). Aangezien formule (4.23) ook van de orde 2 is, zien we dat een geweldige winst geboekt kan worden met gegeneraliseerde Runge-Kuttamethoden. Voor verdere details verwijzen we naar het MC-rapport NW 26/76.

V. MEERSTAPSMETHODEN

Meerstapsmethoden, en zeker methoden met meer dan twee stappen (of wel drie niveaux), worden weinig gebruikt voor partiele differentiaalvergelijkingen in meer dan één plaatsvariabele. De reden is de grote geheugenbe-zetting bij implementatie op een rekenmachine. Mogelijk is dit standpunt langzamerhand achterhaald in het licht van de steeds sneller wordende achtergrond geheugens. Hoe dit ook zij, in de literatuur vindt men voor het twee-dimensionale ondiep-waterprobleem geen numerieke algoritmen met meer dan 2 stappen (3 niveaux). We zullen ons in deze syllabus dan ook beperken tot de constructie van enkele 2-stapsmethoden voor het ondiep-waterprobleem; mogelijk geeft dit enige indicatie voor de constructie van efficiëntere formules gebaseerd op 3 of meer stappen.

5.1. Het "leap-frog" schema

We gaan weer uit van de algemene eerste orde vergelijking

$$(5.1) \quad \frac{d\vec{s}}{dt} = \vec{r}(t, \vec{s}),$$

en we definiëren het algemene lineaire expliciete twee-stapschema

$$(5.2) \quad \vec{s}_{n+1} = a_1 \vec{s}_n + a_2 \vec{s}_{n-1} + b_1 \tau \vec{r}(t_n, \vec{s}_n) + b_2 \tau \vec{r}(t_{n-1}, \vec{s}_n).$$

Hierin zijn a_1 , a_2 , b_1 en b_2 nog nader te bepalen scalaire parameters. Onder andere zullen deze parameters zodanig moeten zijn dat een oplossing $\vec{s}$ van (5.1) op een kleine fout (afbreekfout) na aan het schema (5.2) voldoet. Substitutie van zo'n oplossing $\vec{s}$ geeft ($\vec{s}$ wordt voldoende vaak differentieerbaar naar t verondersteld) het residu

$$(5.3) \quad (1-a_1-a_2) \vec{s} + \tau (1+a_2-b_1-b_2) \frac{d\vec{s}}{dt} + \frac{1}{2}\tau^2 (1-a_2+2b_2) \frac{d^2\vec{s}}{dt^2} + O(\tau^3),$$

waarin $\vec{s}$ en zijn afgeleiden in $t = t_n$ geëvalueerd verondersteld worden te zijn. Hieruit volgt dat de afbreekfout van de orde 2 in τ is als

$$(5.4) \quad a_1 + a_2 = 1, \quad a_2 = b_1 + b_2 - 1;$$

de methode is dan (per definitie) *eerste orde consistent*. Geldt bovendien

$$(5.5) \quad a_2 = 2b_2 + 1,$$

dan is de afbreekfout van de 3^e orde en de methode tweede orde consistent.

Vervolgens gaan we na aan welke differentiaalvergelijking de oplossing $\vec{S}_n$ van schema (5.2) als functie van τ voldoet. Ontwikkeling in Taylorreeksen geeft de vergelijking

$$(5.6) \quad (1-a_1-a_2) \vec{S} + \tau \left[(1+a_2) \frac{d\vec{S}}{dt} - (b_1+b_2) \vec{r}(t, \vec{S}) \right] + O(\tau^2) = 0.$$

We zien hieruit dat schema (10.2) alleen dan de differentiaalvergelijking (10.1) benadert wanneer aan (10.4) voldaan is en wanneer bovendien

$$(5.7) \quad b_1 + b_2 \neq 0.$$

Substitutie van (5.4) in (5.2) geeft het schema

$$(5.2') \quad \begin{aligned} \vec{S}_{n+1} = a_1 \vec{S}_n + (1-a_1) \vec{S}_{n-1} + b_1 \tau \vec{r}(t_n, \vec{S}_n) + \\ + (2-a_1-b_1) \tau \vec{r}(t_{n-1}, \vec{S}_{n-1}), \end{aligned}$$

waarin krachtens (5.7) de parameter a_1 niet gelijk aan 2 mag zijn; verder zijn a_1 en ook b_1 vrij te kiezen parameters.

De stabiliteit van schema (5.2') onderzoeken we, zoals gebruikelijk, aan de hand van de variatie-vergelijking bij (5.2'). Daartoe beschouwen we verstoringen $\tilde{S}_n$ en $\tilde{S}_{n-1}$ van $\vec{S}_n$ en $\vec{S}_{n-1}$, en trekken de resulterende vector $\tilde{S}_{n+1}$ van $\vec{S}_{n+1}$ af. Noteren we kortheidshalve

$$\Delta \vec{S}_n = \vec{S}_n - \tilde{S}_n, \quad n = 0, 1, 2, \dots,$$

dan geldt

$$(5.8) \quad \Delta \vec{S}_{n+1} = a_1 \Delta \vec{S}_n + (1-a_1) \Delta \vec{S}_{n-1} + b_1 \tau \left[\vec{r}(t_n, \vec{S}_n) - \vec{r}(t_n, \tilde{S}_n) \right] +$$

$$\begin{aligned}
& + (2-a_1-b_1) \tau \left[\vec{r}(t_{n-1}, \vec{s}_{n-1}) - \vec{r}(t_{n-1}, \vec{s}_{n-1}^*) \right] \\
& \cong \left[a_1 + b_1 \tau J(t_n, \vec{s}_n) \right] \Delta \vec{s}_n + \left[1 - a_1 + \right. \\
& \left. + (2-a_1-b_1) \tau J(t_{n-1}, \vec{s}_{n-1}) \right] \Delta \vec{s}_{n-1},
\end{aligned}$$

waarin J weer de Jacobiaan van $\vec{r}$ voorstelt. We schrijven (1.8) in de norm

$$(5.8') \quad \begin{pmatrix} \Delta \vec{s}_{n+1} \\ \Delta \vec{s}_n \end{pmatrix} = A_n \begin{pmatrix} \Delta \vec{s}_n \\ \Delta \vec{s}_{n-1} \end{pmatrix},$$

waarin de matrix A_n gegeven wordt door

$$A_n = \begin{pmatrix} a_1 + b_1 \tau J(t_n, \vec{s}_n) & (1-a_1) + (2-a_1-b_1) \tau J(t_{n-1}, \vec{s}_{n-1}) \\ I & 0 \end{pmatrix}$$

We zullen A_n de amplificatiematrix van schema (5.2') noemen en van stabiliteit spreken wanneer de spectrale norm van A_n kleiner dan 1 is (vergelijk paragraaf 2.2). In de praktijk stelt men zich echter ook hier weer tevreden met de voorwaarde dat de eigenwaarden van A_n binnen de eenheidscirkel liggen (*sterke stabiliteit*) of desnoods op de eenheidscirkel (*zwakke stabiliteit*). In het volgende nemen we aan dat de Jacobiaan $J(t_n, \vec{s}_n)$ zo langzaam met t_n en $\vec{s}_n$ varieert dat de eigenwaarden van $J(t_n, \vec{s}_n)$ en $J(t_{n-1}, \vec{s}_{n-1})$ gelijk verondersteld mogen worden; deze eigenwaarden noemen we δ . De eigenwaarden α van de amplificatiematrix A_n kunnen we nu uitdrukken in δ ; er geldt namelijk

$$(5.10) \quad \alpha^2 - (a_1 + b_1 \tau \delta) \alpha - (1-a_1) - (2-a_1-b_1) \tau \delta = 0,$$

waaruit $\alpha_1(\tau\delta)$ en $\alpha_2(\tau\delta)$ af te leiden zijn. De twee functies $\alpha_{1,2}(z)$ zijn de analoge van het stabiliteitspolynoom bij de éénstapsmethoden (vergelijk paragraaf 4.3).

Evenals in paragraaf 4.4 willen we de eigenwaarden van de amplificatiematrix over een zo groot mogelijke imaginair interval van $\tau\delta$ -waarden binnen de eenheidscirkel houden. Daartoe bepalen we in het complex z -vlak het gebied (zogenaamd stabiliteitsgebied)

$$\{z \mid |\alpha_{1,2}(z)| < 1\}.$$

De rand van dit gebied wordt gegeven door de kromme $z = z(\phi)$, $0 \leq \phi \leq 2\pi$, gedefinieerd door substitutie van $\alpha = \exp(i\phi)$ in (5.10), dus

$$(5.11) \quad z = \frac{1 - a_1 + a_1 e^{i\phi} - e^{2i\phi}}{a_1 + b_1 - 2 - b_1 e^{i\phi}}.$$

We zullen hier het bijzondere (en eenvoudige) geval

$$(5.12) \quad a_1 + b_1 - 2 = 0$$

verder uitwerken. Analyse van het algemene geval laten we over aan de geïnteresseerde lezer. Substitutie van (5.12) in (5.11) leidt tot

$$(5.11') \quad z = a_1 \frac{-1 + \cos\phi + i\left(\frac{2}{a_1} - 1\right) \sin\phi}{2 - a_1}.$$

Deze kromme stelt een ellips voor van de vorm zoals aangegeven in figuur 5.1.

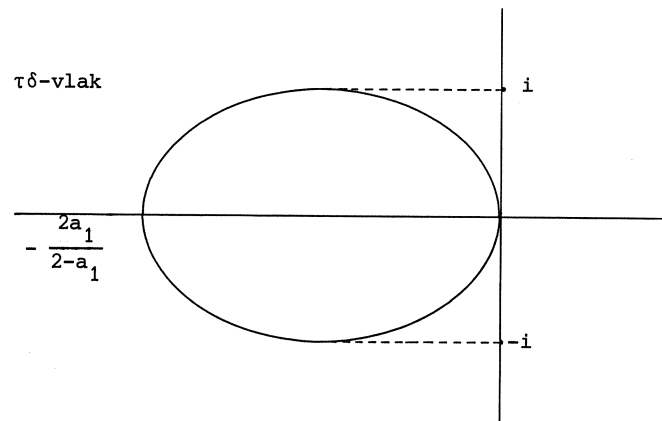


Fig. 5.1. Stabiliteitsgebied van formule (5.2') voor $a_1 + b_1 - 2 = 0$.

Wil men schema (5.2') stabiel maken voor vergelijkingen met imaginaire eigenwaarden (de ondiep-watervergelijkingen) dan moet men blijkbaar

$$(5.13) \quad a_1 = 0$$

kiezen. Het door (5.2'), (5.12) en (5.13) bepaalde schema neemt de norm

$$(5.2'') \quad \vec{S}_{n+1} = \vec{S}_{n-1} + 2 \tau \vec{r}(t_n, \vec{S}_n)$$

aan. Voor imaginaire eigenwaarden geldt de stabiliteitsvoorwaarde (vergelijk (3.24) en (4.13))

$$(5.14) \quad \tau \leq \frac{1}{\sigma(J(t_n, \vec{S}_n))} \approx \min_{\vec{\gamma}} \frac{1}{\sqrt{g_H |\delta_x^2 + \delta_y^2|}} ;$$

de eigenwaarden van A_n liggen dan juist op de eenheidscirkel. Het schema is 2e orde consistent omdat de keuze (5.12) en (5.13) impliceert dat $a_2 = 1$ en $b_2 = 0$ zodat aan (5.5) voldaan is. Een groot nadeel is echter dat voor eigenwaarden δ met een klein negatief reëel deel (bodempwrijving!) instabiliteit optreedt. Het is interessant om te onderzoeken of een andere keuze dan (5.12) het stabiliteitsgebied kan verbeteren.

