

RA

stichting
mathematisch
centrum



REKENAFDELING

CR 17/70

JUNI

CURSUS WETENSCHAPPELIJK REKENEN B

PROF.DR. E. VAN SPIEGEL
NUMERIEKE ANALYSE, DEEL 1b

RA

2e boerhaavestraat 49 amsterdam

BIBLIOTHEEK MATHEMATISCH CENTRUM
 ANSTERDAM

Partiële differentiaalvergelijkingen

We zullen ons in de volgende paragrafen bezighouden met numerieke oplossingsmethoden voor partiële differentiaalvergelijkingen.

Speciaal van praktische betekenis zijn de tweede-orde partiële differentiaalvergelijkingen die we zullen onderscheiden in drie typen: hyperbolische, elliptische en parabolische vergelijkingen.

6. Eerste-orde partiële differentiaalvergelijkingen

Als inleiding op de behandeling van de drie genoemde typen differentiaalvergelijkingen beschouwen we eerst partiële differentiaalvergelijkingen van de eerste orde in twee onafhankelijke variabelen x en y . We gebruiken de behandeling van deze vergelijkingen, die praktisch niet voorkomen, om een aantal begrippen die in de theorie van de partiële differentiaalvergelijkingen een rol spelen, toe te lichten.

De vergelijking

$$a \frac{\partial u}{\partial x} + b \frac{\partial u}{\partial y} = c \quad (6.1)$$

wordt een quasi-lineaire eerste-orde partiële differentiaalvergelijking genoemd als de coëfficiënten a , b en c functies zijn van x , y en u , maar niet afhangen van de afgeleiden $\frac{\partial u}{\partial x}$ en $\frac{\partial u}{\partial y}$. Het woord lineair slaat op het feit dat de afgeleiden van de hoogste orde, hier $\frac{\partial u}{\partial x}$ en $\frac{\partial u}{\partial y}$, lineair voorkomen; quasi betekent dat in de coëfficiënten a en b de onbekende functie u mag voorkomen. Zijn a en b niet afhankelijk van u dan heet de vergelijking (6.1) lineair. Verder betekent eerste-orde dat alleen de eerste afgeleiden van u voorkomen.

We gebruiken in het vervolg de standaard notatie

$$\frac{\partial u}{\partial x} = p \quad \frac{\partial u}{\partial y} = q \quad (6.2)$$

waardoor de differentiaalvergelijking (6.1) overgaat in

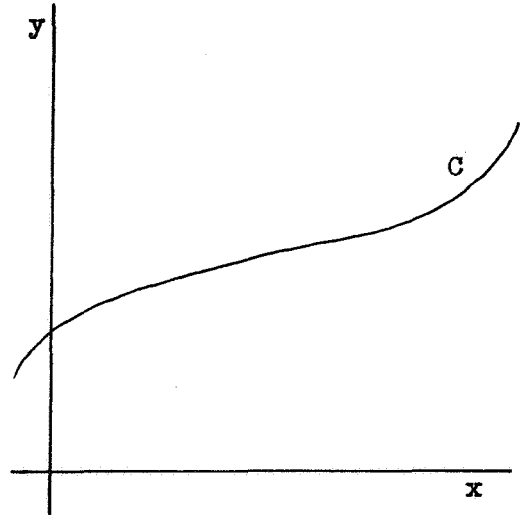
$$ap + bq = c \quad (6.3)$$

Bij een gewone differentiaalvergelijking van de eerste orde hebben we de waarde van de gezochte functie in één punt, het beginpunt, nodig om een éénduidige oplossing te kunnen vinden. In analogie hiermee mogen we verwachten dat bij een eerste orde partiële differentiaalvergelijking het voorschrijven van de onbekende functie op een kromme C , de beginkromme, in het (x,y) -vlak voldoende is om een éénduidig bepaalde oplossing te krijgen. We zullen daarom onder de oplos-

sing van (6.1) met gegeven beginwaarden verstaan: de functie $u = u(x,y)$ die voldoet aan (6.1) en op de kromme C de voorgeschreven waarden aanneemt.

We kunnen het probleem nog anders formuleren.