Schema (5.2'') wordt wel het "leap-frog" schema ("haasje-over" schema) of "step-over" schema (RICHARDSON [1922]) genoemd. Het werd door HARRIS en JELESNIANSKI [1964] op de ondiep-water vergelijkingen toegepast waarbij de bodempwrijving weggelaten werd; als reden werd opgegeven dat wanneer het schema *zonder* bodempwrijving stabiel is, het zeker stabiel zou zijn *met* bodempwrijving. Voorgaande analyse wijst "op papier" anders uit. Ook dit zou nader onderzocht kunnen worden.

VI. DISSIPATIEVE TERMEN EN ANDERE STABILISERENDE FACTOREN

De stabiliteitsvoorwaarden die we voor de verschillende formules afgeleid hebben zijn gebaseerd op een *lineaire, lokale* theorie en garanderen dus geen overall-stabiliteit in een concreet rekenschema. Er zijn diverse voorstellen in de literatuur gedaan om de ontwikkeling van instabiliteiten die in weerwil van de lineaire theorie, tóch in een daadwerkelijk rekenschema

optreden, de kop in te drukken. We zullen achtereenvolgens bespreken: *effeningsoperatoren (smoothing operators)*, *artificiele wrijving* en *artificiele viscositeitstermen*. Deze stabiliserende technieken zijn algemener toepasbaar en worden vaak bij hyperbolische systemen gebruikt.

6.1. Effeningsoperatoren, de Shuman-operatoren

In paragraaf 2.8 (zie figuur 2.4) hebben we al opgemerkt dat de kleinste golflengte van de exponentiele componenten $\vec{a} \exp [i(\vec{\gamma}, \vec{x})]$ waaruit de verstoringen opgebouwd zijn, op een rooster met vierkante mazen van ξ bij ξ de waarde 2ξ heeft. (Hogere frequenties kan zo'n rooster niet meer representeren.) Het mag bekend verondersteld worden dat afrondingsfouten zich manifesteren in de hogere frequenties; een instabiel rekenschema blaast verstoringen op en zal resultaten afleveren waarin juist de *hoog-frequente* componenten onevenredig sterk vertegenwoordigd zijn. We zouden dus een techniek moeten hebben om uit het afgeleverde stromings- en verhogingsveld de hogere frequenties te filteren. Een dergelijke techniek is door SHUMAN [1957] voorgesteld. Voor een roosterfunctie f in één roostervariabele $x_j = j\xi$ stelt hij de effeningsoperator

$$(6.1) \quad E_x f(x_j) = \frac{1}{4} [f(x_{j-1}) + 2f(x_j) + f(x_{j+1})]$$

voor. Dit schrijft men ook wel in "molecuul-notatie" als

$$(6.1') \quad E_x f = \frac{1}{4} [1 \ 2 \ 1] f,$$

waarbij de gewichten van de f -waarden uit formule (6.1) in een array gezet worden op met de roosterpunten corresponderende plaatsen. Voor een tweedimensionaal veld f in de roostervariabelen $x_j = j\xi$ en $y_\ell = \ell\eta$ krijgen we door toepassing van (6.1) in x - en y -richting:

$$(6.2) \quad E f = E_y E_x f = \frac{1}{16} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \cdot [1 \ 2 \ 1] f = \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} f.$$

Intuitief heeft men inderdaad het gevoel dat dit soort "middelingsoperatoren" de hoogfrequente golfjes uit het veld $f(x_j, y_\ell)$ filtert (zie figuur 6.1).

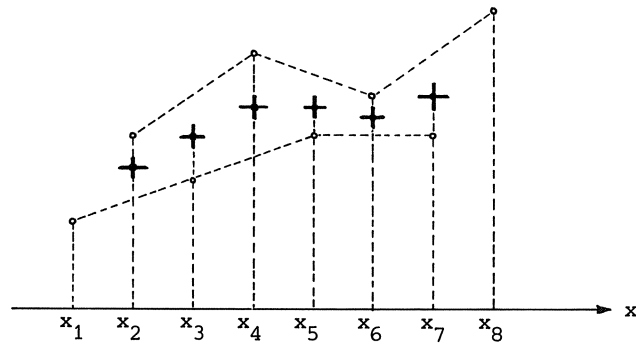


Fig. 6.1. $\circ$ numerieke resultaten
 $+$ resultaten na toepassing van (6.1).

We zullen nu wat quantitatiever het effect van effeningsoperatoren op de hogere harmonischen nagaan. Daarbij gaan we uit van een wat algemenere operator

$$(6.3) \quad E = \frac{1}{c+4(a+b)} \begin{pmatrix} a & b & a \\ b & c & b \\ a & b & a \end{pmatrix},$$

waarin a , b en c nog vrij zijn. Passen we E toe op de component

$$\exp [i(\vec{\gamma} \cdot \vec{x})] = \exp [i(\gamma_1 j \xi + \gamma_2 l \eta)]$$

dan zien we dat deze functie een eigenfunctie van E is met eigenwaarde

$$(6.4) \quad \mu(\vec{\gamma}) = \frac{c+4a \cos \gamma_1 \xi \cos \gamma_2 \eta + 2b(\cos \gamma_1 \xi + \cos \gamma_2 \eta)}{c + 4a + 4b}$$

De hoogste frequentie op dit rooster (golflengte 2ξ en 2η in x - en y -richting) heeft de $\vec{\gamma}$ -index:

$$\gamma_1 = \frac{2\pi}{2\xi}, \quad \gamma_2 = \frac{2\pi}{2\eta}.$$

Uit (6.4) volgt

$$(6.4') \quad \mu\left(\frac{\pi}{\xi}, \frac{\pi}{\eta}\right) = \frac{c + 4a - 4b}{c + 4a + 4b}.$$

De hoogste frequentie wordt door E volledig geëlimineerd wanneer

$$(6.5) \quad c + 4a - 4b = 0.$$

Substitutie in (6.4) geeft

$$(6.6) \quad \mu(\vec{\gamma}) = \frac{2b - 2a + 2a pq + bp + bq}{4b},$$

waarin we $\cos \gamma_1 \xi = p$ en $\cos \gamma_2 \eta = q$ gesteld hebben. In figuur 6.2 is het oppervlak $\mu(\vec{\gamma})$ als functie van p en q geschetst. Voor p en q constant is μ een lineaire functie van q respectievelijk p ; het maximum van $|\mu|$ wordt dus op de randen $p = \pm 1$ en $q = \pm 1$ aangenomen. Hieruit volgt dat μ optimaal is wanneer μ nul is voor $p = -1$ en $q = -1$, i.e. $b = 2a$ (zie figuur 6.2). Deze relatie gecombineerd met (6.5) levert juist de operator van Shuman. Deze

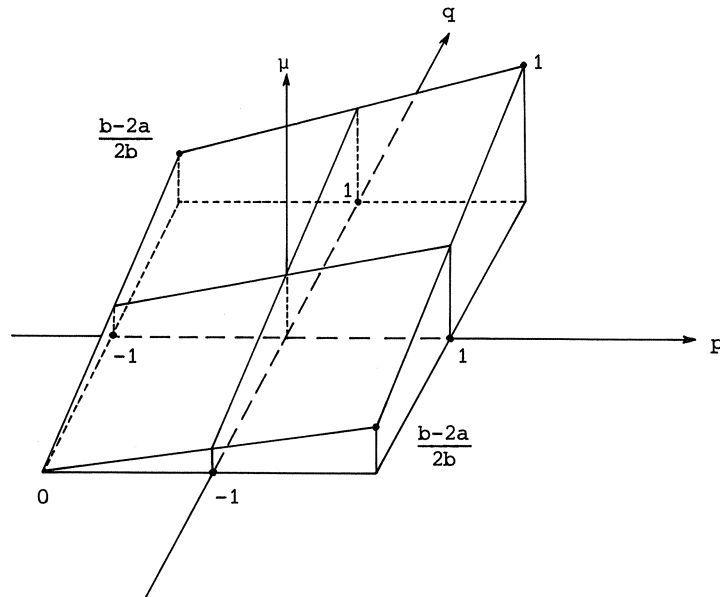


Fig. 6.2. Het oppervlak $\mu(\gamma)$ als functie van $p = \cos \gamma_1 \xi$ en $q = \cos \gamma_2 \eta$.

operator dempt het sterkst in het 2^e, 3^e en 4^e quadrant van het (p,q)-vlak en dempt weinig in het 1^e quadrant. Men kan zich afvragen of het rekenschema door het toepassen van (6.2) niet *overstabil* wordt omdat *alle* componenten behalve die met $\vec{\gamma} = \vec{0}$ gedempt worden. Aangezien echter voor ξ en $\eta \rightarrow 0$ de damping relatief minder wordt, dat wil zeggen de lagere harmonischen waaruit $\vec{u}$, $\vec{v}$ en h hoofdzakelijk opgebouwd zijn, worden steeds minder sterk gedempt, verstoort de operator E de consistentie niet.

Bovenstaande analyse betreft effeningsoperatoren op *uniforme* roosters. Voor *niet-uniforme* roosters zou men de Shuman-operator als volgt kunnen definiëren: Stel dat de roosterpunten op "gebroken" roosterlijnen liggen zoals geschetst in figuur 6.3. Interpreteren we deze roosterlijnen als "horizontale" en "verticale" coördinaatlijnen, dan zou men achtereenvolgens langs deze coördinaatlijnen kunnen middelen en aldus een effeningsoperator op niet-uniforme roosters krijgen. Geven we deze middelingsoperatoren met E_x respectievelijk E_y aan, dan geldt (zie figuur 6.4) in het punt P.

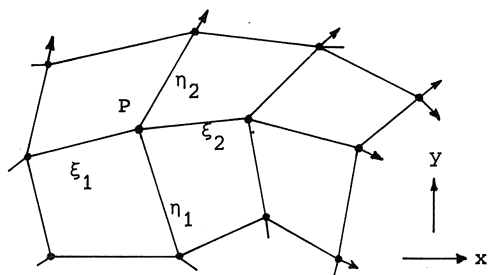


Fig. 6.3. Niet uniform rooster.

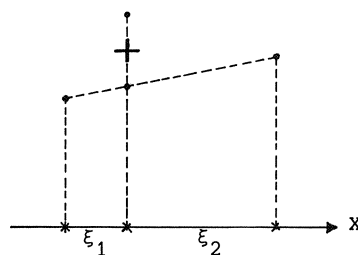


Fig. 6.4. * numerieke
resultaten
+ resultaten na
toepassing van
(6.1").

$$(6.1'') \quad E_X f = \frac{1}{4\bar{\xi}} [\xi_2 \quad 2\bar{\xi} \quad \xi_1] f, \quad \bar{\xi} = \frac{1}{2}(\xi_1 + \xi_2)$$

en iets dergelijks voor E_Y . Op deze wijze kunnen we de operatoren $E_X E_Y$ en $E_Y E_X$ definiëren; het is eenvoudig te verifiëren dat deze productoperatoren *niet* gelijk zijn, tenzij het rooster uniform is. Dit suggereert om op niet-uniforme roosters de Shuman-operator te definiëren als

$$(6.2') \quad E f = \frac{1}{2} (E_X E_Y + E_Y E_X) f.$$

Evenals het geval is voor uniforme roosters verstoort deze operator de consistentie niet, alhoewel de afbreekfout een orde lager is dan voor uniforme roosters ($O(\bar{\xi}) + O(\bar{\eta})$) in plaats van $O(\xi^2) + O(\eta^2)$. Voor niet te grote afwijkingen van de uniformiteit zijn de orde-constanten echter erg klein.

We besluiten de discussie van effeningsoperatoren met de beschrijving van een nog niet uitgetroefde strategie bij de toepassing van effeningsoperatoren. Stel dat we in de successieve integratiestappen *verschillende* operatoren E_n gebruiken en wel zodanig dat de bijbehorende eigenwaardenspectra $\mu_n(\vec{\gamma})$ in het (p, q) -vlak zijn nulpunten heeft op lijnen (uniforme roosters)

$$(6.7) \quad 2b_n - 2a_n + 2a_n p q + b_n p + b_n q = 0$$

zoals geschetst in figuur 6.5. We willen deze lijnen zo kiezen dat $\prod_n |\mu_n(\vec{\gamma})|$

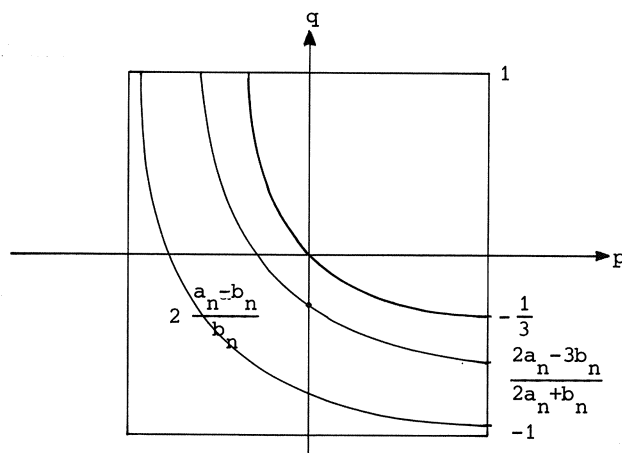


Fig. 6.5. Lijnen waarop $\mu_n(\vec{\gamma})$ nul is.

geminimaliseerd wordt. Omdat het maximum van $|\mu_n(\vec{\gamma})|$ langs de randen wordt aangenomen geldt dit ook voor het maximum van $\prod_n |\mu_n(\vec{\gamma})|$. Zij nu $\vec{a} \exp [i(\vec{\gamma} \cdot \vec{x})]$ een component van de een of andere verstoring $\Delta \vec{S}_n$, dan geldt volgens de locale theorie dat deze in een roosterpunt $\vec{x}$ met de eigenwaarde $\alpha_n(\vec{x}, \vec{\gamma})$ wordt vermenigvuldigd. Passen we vervolgens de operator E_n toe dan wordt de versterkingsfactor $\mu_n(\vec{\gamma}) \propto \alpha_n(\vec{x}, \vec{\gamma})$. De versterkingsfactor na v integratiestappen wordt dan

$$\prod_{j=0}^{v-1} \mu_{n+j}(\vec{\gamma}) \propto \alpha_{n+j}(\vec{x}, \vec{\gamma})$$

Deze factor kan geminimaliseerd worden door $\prod_{j=0}^{v-1} |\mu_{n+j}(\vec{\gamma})|$ op de randen in het (p,q) -vlak te minimaliseren. Aangezien we in het bijzonder de hogere frequenties willen dempen doen we dit op de randstukken

$$q = 1, -1 \leq p \leq 0 \text{ en } p = -1, -1 \leq q \leq 1.$$

Dit betekent dat de nulpunten van $\mu_{n+j}(\vec{\gamma})$ voor $q = 1$ tussen $p = -1$ en $p = 0$ moeten liggen. Verder geldt dat het maximum van $|\mu_{n+j}(\vec{\gamma})|$ voor $p = -1$ altijd kleiner of gelijk is aan de waarde van $|\mu_{n+j}|$ in $(p,q) = (-1,1)$, zodat we het probleem gereduceerd hebben tot het volgende minimax-probleem:

Minimaliseer de maximum norm van het v -de graads polynoom

$$(6.8) \quad \prod_{j=0}^{v-1} [A_j + (1-A_j)p], \quad A_j = \frac{3b_{n+j} - 2a_{n+j}}{4b_{n+j}}$$

op het interval $-1 \leq p \leq 0$. De oplossing van dit probleem wordt gegeven door de nulpunten van het polynoom te identificeren met de nulpunten van het naar het interval $[-1,0]$ getransformeerde Chebyshev-polynoom van de graad v . Deze identificatie levert

$$(6.9) \quad \prod_{j=0}^{v-1} [A_j + (1-A_j)p] = \frac{T_v(1+2p)}{T_v(3)}$$

waaruit direct volgt

$$(6.10) \quad \frac{A_j}{A_{j-1}} = -\frac{1}{2} + \frac{1}{2} \cos\left(\frac{2j+1}{v}\pi\right).$$

Voor a_{n+j} vinden we tenslotte

$$(6.10') \quad a_{n+j} = b_{n+j} \frac{5 + \cos\left(\frac{2j+1}{v}\pi\right)}{6 - 2\cos\left(\frac{2j+1}{v}\pi\right)}, \quad j = 0, 1, \dots, v-1.$$

Substitutie van (6.5) en (6.10') in (6.3) met $b_{n+j} = 6 - 2\cos\left(\frac{2j+1}{v}\pi\right)$ levert de rij effenings-operatoren:

$$(6.11) \quad E_{n+j} = \frac{1}{16(3-c_j)} \begin{vmatrix} 5 + c_j & 2(3-c_j) & 5 + c_j \\ 2(3-c_j) & 4(1-3c_j) & 2(3-c_j) \\ 5 + c_j & 2(3-c_j) & 5 + c_j \end{vmatrix}, \quad j = 0, 1, \dots, v-1,$$

waarin $c_j = \cos\left(\frac{2j+1}{v}\pi\right)$. Deze rij operatoren zorgt er voor dat alle harmonischen met (p,q) -waarden in het 2^e , 3^e en 4^e quadrant gedempt worden met *minstens* een factor

$$(6.12) \quad T_v^{-1}(3) = [\cosh(v \operatorname{arcosh}(3))]^{-1} \approx [\cosh\left(\frac{7}{4}v\right)]^{-1} \\ \approx 2 \exp\left(-\frac{7}{4}v\right) \approx 2(.175)^v$$

Ter vergelijking berekenen we de dempingsfactor van de Shuman-operator (6.2) na v integratie-stappen. Uit figuur 6.2 zien we direct dat voor $b = 2a$ (dit levert de Shuman-operator) het maximum in het 2^e , 3^e en 4^e quadrant in $(p,q) = (0,1)$ aangenomen wordt en wel de waarde $\frac{1}{2}$. Dus na v integratie-stappen $(.5)^v$. Hieruit volgt dat voor $v = 5$ de demping al bijna $2\frac{1}{2}$ maal minder is dan de hierboven beschreven strategie. Voor grotere waarden van v loopt dit op tot bijna een factor 3.

Tenslotte merken we nog op dat effeningsoperatoren niet noodzakelijk voor het gehele stromingsveld of verhogingsveld toegepast hoeven te worden. Men kan volstaan met het toepassen in die gebieden van de beschouwde zee waar instabiliteiten makkelijk ontwikkeld worden, bijvoorbeeld aan de kunsten en in hoekpunten (zoals de ervaring geleerd heeft).

6.2. Dissipatieve termen, het Lax-Wendroff-schema

Eenvoudigheidshalve beperken we ons eerst tot de vergelijking

$$(6.13) \quad \frac{\partial \vec{u}}{\partial t} + \mathcal{D} \frac{\partial \vec{u}}{\partial x} = 0, \quad \mathcal{D} \text{ matrix met reële eigenwaarden.}$$

Het meest eenvoudige differentieschema hiervoor heeft de vorm (*Eulerschema* en *centrale plaatsdifferenties*)

$$(6.14) \quad \frac{\vec{U}(x, t+\tau) - \vec{U}(x, t)}{\tau} = - \frac{\mathcal{D}}{2\xi} (\vec{U}(x+\xi, t) - \vec{U}(x-\xi, t)),$$

We zullen nu nagaan aan welke differentiaalvergelijking de functie $\vec{U}(x, t)$ bepaald door (6.14), een oplossing is; daartoe veronderstellen we ξ en τ constant en laten we x alle punten van de x -as doorlopen en t de t -as. Ontwikkeling in een Taylorreeks in het punt (x, t) geeft de differentiaalvergelijking (tot en met de lagere orde termen)

$$(6.15) \quad \frac{\partial \vec{U}}{\partial t} + \frac{1}{2} \tau \frac{\partial^2 \vec{U}}{\partial t^2} = - \mathcal{D} \frac{\partial \vec{U}}{\partial x} - \frac{1}{6} \xi^2 \mathcal{D} \frac{\partial^3 \vec{U}}{\partial x^3} + O(\tau^2) + O(\xi^3).$$

Voor $\xi \rightarrow 0$ en $\tau \rightarrow 0$ voldoet de differentie-oplossing dus inderdaad aan vergelijking (6.13). Voor ξ en τ groter dan 0 voldoet $\vec{U}$ echter aan een verstoorte vergelijking. Men kan zich nu afvragen of deze stoortermen *dissipatief* zijn (dat wil zeggen een beginfunctie van de vorm $\exp(i\gamma x)$ wordt door vergelijking (6.15) voor elke γ uitgedempt). Zo ja, dan zal dit de stabiliteit van het differentieschema ten goede komen. Zo niet, dan mag men nauwelijks verwachten dat het differentieschema een gunstig stabiliteitsgedrag heeft. Dit suggereert om het differentieschema zodanig te kiezen dat het equivalent is met de beschouwde differentiaalvergelijking plus een aantal dissipatieve termen die naar 0 gaan als ξ en τ naar 0 gaan. Het loont niet de moeite om bovenstaande vergelijking op dissipativiteit te onderzoeken; we weten immers al dat het Eulerschema (6.14) voor differentiaalvergelijkingen waarvan de Jacobiaan imaginaire eigenwaarden heeft instabiel is (zie paragraaf 3.1).