In de drie-dimensionale (x,y,u) -ruimte is een kromme gegeven, waarvan C de projectie op het (x,y) -vlak is. Gevraagd wordt dan het oppervlak $u = u(x,y)$ te bepalen, dat gaat door de gegeven kromme, zodanig dat in ieder punt van het oppervlak aan de differentiaalvergelijking (6.1) voldaan is.



Nu zijn bij een willekeurig oppervlak $u = u(x,y)$ de getallen $p = \frac{\partial u}{\partial x}$, $q = \frac{\partial u}{\partial y}$ en -1 in het punt (x,y,u) de richtingsgetallen van de normaal op het oppervlak in dit punt. De vergelijking (6.1) of (6.3) betekent dan meetkundig dat de rechte met richtingsgetallen a , b en c loodrecht staat op de normaal en dus ligt in het raakvlak door (x,y,u) aan het oppervlak. We kunnen een willekeurige kromme op het oppervlak aangeven met de parameteraanpak

$$x = x(s), \quad y = y(s), \quad u = u(s) = u(x(s), y(s)). \quad (6.4)$$

De raaklijn aan deze kromme, die tevens raaklijn aan het oppervlak is heeft de richtingsgetallen $\frac{dx}{ds}$, $\frac{dy}{ds}$ en $\frac{du}{ds}$. De rechte met richtingsgetallen a , b en c zal daarom precies dan raaklijn zijn aan het oppervlak $u = u(x,y)$ als er een kromme (6.4) bestaat waarvoor geldt

$$\frac{dx}{ds} = a, \quad \frac{dy}{ds} = b, \quad \frac{du}{ds} = c. \quad (6.5)$$

De gezochte oplossing van de differentiaalvergelijking zal daarom de functie $u = u(x,y)$ zijn die aan deze drie vergelijkingen voldoet en bovendien op C de voorgeschreven waarden aanneemt.

We nemen nu aan dat de vergelijking (6.1) lineair is d.w.z. a en b zijn onafhankelijk van u . De eerste twee vergelijkingen van (6.5) kunnen we dan vervangen door

$$\frac{dy}{dx} = \frac{b}{a} = f(x,y), \quad (\text{ev. } \frac{dx}{dy} = \frac{a}{b} \text{ als } a = 0). \quad (6.6)$$

Deze vergelijking definieert een schaar krommen in het (x,y) -vlak die we de karakteristieken van de differentiaalvergelijking noemen.

De vergelijking zelf heet de karakteristieke vergelijking. Op iedere karakteristiek moet volgens (6.5) gelden

$$\frac{du}{dx} = \frac{c}{a}, \quad (\text{ev. } \frac{du}{dy} = \frac{c}{b}). \quad (6.7)$$

Deze vergelijking zullen we de karakteristieke voorwaarde noemen.

We zien dus dat het oplossen van (6.1) in het geval van een lineaire partiële differentiaalvergelijking neerkomt op het oplossen van twee gewone eerste-orde differentiaalvergelijkingen, en wel op de volgende wijze: Zoek bij een willekeurig punt (x, y) van C (beginwaarde) de karakteristiek met (6.6) en integreer vervolgens langs deze karakteristiek de karakteristieke voorwaarde (6.7) met als beginvoorwaarde de voorgeschreven waarde van u in het punt (x, y) .

In het algemeen zal er één oplossing $u = u(x, y)$ bestaan bij gegeven functie-waarden op C . Als echter de kromme C zelf een karakteristiek is, d.w.z. als de kromme C voldoet aan (6.6) dan

kunnen we met de karakteristieke vergelijking niet van de kromme C afkomen. Er zal dan zeker geen oplossing bestaan als op C de voorgeschreven u niet aan de karakteristieke voorwaarde voldoet, terwijl er oneindig veel oplossingen zullen bestaan als u op C wel voldoet aan (6.7). We zullen dit resultaat nu afleiden op een manier, die gemakkelijk uitgebreid kan worden tot het geval van de tweede-orde partiële differentiaalvergelijkingen.