In de praktijk is een mogelijke aanpak om uit te gaan van de gegeven differentiaalvergelijking en hieraan dissipatieve termen toe te voegen. Bijvoorbeeld in het geval van (6.13) kunnen we als uitgangsvergelijking kiezen:

$$(6.16) \quad \frac{\partial \vec{u}}{\partial t} = -D \frac{\partial \vec{u}}{\partial x} + \epsilon D^2 \frac{\partial^2 \vec{u}}{\partial x^2},$$

waarin ϵ een positieve parameter is die bij discretisering van deze vergelijking naar nul verondersteld wordt te gaan als de roosterparameters naar nul gaan. Het is duidelijk dat de diffusie-term $\epsilon D^2 \partial^2 \vec{u} / \partial x^2$ dissipatief is. Discretisering van (6.16) volgens Euler met centrale plaatsdifferenties geeft het schema (in verkorte notatie)

$$(6.17) \quad \frac{\vec{u}_{n+1} - \vec{u}_n}{\tau} = -\frac{D}{2\xi} (X_+ - X_-) \vec{u}_n + \epsilon \frac{D^2}{\xi^2} (X_+ - 2X_-) \vec{u}_n,$$

waarin X_+ en X_- "schuifoperatoren" zijn langs de x-as (zie paragraaf 7.1).

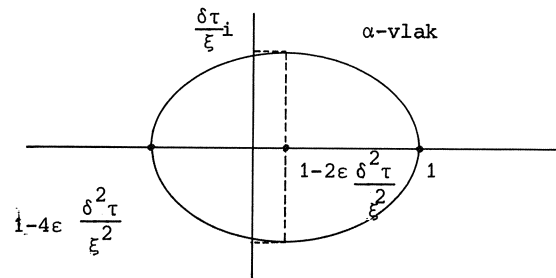
Schema (6.17) is voor $\epsilon = \frac{\tau}{2}$ juist het *Lax-Wendroff-schema* voor hyperbolische vergelijkingen van de vorm (6.13) (zie RICHTMYER en MORTON [1967, p. 302]). Het is eenvoudig te verifiëren dat het Lax-Wendroff-schema tweede orde nauwkeurig in τ en ξ is wanneer de verhouding τ/ξ constant gehouden wordt. Immers substitutie van een oplossing $\vec{u}$ van (6.13) in (6.17) en gebruikmaking van (6.13) levert het residu (afbreekfout)

$$\begin{aligned} & \frac{\tau}{2} \frac{\partial^2 \vec{u}}{\partial t^2} + \frac{\xi^2}{6} D \frac{\partial^3 \vec{u}}{\partial x^3} - \frac{\tau}{2} D^2 \frac{\partial^2 \vec{u}}{\partial x^2} + O(\tau^2) + O(\xi^3) = \\ & = \frac{\xi^2}{6} D \frac{\partial^3 \vec{u}}{\partial x^3} + O(\tau^2). \end{aligned}$$

Vervolgens onderzoeken we de stabiliteit van schema (6.17). De eigenwaarden van de amplificatiematrix van schema (6.17) worden gegeven door

$$\alpha = 1 - i \frac{\delta \tau}{\xi} \sin \gamma \xi - 2\epsilon \frac{\delta^2 \tau}{\xi^2} (1 - \cos \gamma \xi),$$

waarin δ de eigenwaarden van de matrix D doorloopt en γ de index van de eigenfuncties van A voorstelt. Wanneer $\gamma \xi$ van 0 tot 2π loopt, doorloopt α een ellips in het complexe vlak (zie figuur 6.6).

Fig. 6.6. Eigenwaarde α .

Hieruit volgt dat de stabiliteitsvoorwaarden zijn

$$(6.18) \quad \tau < \frac{\xi}{\sigma(\mathcal{D})}, \quad \tau < \frac{\xi^2}{2\epsilon\sigma^2(\mathcal{D})},$$

zodat de waarde $\epsilon = \tau/2$ geen beperking voor de stabiliteitsvoorwaarde (6.18) is. In RICHTMYER en MORTON [1967, p. 303] wordt een variant van schema (6.17) gegeven die gebruikt kan worden voor algemene hyperbolische systemen van de vorm

$$(6.13') \quad \frac{\partial \vec{u}}{\partial t} + \frac{\partial}{\partial x} \vec{f}(\vec{u}) = \vec{0}.$$

Het is een "Runge-Kutta-achtige" methode van de vorm

$$(6.17') \quad \begin{aligned} \vec{u}_{n+1}^{(1)}(x_{j+\frac{1}{2}}) &= \frac{1}{2} \left(\vec{u}_n(x_{j+1}) + \vec{u}_n(x_j) \right) - \frac{\tau}{2\xi} \left[\left(\vec{F}(\vec{u}_n(x_{j+1})) - \vec{F}(\vec{u}_n(x_j)) \right) \right] \\ \vec{u}_{n+1}(x_j) &= \vec{u}_n(x_j) - \frac{\tau}{\xi} \left[\vec{F}(\vec{u}_{n+1}^{(1)}(x_{j+\frac{1}{2}})) - \vec{F}(\vec{u}_{n+1}^{(1)}(x_{j-\frac{1}{2}})) \right]. \end{aligned}$$

Het is eenvoudig te zien dat dit schema tot schema (6.17) reduceert wanneer het op (6.13) wordt toegepast. Hieruit volgt dat de stabiliteitsvoorwaarde in de zin van de *lineaire* stabiliteitstheorie identiek zijn aan (6.18) met $\epsilon = \frac{\tau}{2}$ waarin D de Jacobiaan van de functie $\vec{f}$ voorstelt.

LAX en WENDROFF [1960] hebben hun schema ook toegepast op problemen waarin zich *schokgolven* ontwikkelen; het bleek echter dat dan additionele dissipatieve termen toegevoegd dienen te worden. Zij stelden voor (zie RICHTMYER en MORTON [1960, p.335]) een artificiele viscositeitsterm van de vorm

$$\frac{\tau}{2\xi} \left[Q^+ (\vec{U}_n(x_{j+1}) - \vec{U}_n(x_j)) - Q^- (\vec{U}_n(x_j) - \vec{U}_n(x_{j-1})) \right]$$

in het schema op te nemen; hierin is $Q^+ = Q(\vec{U}_n(\vec{x}_j), \vec{U}_n(\vec{x}_{j+1}))$ een geschikt gekozen matrix die nagenoeg de eenheidsmatrix is wanneer $\vec{U}_n(x_j)$ en $\vec{U}_n(x_{j+1})$ weinig verschillen en dissipatief werken wanneer deze veel verschillen (in het schokgebied).

In verband hiermee merken we op dat de gegeneraliseerde Euler-methoden ook verkregen kunnen worden door toevoeging van artificiele viscositeitstermen aan het Euler-schema. Dit kan men inzien door schema (3.0) te schrijven als

$$(6.19) \quad \vec{S}_{n+1} = [I + \tau \mathcal{D}_n + \tau^2 \mathcal{D}_n Q \mathcal{D}_n] \vec{S}_n,$$

waarin de windveld-term eenvoudigheidshalve weggelaten is, en waarin de artificiele viscositeitsterm $\tau^2 \mathcal{D}_n Q \mathcal{D}_n$ bepaald wordt door dit schema te identificeren met schema (3.0), dus door de relatie

$$I + \tau \mathcal{D}_n + \tau^2 \mathcal{D}_n Q \mathcal{D}_n \equiv (I - \tau E_n)^{-1} (I + \tau (\mathcal{D}_n - E_n)).$$

Uitwerking geeft voor de matrix Q de uitdrukking

$$(6.20) \quad Q = \mathcal{D}_n^{-1} (I - \tau E_n)^{-1} E_n.$$

Als voorbeeld berekenen we de "viscositeitsmatrix" Q van het Fischerschema (3.12) toegepast op een lineair model zonder bodemwrijving, Corioliskracht en kustvoorwaarden (randvoorwaarden in oneindig). In de operatoren $\mathcal{D}_n$ en E_n kunnen we dan $A_{xy} = A_x = A_y = \Omega_x$, $\Omega_y = 0$ en $A_x = A_y = I$ stellen, zodat

$$\mathcal{D}_n = - \begin{pmatrix} 0 & 0 & g \mathcal{D}_k \\ 0 & 0 & g \mathcal{D}_y \\ B_x + B \mathcal{D}_x & B_y + B \mathcal{D}_y & 0 \end{pmatrix}, \quad E_n = - \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ B_x + B \mathcal{D}_x & B_y + B \mathcal{D}_y & 0 \end{pmatrix}.$$

Substitutie in (6.20) geeft na enige herleiding

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Dat de door (6.21) gedefinieerde viscositeitsterm inderdaad dissipatief werkt kunnen we zien door weer de differentiaalvergelijkingen op te stellen waaraan de differentieoplossing van (6.19) als functie van ξ , η en τ voldoet. Aannemende dat

$$\mathcal{D}_n = \mathcal{D} + o(\xi^2) + o(\eta^2),$$

vinden we de partiele differentiaalvergelijking (vergelijk (2.32) met $\vec{f} = 0$)

$$\frac{\partial \vec{s}}{\partial t} + \frac{1}{2} \tau^2 \frac{\partial^2 \vec{s}}{\partial t^2} + o(\tau^3) = \mathcal{D} \vec{s} + \tau \mathcal{D} Q \vec{s} + o(\xi^2).$$

Beperken we ons weer tot verstoringen opgebouwd uit complexe e-machten, dan concluderen we hieruit dat $\mathcal{D}Q\mathcal{D}$ *niet-positieve* eigenwaarden moet hebben wil er van dissipatie van energie sprake zijn. Aangezien $\mathcal{D}$ zuiver *imaginaire* eigenwaarden heeft in het beschouwde model, moet Q *niet-negatieve* eigenwaarden hebben. Hieraan voldoet (6.21). De gegeneraliseerde Eulermethoden van Fischer, Sielecki, enz., kunnen opgevat worden als Eulermethoden toegepast op de oorspronkelijke differentiaalvergelijking waaraan een dissipatieve term is toegevoegd.

VII. DISCRETISERING VAN $\partial/\partial x$ EN $\partial/\partial y$

Tot dusver hebben we nauwelijks aandacht besteed aan de benadering van de differentiaaloperatoren $\partial/\partial x$ en $\partial/\partial y$ door differentie-operatoren $[\partial/\partial x]$ en $[\partial/\partial y]$. Alleen de benadering door centrale differenties is in hoofdstuk 2 en 3 terloops ter sprake gekomen. Dit is echter één van de vele mogelijkheden; in dit hoofdstuk zullen de verschillende aspecten van de discretisering van $\partial/\partial x$ en $\partial/\partial y$ aan de orde komen. We beginnen met de discretisering op *uniforme* roosters.

7.1. Gewogen centrale differenties

Beschouw een uniform rooster met roosterpuntafstanden ξ en η in de respectieve x- en y-richtingen (zie figuur 7.1).

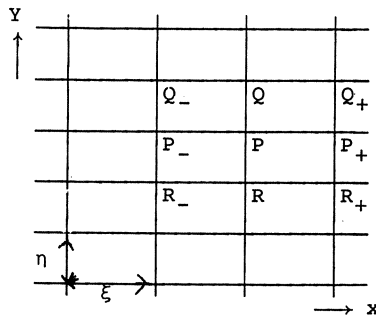


Fig. 7.1. Uniform rooster met mazen van ξ bij η .

Stel dat men de afgeleide $\partial S(x,y)/\partial x$ in een punt P wil benaderen door alleen van de waarden van $S(x,y)$ in de roosterpunten van dit rooster gebruik te maken. Men kan dit doen door het zogenaamde *gewogen centrale differentiequotient*, dat wil zeggen een lineaire combinatie van centrale differentiequotienten in de punten P, Q en R (zie figuur 7.1). Zijn (x,y) de coördinaten van P dan is het gewogen centrale differentiequotient in P gedefinieerd door

$$(7.1) \quad \left[\frac{\partial}{\partial x} \right] S(x,y) = \frac{b}{2a+b} \left[\frac{S(x+\xi,y) - S(x-\xi,y)}{2\xi} \right] + \\ + \frac{2a}{2a+b} \left[\frac{S(x+\xi,y+\eta) - S(x-\xi,y+\eta)}{2\xi} + \frac{S(x+\xi,y-\eta) - S(x-\xi,y-\eta)}{2\xi} \right].$$

Hierin zijn a en b willekeurige gewichten. In de literatuur worden dergelijke differentieformules wat compacter geschreven door gebruik te maken van "schuifoperatoren" X_+ , X_- , Y_+ en Y_- ; X_+ wordt bijvoorbeeld gedefinieerd door

$$(7.2) \quad X_+ S(x,y) = S(x+\xi,y),$$

en soortgelijke definities voor X_- en $Y_{\pm}$. Formule (7.1) kan dan geschreven worden als

$$(7.1') \quad \left[\frac{\partial}{\partial x} \right] S(x, y) = \frac{aY_+ + b + aY_-}{2a+b} \frac{X_+ - X_-}{2\xi} S(x, y),$$

of nog compacten (vergelijk de in paragraaf 6.1 geïntroduceerde notatie)

$$(7.1'') \quad \left[\frac{\partial}{\partial x} \right] S(x, y) = \frac{1}{2\xi(2a+b)} \begin{bmatrix} -a & 0 & a \\ -b & 0 & b \\ -a & 0 & a \end{bmatrix} S(x, y)$$

waarin weer de gewichten van de S-waarden in een array gezet zijn op met de roosterpunten corresponderende plaatsen.

Bijzondere gevallen van gewogen centrale differenties zijn de *centrale differenties* ($a = 0$):

$$(7.3) \quad \left[\frac{\partial}{\partial x} \right] = \frac{1}{2\xi} \begin{bmatrix} -1 & 0 & 1 \end{bmatrix},$$

en de *gemiddelde centrale differenties* ($b = 0$):

$$(7.4) \quad \left[\frac{\partial}{\partial x} \right] = \frac{1}{4\xi} \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}.$$

Voor de differentie-operator $[\partial/\partial y]$ kan men op overeenkomstige wijze de gewogen centrale differenties definiëren.

7.2. De afbreekfout van de gewogen centrale differentie-operator

Wanneer $S(x, y)$ voldoende vaak naar x en y differentieerbaar is, kan men met behulp van Taylor-ontwikkelingen schrijven

$$x_{\pm} S(x, y) = S(x \pm \xi, y) = \left[1 \pm \xi \frac{\partial}{\partial x} + \frac{1}{2} \xi^2 \frac{\partial^2}{\partial x^2} \pm \frac{1}{6} \xi^3 \frac{\partial^3}{\partial x^3} + O(\xi^4) \right] S(x, y)$$

en

$$y_{\pm} S(x, y) = S(x, y \pm \eta) = \left[1 \pm \eta \frac{\partial}{\partial y} + \frac{1}{2} \eta^2 \frac{\partial^2}{\partial y^2} \pm \frac{1}{6} \eta^3 \frac{\partial^3}{\partial y^3} + O(\eta^4) \right] S(x, y).$$

Substitutie in (7.1') geeft

$$\begin{aligned} \left[\frac{\partial}{\partial x} \right] S(x, y) &= \frac{1}{2(2a+b)\xi} \left[2a+b+a\eta^2 \frac{\partial^2}{\partial y^2} + O(\eta^4) \right] \cdot \\ &\quad \cdot \left[2\xi \frac{\partial}{\partial x} + \frac{1}{3}\xi^3 \frac{\partial^3}{\partial x^3} + O(\xi^5) \right] S(x, y) = \\ &= \left[\frac{\partial}{\partial x} + \left(\frac{1}{6}\xi^2 \frac{\partial^2}{\partial x^2} + \frac{a}{2a+b} \eta^2 \frac{\partial^2}{\partial y^2} \right) \frac{\partial}{\partial x} + O(\xi^4 + \eta^4) \right] S(x, y). \end{aligned}$$

Dus door $\partial S / \partial x$ te benaderen met $[\partial / \partial x]S$ maken we een (afbreek)-fout van de orde

$$(7.3) \quad \frac{1}{6} \xi^2 \frac{\partial^3 S}{\partial x^3} + \frac{a}{2a+b} \eta^2 \frac{\partial^3 S}{\partial y^2 \partial x},$$

en iets dergelijks voor de fout in $[\partial / \partial y]S$. Hieruit concluderen we dat voor willekeurige functies S de afbreekfout niet geminimaliseerd kan worden door een bijzondere keuze van a en b .

7.3. Het spectrum van gewogen centrale differentie-operatoren

In de stabiliteitsanalyse van de diverse schema's hebben we gezien dat de maximale waarde van de som van de kwadraten van de eigenwaarden van $[\partial / \partial x]$ en $[\partial / \partial y]$, dus

$$(7.5) \quad \text{Max} |\delta_x^2 + \delta_y^2|,$$

een belangrijke rol speelt, althans in *expliciete* schema's: hoe kleiner deze waarde des te groter de toegestane integratiestappen. Het ligt dus voor de hand na te gaan voor welke waarden van a en b uitdrukking (7.5) geminimaliseerd wordt. De eigenwaarden δ_x en δ_y worden bepaald door de vergelijkingen

$$(7.6) \quad \begin{aligned} \left[\frac{\partial}{\partial x} \right] \exp[i(\vec{\gamma}, \vec{x})] &= \delta_x \exp[i(\vec{\gamma}, \vec{x})] \\ \left[\frac{\partial}{\partial y} \right] \exp[i(\vec{\gamma}, \vec{x})] &= \delta_y \exp[i(\vec{\gamma}, \vec{x})] \end{aligned}$$

waarin $\vec{x} = (x, y)$ de roosterpunten doorloopt. Uit de definities van $[\partial/\partial x]$ en $[\partial/\partial y]$ volgt op eenvoudige wijze dat

$$\delta_x = i \frac{\sin(\gamma_1 \xi) (b + 2a \cos(\gamma_2 \eta))}{(b + 2a) \xi} \quad (7.7)$$

$$\delta_y = i \frac{\sin(\gamma_2 \eta) (b + 2a \cos(\gamma_1 \xi))}{(b + 2a) \eta}$$

Stelling 7.1. Zij $\delta(a, b) = \max_{\vec{\gamma}} \sqrt{|\delta_x^2 + \delta_y^2|}$, dan geldt

$$\frac{1}{\min(\xi, \eta)} = \delta(a, 0) \leq \delta(a, b) \leq \delta(0, b) = \sqrt{\frac{1}{\xi^2} + \frac{1}{\eta^2}}.$$

Bewijs. Zie VAN DER HOUWEN [1968].

Blijkbaar zijn de gemiddelde centrale differentie-operatoren optimaal in de zin van een minimale δ -waarde. Voor $\xi = \eta$ laten ze in expliciete schema's een $\sqrt{2}$ maal grotere integratie-stap toe dan de centrale differentie-operatoren. Bovendien kan evenals bij de centrale differenties een "space-staggered" rooster gebruikt worden (vergelijk figuur 3.7), maar nu van de vorm zoals in figuur 7.2 aangegeven is.

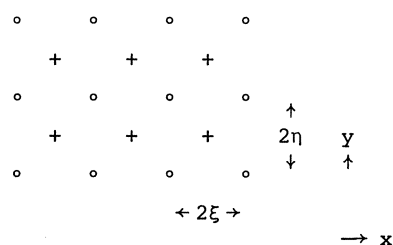


Fig. 7.2 o verhogingspunten
 + stromingspunten

De hoofdtermen in de ondiep-water-vergelijkingen kunnen zonder meer op dit rooster gerepresenteerd worden; voor de 2^e orde termen zoals de convectieve

traagheidstermen moet men waar nodig door middeling de stroming in de verhogingspunten berekenen (vergelijk de discussie in paragraaf 3.7).

7.4. Hogere orde discretisaties

De gewogen centrale differentie-operator levert een 2^e orde benadering voor $\partial/\partial x$ en $\partial/\partial y$. Hogere orde benaderingen kunnen verkregen worden door meer roosterpunten te gebruiken. Later we een 12-puntsformule, gedefinieerd op het verschoven rooster uit figuur 7.2, beschouwen:

$$(7.8) \quad \left[\frac{\partial}{\partial x} \right] = \frac{1}{\xi} \begin{pmatrix} 0 & 0 & -c & 0 & c & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -a & 0 & -b & 0 & b & 0 & a \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -a & 0 & -b & 0 & b & 0 & a \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -c & 0 & c & 0 & 0 \end{pmatrix}.$$

Met behulp van Taylor-reeksen kan afgeleid worden dat hiervoor geldt

$$(7.9) \quad \left[\frac{\partial}{\partial x} \right] = \frac{\partial}{\partial x} + O(\xi^4 + \eta^4)$$

indien men

$$(7.10) \quad a = -\frac{1}{96}, \quad b = \frac{30}{96}, \quad c = \frac{3}{96}$$

kiest.

De eigenwaarden δ_x van (7.8) en δ_y van de overeenkomstige 12-puntsformule voor $[\partial/\partial y]$ kunnen eenvoudig berekend worden; we vinden

$$(7.11) \quad \delta_x = \frac{i}{6\xi} \sin(\gamma_1 \xi) \cos(\gamma_2 \eta) (6 + \sin^2(\gamma_1 \xi) - 3\cos^2(\gamma_2 \eta))$$

$$\delta_y = \frac{i}{6\eta} \sin(\gamma_2 \eta) \cos(\gamma_1 \xi) (6 + \sin^2(\gamma_2 \eta) - 3\cos^2(\gamma_1 \xi))$$

De berekening van de waarde van $\delta = \max_{\gamma} \sqrt{|\delta_x^2 + \delta_y^2|}$ is minder eenvoudig. Men kan echter aantonen dat

$$(7.12) \quad \delta = \frac{\sqrt{343}}{27\text{Min}(\xi, \eta)} \cong \frac{2}{3\text{Min}(\xi, \eta)}.$$

Deze waarde is ongeveer 2/3 van de δ -waarde van de gemiddelde centrale differentie-operatoren. Hieruit volgt dat de 12-puntsformules van het type (7.8) anderhalf maal grotere integratiestappen toestaat dan de 4-puntsformules van het type (7.4), uiteraard wanneer expliciete schema's gebruikt worden.

Het gebruik van hogere orde discretisaties kan nuttig zijn in die delen van de beschouwde zee waar een grotere nauwkeurigheid gevraagd wordt.

7.5. Niet-uniforme roosters

Het voordeel van niet-uniforme roosters is een betere representatie van de kusten, economischere integratiestappen bij expliciete schema's door aanpassing van de mazen aan de diepte (zie b.v. voorwaarde (2.44') of (3.24)), en de mogelijkheid om in delen van de zee waar men bijzonder in geïnteresseerd is, kleinere mazen te kiezen.

Stel dat men $\partial S / \partial x$ in het punt Q van figuur 7.3 wil benaderen met gebruikmaking van de punten $P_{\pm}$, Q , $Q_{\pm}$, R en $R_{\pm}$.

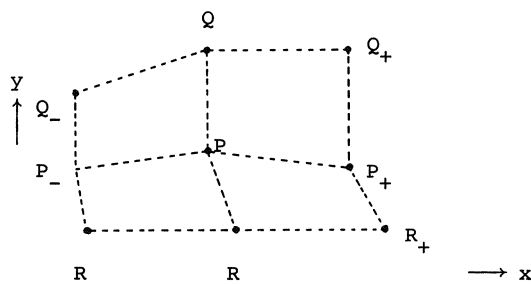


Fig. 7.3. Niet uniform rooster

Een mogelijke aanpak is de volgende: we schrijven $[\partial / \partial x]$ als een lineaire combinatie van de S -waarden in de genoemde buurpunten van Q , dus

$$(7.13) \quad \left[\frac{\partial}{\partial x} \right] S_P = p_+ S_{P_+} + p_- S_{P_-} + q S_Q + q_+ S_{Q_+} + q_- S_{Q_-} + \\ + r S_R + r_+ S_{R_+} + r_- S_{R_-}.$$

Hierin zijn de coëfficiënten $p_{\pm}, q, \dots$ nader te bepalen gewichten en stellen $S_{P_{\pm}}, S_Q, \dots$ de S -waarden in $P_{\pm}, Q, \dots$ voor. Vervolgens ontwikkelen we $S_{P_{\pm}}, S_Q, \dots$ in Taylor-reeksen t.o.v. het punt, waardoor $[\partial/\partial x]S_P$ overgaat in een reeks van partiële afgeleiden van S in P . Tenslotte stellen we de coëfficiënt van $\partial S/\partial x$ gelijk 1 en zoveel mogelijk coëfficiënten van de overige afgeleiden, te beginnen bij de laagste afgeleiden, gelijk nul. Dit levert 8 lineaire vergelijkingen voor de gewichten $p_{\pm}, q, \dots$. De afgeleiden waarvan de coëfficiënten niet nul gemaakt kunnen worden vormen tesamen de afbreekfout (vergelijk de afleiding in paragraaf 7.2). Oplossing van de vergelijkingen voor de gewichten leidt uiteindelijk tot de gezochte differentie-formule op een niet uniform rooster.

De hierboven geschetste procedure zullen we niet volgen. Het is aanzienlijk eenvoudiger om als volgt te werk te gaan. We denken ons de roosterpunten gelegen op roosterlijnen (zie figuur 7.4), die gegeven worden door

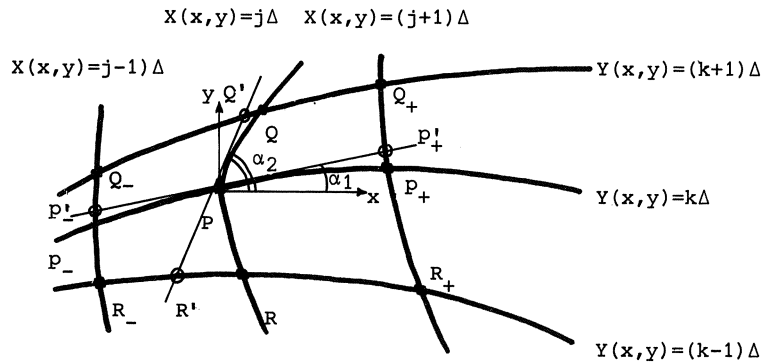


Fig. 7.4. Kromlijnige coördinaten X en Y

de vergelijkingen

$$(7.14) \quad X(x,y) = j\Delta, \quad Y(x,y) = k\Delta,$$

waarin Δ een positieve constante is en j, k gehele getallen zijn waarmee de

roosterlijnen genummerd worden. Verder nemen we aan dat de functies $X(x,y)$ en $Y(x,y)$ differentieerbaar zijn. Er geldt dan

$$(7.15) \quad \frac{\partial}{\partial x} = \frac{\partial X}{\partial x} \frac{\partial}{\partial X} + \frac{\partial Y}{\partial x} \frac{\partial}{\partial Y}, \quad \frac{\partial}{\partial y} = \frac{\partial X}{\partial y} \frac{\partial}{\partial X} + \frac{\partial Y}{\partial y} \frac{\partial}{\partial Y}.$$

Hiermee worden de differentiaaloperatoren $\partial/\partial x$ en $\partial/\partial y$ in de *carthesische* coördinaten x en y uitgedrukt in differentiaaloperatoren in de *kromlijnige* coördinaten X en Y . Deze operatoren $\partial/\partial X$ en $\partial/\partial Y$ zijn echter eenvoudig te discretiseren, omdat in het (X,Y) -vlak de roosterpunten een uniform rooster vormen met maaswijdte Δ (zie figuur 7.5).

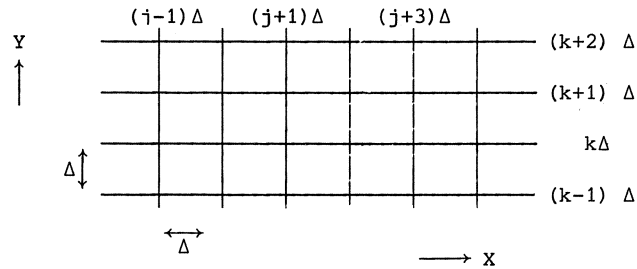


Fig. 7.5. Uniform rooster in het (X,Y) -vlak

Rest nog de partiële afgeleiden $\partial X/\partial x$, $\partial Y/\partial x$, $\partial X/\partial y$ en $\partial Y/\partial y$ uit te rekenen. Dit kunnen we doen door de *richtingsafgeleiden* van de kromlijnige coördinaten in figuur 7.4 te bepalen. Zij α de hoek die een vector $\vec{v}$ maakt met de positieve x -as, dan wordt de afgeleide in de richting van $\vec{v}$ van een functie F (de richtingsafgeleide $D_{\alpha} F$) gegeven door (zie e.g. COURANT [1936]).

$$(7.16) \quad D_{\alpha} F = \cos \alpha \frac{\partial F}{\partial x} + \sin \alpha \frac{\partial F}{\partial y}.$$

We passen deze formule toe op de functies X en Y met voor α respectievelijk de hoeken die de roosterlijnen $X(x,y) = \text{constant}$ en $Y(x,y) = \text{constant}$ maken met de positieve x -as in het punt Q (zie figuur 7.4). We vinden dan voor de partiële afgeleiden van X en Y de vier vergelijkingen

$$(7.17) \quad \begin{aligned} \cos \alpha_1 \frac{\partial X}{\partial x} + \sin \alpha_1 \frac{\partial X}{\partial y} &= D_{\alpha_1} X \approx \frac{\Delta}{\xi}, \\ \cos \alpha_2 \frac{\partial X}{\partial x} + \sin \alpha_2 \frac{\partial X}{\partial y} &= D_{\alpha_2} X = 0, \end{aligned}$$

$$\cos\alpha_1 \frac{\partial Y}{\partial x} + \sin\alpha_1 \frac{\partial Y}{\partial y} = D_{\alpha_1} Y = 0, \quad (7.17)$$

$$\cos\alpha_2 \frac{\partial Y}{\partial x} + \sin\alpha_2 \frac{\partial Y}{\partial y} = D_{\alpha_2} Y \approx \frac{\Delta}{\bar{\eta}},$$

waarin $\bar{\xi}$ en $\bar{\eta}$ de gemiddelde afstanden zijn van $\overline{PP_-}$ en $\overline{PP_+}$ respectievelijk $\overline{PQ}$ en $\overline{PR}$. Stelsel (7.17) is eenvoudig op te lossen en geeft dan

$$\begin{aligned} \frac{\partial X}{\partial x} &= \frac{\Delta}{\bar{\xi}} \frac{\sin\alpha_2}{\sin(\alpha_2 - \alpha_1)}, \quad \frac{\partial X}{\partial y} = \frac{\Delta}{\bar{\xi}} \frac{\cos\alpha_2}{\sin(\alpha_1 - \alpha_2)}, \\ (7.18) \quad \frac{\partial Y}{\partial x} &= \frac{\Delta}{\bar{\eta}} \frac{\sin\alpha_1}{\sin(\alpha_1 - \alpha_2)}, \quad \frac{\partial Y}{\partial y} = \frac{\Delta}{\bar{\eta}} \frac{\cos\alpha_1}{\sin(\alpha_2 - \alpha_1)}. \end{aligned}$$

Kiest men nu een of andere discretisering voor $\partial/\partial X$ en $\partial/\partial Y$ op het uniforme rooster van figuur 7.5, dan bepalen (7.15) en (7.18) volledig de discretisering van $\partial/\partial x$ en $\partial/\partial y$ op het niet-uniforme rooster van figuur 7.4.

We benadrukken dat de benadering van de richtingsafgeleiden $D_{\alpha_1} X$ en $D_{\alpha_2} Y$ door $\Delta/\bar{\xi}$ respectievelijk $\Delta/\bar{\eta}$ alleen dan aanvaardbaar is wanneer de roosterlijnen ten op zichte van de roosterpunten voldoende "recht" zijn (overigens zouden deze benaderingen nog iets verbeterd kunnen worden door $\bar{\xi}$ en $\bar{\eta}$ te vervangen door de gemiddelden van de afstanden $\overline{PP_-^1}$ en $\overline{PP_+^1}$ respectievelijk $\overline{PQ^1}$ en $\overline{PR^1}$ (zie figuur 7.4). Een alternatief voor de berekening van de partiele afgeleiden van de functies X en Y zou kunnen zijn de analytische voorstelling van X en Y met behulp van splines. Dit valt echter buiten het kader van deze syllabus.

7.6. Het spectrum van de centrale differentie-operatoren op een niet-uniform rooster

Op een niet-uniform rooster worden de eigenwaarden δ_x en δ_y van $[\partial/\partial x]$ en $[\partial/\partial y]$ volgens (7.15) gegeven door

$$\begin{aligned} \delta_x &= \frac{\partial X}{\partial x} \delta_x + \frac{\partial Y}{\partial x} \delta_y \\ (7.19) \quad \delta_y &= \frac{\partial X}{\partial y} \delta_x + \frac{\partial Y}{\partial y} \delta_y \end{aligned}$$

waarin δ_X en δ_Y de eigenwaarden van $[\partial/\partial X]$ en $[\partial/\partial Y]$ zijn. In het bijzonder zijn we geïnteresseerd in het maximum van $|\delta_X^2 + \delta_Y^2|$. Substitutie van (7.18) geeft na enige herleiding

$$(7.20) \quad \delta_X^2 + \delta_Y^2 = \frac{\Delta^2}{\sin^2(\alpha_2 - \alpha_1)} \left[\frac{\delta_X^2}{\xi^2} + 2\cos(\alpha_2 - \alpha_1) \frac{\delta_X \delta_Y}{\xi \bar{\eta}} + \frac{\delta_Y^2}{\bar{\eta}^2} \right].$$

Volgens (7.7) geldt voor centrale differenties $[\partial/\partial X]$ en $[\partial/\partial Y]$

$$(7.21) \quad \delta_X = i \frac{\sin(\gamma_1 \Delta) \cos(\gamma_2 \Delta)}{\Delta}, \quad \delta_Y = i \frac{\sin(\gamma_2 \Delta) \cos(\gamma_1 \Delta)}{\Delta}$$

en volgens stelling 7.1

$$\max_Y (\delta_X^2 + \delta_Y^2) \geq -\frac{1}{\Delta^2}.$$

Verder geldt

$$\delta_X \delta_Y = -\frac{\sin 2\gamma_1 \Delta \sin 2\gamma_2 \Delta}{4\Delta^2} \geq -\frac{1}{4\Delta^2}$$

zodat

$$\delta_X^2 + \delta_Y^2 \geq -\frac{1}{\sin^2(\alpha_2 - \alpha_1)} \left[\frac{1}{\min^2(\bar{\xi}, \bar{\eta})} + \frac{|\cos(\alpha_2 - \alpha_1)|}{2\bar{\xi}\bar{\eta}} \right].$$

Derhalve geldt

$$(7.22) \quad \delta = \max_Y \sqrt{|\delta_X^2 + \delta_Y^2|} = \frac{1}{|\sin(\alpha_2 - \alpha_1)|} \left[\frac{1}{\min^2(\bar{\xi}, \bar{\eta})} + \frac{|\cos(\alpha_2 - \alpha_1)|}{2\bar{\xi}\bar{\eta}} \right]^{\frac{1}{2}}.$$

Hieruit zien we dat de kleinste waarde voor δ ontstaat wanneer het rooster gevormd wordt door orthogonale kromlijnige coördinaten.

Om te zien wat de afwijking van orthonormaliteit betekent voor de grootte van de integratiestappen, beschouwen we nog eens de stabiliteitsvoorwaarde (3.24) voor het gesymmetriseerde schema. Substitutie van (7.22) in de voorwaarde voor τ geeft

$$(7.23) \quad \tau \leq \frac{2\sqrt{2} |\sin(\alpha_2 - \alpha_1)|}{\sqrt{gH(2 + |\cos(\alpha_2 - \alpha_1)|)}} \bar{\xi},$$

waarin we eenvoudigheidshalve $\bar{\eta} = \bar{\xi}$ genomen hebben.

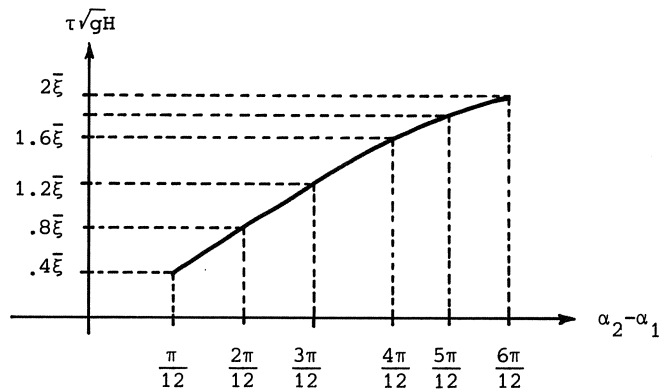


Fig. 7.6. Maximale tijdstap als functie van $\alpha_2 - \alpha_1$

We merken op dat voor orthogonale roosters de stabiliteitsvoorwaarde overgaat in

$$(7.24) \quad \tau \leq \frac{2\text{Min}(\bar{\xi}, \bar{\eta})}{\sqrt{gH}},$$

hetgeen een factor $\sqrt{2}$ grotere stappen toelaat dan het Fischer- en Sielecki-schema.

7.7. Discretisering van de randvoorwaarden

De voorgaande paragrafen betroffen steeds de discretisering van $\partial/\partial x$ en $\partial/\partial y$ in inwendige roosterpunten, dus punten omringd met andere roosterpunten. Op de randen is dit niet het geval en moeten we andere differentieformules definiëren. Laten we in het (X, Y) -vlak een rooster van het type zoals geschetst in figuur 7.2 beschouwen en wel in het bijzonder de stromingspunten op de rand van zo'n rooster (zie figuur 7.7).

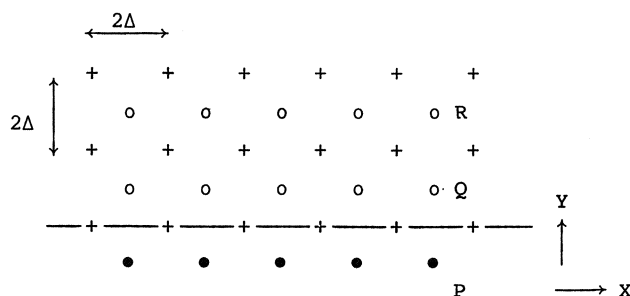


Fig. 7.7. Discretisering op de randen met geextrapoleerde centrale differentie.

In alle *inwendige* punten van dit rooster kunnen gemiddelde centrale differenties gebruikt worden. In de *randpunten* kan dit ook als we *eerst de verhogingspunten over de rand heen extrapoleren*. Dat wil zeggen we definiëren *fictieve* verhogingspunten (in figuur 7.7 met • aangegeven), berekenen door lineaire extrapolatie "landinwaarts" de verhogingen in deze punten en passen dan gemiddelde centrale differenties toe. Zo vinden we voor de waarde van een functie H in een punt P uitgedrukt in de waarden van H in de punten Q en R (zie figuur 7.7)

$$H_P = 2 H_Q - H_R.$$

In de stromingspunten op de rand vinden we op deze manier in de "molecuulnotatie" (zoals geïntroduceerd in paragraaf 7.1):

$$(7.24) \quad \left[\frac{\partial}{\partial X} \right] = \frac{1}{4\Delta} \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} = \frac{1}{4\Delta} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -3 & 0 & 3 \end{bmatrix},$$

$$(7.25) \quad \left[\frac{\partial}{\partial Y} \right] = \frac{1}{4\Delta} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & -1 \end{bmatrix} = \frac{1}{4\Delta} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & -1 \end{bmatrix}.$$

Het wordt aan de lezer overgelaten de eigenwaarden δ_X en δ_Y van deze differentie-operatoren te bepalen en het maximum van $|\delta_X^2 + \delta_Y^2|$ te berekenen.

VIII. PROGRAMMATUUR VOOR ONDIEP-WATER-PROBLEMEN

In de praktijk worden voornamelijk de oplossingsmethoden van Hansen, Fischer en Leendertse, of varianten hiervan gebruikt (VAN DER HOUWEN [1966], SIELECKI [1968], HEAPS [1969], LEENDERTSE [1971-1974]). Programmatuur voor *willekeurige kustconfiguraties* kan men vinden in LEENDERTSE [1967, 1971-74]. Op de diverse instituten, met name in Nederland, zijn aanpassingen aan locale inzichten van deze programmatuur in gebruik.

Naar aanleiding van het college aan de Universiteit van Amsterdam waaraan deze syllabus gewijd is, is door SOMMEIJER [1977] en ROZENHART [1977] programmatuur ontwikkeld, respectievelijk gebaseerd op het gesymmetriseerde Leendertse schema (3.25) en het gesymmetriseerde Hansen-Fischer schema (3.16). Uitvoerige documentatie van deze programmatuur en een groot aantal numerieke experimenten kan men in genoemde referenties vinden.

REFERENTIES

- COURANT, R. [1936]: *Differential and Integral calculus*, Blackie & Son, London.
- DRONKERS, J.J. [1964]: *Tidal computations in rivers and coastal waters*, John Wiley & Sons, Inc., New York.
- FISCHER, G. [1959]: *Ein numerisches Verfahren zur Errechnung von Windstau und Gezeiten in Randmeeren*, Tellus, 11, 60-76.
- HANSEN, W. [1956]: *Theorie zur Errechnung des Wasserstandes und der Strömungen in Randmeeren nebst Anwendungen*, Tellus, 8, 289-300.
- [1961]: *Hydrodynamical methods applied to oceanographic problems*, Proceedings of the symposium on mathematical-hydrodynamical methods of physical oceanography, Institut für Meereskunde der Universität Hamburg, 1962.
- HEAPS, N.S. [1969]: *A two-dimensional numerical sea model*, Phil. Trans., A 257.
- HOUWEN, P.J. VAN DER [1966]: *On the stability of a difference scheme for the North Sea Problem*, Report TW 100, Mathematisch Centrum, Amsterdam.

- HOUWEN, P.J. VAN DER [1968]: *Finite difference methods for solving partial differential equations*, MC Tract 20, Mathematisch Centrum, Amsterdam.
- [1971]: *Stabilized Runge-Kutta methods with limited storage requirements*, Report TW 124, Mathematisch Centrum, Amsterdam.
- [1974]: *Numerieke integratie van differentiaalvergelijkingen. Deel 1: Eenstapsmethoden*, MC Syllabus 24.1, Mathematisch Centrum, Amsterdam.
- [1975]: *Stabilized Runge-Kutta Methods for second order differential equations without first derivatives*, Report NW 26/75, Mathematisch Centrum, Amsterdam.
- LEENDERTSE, J.J. [1967]: *Aspects of a computational model for long-period water-wave propagation*, Rand Corp., Santa Monica, Mem. RM-5294.
- [1971-74]: *A water quality simulation model for well-mixed estuaries and coastal seas, Vol. I-VI*, Rand Corp., Santa Monica.
- LANCASTER, P. [1969]: *Theory of matrices*, Academic Press, New York.
- LAUWERIER, H.A. & B.R. DAMSTÉ [1963]: *The North Sea Problem VIII, A numerical treatment*, Proc. Kon. Ned. Ak. v. Wetensch., A 66, 167-184.
- PLATZMAN, G.W. [1959]: *A numerical computation of the surge of 26 June 1954 on Lake Michigan*, Geophysics, 6, 407-438.
- PROUDMAN, J. [1953]: *Dynamical oceanography*, John Wiley & Sons, Inc., New York.
- RICHTMYER, R.D. & K.W. MORTON [1967]: *Difference methods for initial-value problems*, Interscience Publishers, New York.
- ROZENHART, H.G.J. [1977]: *Een expliciet schema met variabele mazen voor de ondiep-water-vergelijkingen* (verschijnt in de NN-reeks van het Mathematisch Centrum).
- SIELECKI, A. [1968]: *Math. Weath. Rev. U.S. Dep. Agric.*, 96, 150-156.

- SOMMEIJER, B.P [1977]: *Een impliciet schema voor de ondiep-water-vergelijkingen* (verschijnt in de NN-reeks van het Mathematisch Centrum).
- STRANG, G. [1964]: *Accurate partial difference methods. II. Non-linear problems*, Numer. Math. 6, 37-46.
- VARGA, R.S. [1962]: *Matrix iterative analysis*, Prentice Hall, New Jersey.

UITGAVEN IN DE SERIE MC SYLLABUS

Onderstaande uitgaven zijn verkrijgbaar bij het Mathematisch Centrum,
2e Boerhaavestraat 49 te Amsterdam-1005, tel. 020-947272.

-
- | | |
|----------|---|
| MCS 1.1 | F. GÖBEL & J. VAN DE LUNE, <i>Leergang Besliskunde, deel 1: Wiskundige basiskennis</i> , 1965. ISBN 90 6196 014 2. |
| MCS 1.2 | J. HEMELRIJK & J. KRIENS, <i>Leergang Besliskunde, deel 2: Kansberekening</i> , 1965. ISBN 90 6196 015 0. |
| MCS 1.3 | J. HEMELRIJK & J. KRIENS, <i>Leergang Besliskunde, deel 3: Statistiek</i> , 1966. ISBN 90 6196 016 9. |
| MCS 1.4 | G. DE LEVE & W. MOLENAAR, <i>Leergang Besliskunde, deel 4: Markovketens en wachttijden</i> , 1966. ISBN 90 6196 017 7. |
| MCS 1.5 | J. KRIENS & G. DE LEVE, <i>Leergang Besliskunde, deel 5: Inleiding tot de mathematische besliskunde</i> , 1966. ISBN 90 6196 018 5. |
| MCS 1.6a | B. DORHOUT & J. KRIENS, <i>Leergang Besliskunde, deel 6a: Wiskundige programmering 1</i> , 1968. ISBN 90 6196 032 0. |
| MCS 1.6b | B. DORHOUT, J. KRIENS & J.TH. VAN LIESHOUT, <i>Leergang Besliskunde, deel 6b: Wiskundige programmering 2</i> , 1977. ISBN 90 6196 150 5. |
| MCS 1.7a | G. DE LEVE, <i>Leergang Besliskunde, deel 7a: Dynamische programmering 1</i> , 1968. ISBN 90 6196 033 9. |
| MCS 1.7b | G. DE LEVE & H.C. TIJMS, <i>Leergang Besliskunde, deel 7b: Dynamische programmering 2</i> , 1970. ISBN 90 6196 055 X. |
| MCS 1.7c | G. DE LEVE & H.C. TIJMS, <i>Leergang Besliskunde, deel 7c: Dynamische programmering 3</i> , 1971. ISBN 90 6196 066 5. |
| MCS 1.8 | J. KRIENS, F. GÖBEL & W. MOLENAAR, <i>Leergang Besliskunde, deel 8: Minimaxmethode, netwerkplanning, simulatie</i> , 1968. ISBN 90 6196 034 7. |
| MCS 2.1 | G.J.R. FÖRCH, P.J. VAN DER HOUWEN & R.P. VAN DE RIET, <i>Colloquium Stabiliteit van differentieschema's, deel 1</i> , 1967. ISBN 90 6196 023 1. |
| MCS 2.2 | L. DEKKER, T.J. DEKKER, P.J. VAN DER HOUWEN & M.N. SPIJKER, <i>Colloquium Stabiliteit van differentieschema's, deel 2</i> , 1968. ISBN 90 6196 035 5. |
| MCS 3.1 | H.A. LAUWERIER, <i>Randwaardeproblemen, deel 1</i> , 1967. ISBN 90 6196 024 X. |
| MCS 3.2 | H.A. LAUWERIER, <i>Randwaardeproblemen, deel 2</i> , 1968. ISBN 90 6196 036 3. |
| MCS 3.3 | H.A. LAUWERIER, <i>Randwaardeproblemen, deel 3</i> , 1968. ISBN 90 6196 043 6. |
| MCS 4 | H.A. LAUWERIER, <i>Representaties van groepen</i> , 1968. ISBN 90 6196 037 1. |

- MCS 5 J.H. VAN LINT, J.J. SEIDEL & P.C. BAAYEN, *Colloquium Discrete wiskunde*, 1968. ISBN 90 6196 044 4.
- MCS 6 K.K. KOKSMA, *Cursus ALGOL 60*, 1969. ISBN 90 6196 045 2.
- MCS 7.1 *Colloquium Moderne rekenmachines, deel 1*, 1969. ISBN 90 6196 046 0.
- MCS 7.2 *Colloquium Moderne rekenmachines, deel 2*, 1969. ISBN 90 6196 047 9.
- MCS 8 H. BAVINCK & J. GRASMAN, *Relaxatietrillingen*, 1969. ISBN 90 6196 056 8.
- MCS 9.1 T.M.T. COOLEN, G.J.R. FÖRCH, E.M. DE JAGER & H.G.J. PIJLS, *Elliptische differentiaalvergelijkingen, deel 1*, 1970. ISBN 90 6196 048 7.
- MCS 9.2 W.P. VAN DEN BRINK, T.M.T. COOLEN, B. DIJKHUIS, P.P.N. DE GROEN, P.J. VAN DER HOUWEN, E.M. DE JAGER, N.M. TEMME & R.J. DE VOGELAERE, *Colloquium Elliptische differentiaalvergelijkingen, deel 2*, 1970. ISBN 90 6196 049 5.
- MCS 10 J. FABIUS & W.R. VAN ZWET, *Grondbegrippen van de waarschijnlijkheidsrekening*, 1970. ISBN 90 6196 057 6.
- MCS 11 H. BART, M.A. KAASHOEK, H.G.J. PIJLS, W.J. DE SCHIPPER & J. DE VRIES, *Colloquium Halfalgebra's en positieve operatoren*, 1971. ISBN 90 6196 067 3.
- MCS 12 T.J. DEKKER, *Numerieke algebra*, 1971. ISBN 90 6196 068 1.
- MCS 13 F.E.J. KRUSEMAN ARETZ, *Programmeren voor rekenautomaten; De MC ALGOL 60 vertaler voor de EL X8*, 1971. ISBN 90 6196 069 X.
- MCS 14 H. BAVINCK, W. GAUTSCHI & G.M. WILLEMS, *Colloquium Approximatie-theorie*, 1971. ISBN 90 6196 070 3.
- MCS 15.1 T.J. DEKKER, P.W. HEMKER & P.J. VAN DER HOUWEN, *Colloquium Stijve differentiaalvergelijkingen, deel 1*, 1972. ISBN 90 6196 078 9.
- MCS 15.2 P.A. BEENTJES, K. DEKKER, H.C. HEMKER, S.P.N. VAN KAMPEN & G.M. WILLEMS, *Colloquium Stijve differentiaalvergelijkingen, deel 2*, 1973. ISBN 90 6196 079 7.
- MCS 15.3 P.A. BEENTJES, K. DEKKER, P.W. HEMKER & M. VAN VELDHUIZEN, *Colloquium Stijve differentiaalvergelijkingen, deel 3*, 1975. ISBN 90 6196 118 1.
- MCS 16.1 L. GEURTS, *Cursus Programmeren, deel 1: De elementen van het programmeren*, 1973. ISBN 90 6196 080 0.
- MCS 16.2 L. GEURTS, *Cursus Programmeren, deel 2: De programmeertaal ALGOL 60*, 1973. ISBN 90 6196 087 8.
- MCS 17.1 P.S. STOBBE, *Lineaire algebra, deel 1*, 1974. ISBN 90 6196 090 8.
- MCS 17.2 P.S. STOBBE, *Lineaire algebra, deel 2*, 1974. ISBN 90 6196 091 6.
- MCS 17.3 N.M. TEMME, *Lineaire algebra, deel 3*, 1976. ISBN 90 6196 123 8.
- MCS 18 F. VAN DER BLIJ, H. FREUDENTHAL, J.J. DE IONGH, J.J. SEIDEL & A. VAN WIJNGAARDEN, *Een kwart eeuw wiskunde 1946-1971, Syllabus van de Vakantiecursus 1971*, 1974. ISBN 90 6196 092 4.
- MCS 19 A. HORDIJK, R. POTHARST & J.Th. RUNNENBURG, *Optimaal stoppen van Markovketens*, 1974. ISBN 90 6196 093 2.

- MCS 20 T.M.T. COOLEN, P.W. HEMKER, P.J. VAN DER HOUWEN & E. SLAGT, *ALGOL 60 procedures voor begin- en randwaardeproblemen*, 1976. ISBN 90 6196 094 0.
- MCS 21 J.W. DE BAKKER (red.), *Colloquium Programmacorrectheid*, 1975. ISBN 90 6196 103 3.
- MCS 22 R. HELMERS, F.H. RUYMGAART, M.C.A. VAN ZUYLEN & J. OOSTERHOFF, *Asymptotische methoden in de toetsingstheorie; Toepassingen van naburigheid*, 1976. ISBN 90 6196 104 1.
- MCS 23.1 J.W. DE ROEVER (red.), *Colloquium Onderwerpen uit de biomathe-matica, deel 1*, 1976. ISBN 90 6196 105 X.
- * MCS 23.2 J.W. DE ROEVER (red.), *Colloquium Onderwerpen uit de biomathe-matica, deel 2*, 1976. ISBN 90 6196 115 7.
- MCS 24.1 P.J. VAN DER HOUWEN, *Numerieke integratie van differentiaalver-gelijkingen, deel 1: Eenstapsmethoden*, 1974. ISBN 90 6196 106 8.
- MCS 25 *Colloquium Structuur van programmeertalen*, 1976. ISBN 90 6196 116 5.
- MCS 26.1 N.M. TEMME (ed.), *Nonlinear analysis, volume 1*, 1976. ISBN 90 6196 117 3.
- MCS 26.2 N.M. TEMME (ed.), *Nonlinear analysis, volume 2*, 1976. ISBN 90 6196 121 1.
- MCS 27 M. BAKKER, P.W. HEMKER, P.J. VAN DER HOUWEN, S.J. POLAK & M. VAN VELDHUIZEN, *Colloquium Discretiseringsmethoden*, 1976. ISBN 90 6196 124 6.
- MCS 28 O. DIEKMANN, N.M. TEMME (EDS), *Nonlinear Diffusion Problems*, 1976. ISBN 90 6196 126 2.
- MCS 29.1 J.C.P. BUS (red.), *Colloquium Numerieke programmatuur, deel 1A, deel 1B*, 1976. ISBN 90 6196 128 9.
- * MCS 29.2 H.J.J. TE RIELE (red.), *Colloquium Numerieke programmatuur, deel 2*, 1976. ISBN 144 0.
- * MCS 30 P. GROENEBOOM, R. HELMERS, J. OOSTERHOFF & R. POTHARST, *Effi-ciency begrippen in de statistiek*, 1977. ISBN 90 6196 149 1.
- MCS 31 J.H. VAN LINT (red.), *Inleiding in de coderingstheorie*, 1976. ISBN 90 6196 136 X.
- MCS 32 L. GEURTS (red.), *Colloquium Bedrijfssystemen*, 1976. ISBN 90 6196 137 8.
- MCS 33 P.J. VAN DER HOUWEN, *Differentieschema's voor de berekening van waterstanden in zeeën en rivieren*, ISBN 90 6196 138 6.
- * MCS 34 J. HEMELRIJK, *Oriënterende cursus mathematische statistiek*, ISBN 90 6196 139 4.
- * MCS 35 P.J.W. TEN HAGEN (red.), *Colloquium Computer Graphics*, 1977. ISBN 90 6196 142 4.
- * MCS 36 J.M. AARTS, J. DE VRIES (red.), *Colloquium Topologische Dynamische Systemen*, 1977. ISBN 90 6196 143 2.

De met een * gemerkte uitgaven moeten nog verschijnen.