Uit de Taylorreeksontwikkeling

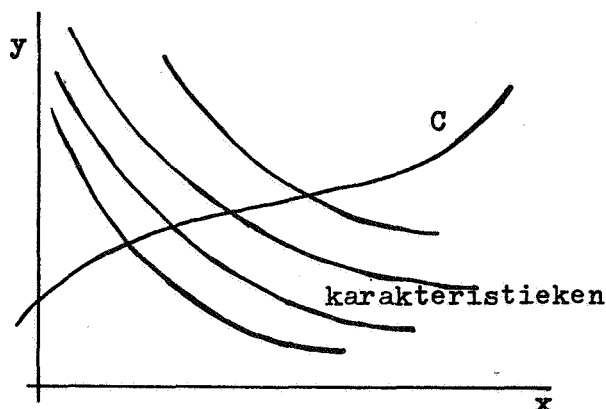
$$u(x, y) = u(x_0, y_0) + (x - x_0) \frac{\partial u}{\partial x} + (y - y_0) \frac{\partial u}{\partial y} + \dots$$

volgt dat $u(x, y)$ in een omgeving van het punt (x_0, y_0) op C alleen maar bepaald (en éénvoudig) is als in dat punt de getallen $p = \frac{\partial u}{\partial x}$ en $q = \frac{\partial u}{\partial y}$ (éénvoudig) te bepalen zijn. Dus uitgaande van de gegeven functie u op C vragen we naar de mogelijkheid om p en q op deze kromme te bepalen zodanig dat aan de differentiaalvergelijking voldaan is.

Een parameterstelling van C is

$$x = x(\lambda), \quad y = y(\lambda)$$

waarbij λ een parameter, bijvoorbeeld de booglengte vanaf een zeker punt, van de kromme C is. Als p en q bestaan dan geldt voor de functie $u = u(\lambda)$ op C



$$\frac{du}{d\lambda} = \frac{\partial u}{\partial x} \frac{dx}{d\lambda} + \frac{\partial u}{\partial y} \frac{dy}{d\lambda} .$$

Deze vergelijking samen met (6.3) geeft voor p en q het volgende stelsel

$$p \frac{dx}{d\lambda} + q \frac{dy}{d\lambda} = \frac{du}{d\lambda} ,$$

(6.8)

$$pa + qb = c .$$

Dit stelsel vergelijkingen heeft precies één oplossing

$$p = \frac{c \frac{dy}{d\lambda} - b \frac{du}{d\lambda}}{a \frac{dy}{d\lambda} - b \frac{dx}{d\lambda}} , \quad q = \frac{a \frac{du}{d\lambda} - c \frac{dx}{d\lambda}}{a \frac{dy}{d\lambda} - b \frac{dx}{d\lambda}}$$

als de determinant

$$a \frac{dy}{d\lambda} - b \frac{dx}{d\lambda} \neq 0 .$$

Is $a \frac{dy}{d\lambda} - b \frac{dx}{d\lambda} = 0$ dan heeft het stelsel (6.8) geen oplossing tenzij ook $c \frac{dy}{d\lambda} - b \frac{du}{d\lambda} = 0$. In dat geval is tevens $a \frac{du}{d\lambda} - c \frac{dx}{d\lambda} = 0$ en zijn er oneindig veel oplossingen voor p en q. Nu is $a \frac{dy}{d\lambda} - b \frac{dx}{d\lambda} = 0$ equivalent met de karakteristieke vergelijking en $c \frac{dy}{d\lambda} - b \frac{du}{d\lambda} = 0$ equivalent met de karakteristieke voorwaarde. Hieruit zien we dan tenslotte dat in het geval dat de geveven kromme C een karakteristiek is de differentiaalvergelijking dan en slechts dan oplosbaar is (en dan oneindig veel oplossingen heeft) als de geveven functie u op C voldoet aan de karakteristieke voorwaarde.

Voorbeeld

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = 1 \quad \text{ofwel} \quad p + q = 1 .$$

De karakteristieken worden geveven door $\frac{dy}{dx} = 1$, de karakteristieke voorwaarde is $\frac{du}{dx} = 1$.

De vergelijking van de karakteristieken is dus $y = x + a$ met a een willekeurige constante. Langs deze karakteristiek krijgen we $u = x + b = y + b - a$.

We nemen als beginkromme C het interval (0,1) op de x-as.

Op C wordt voorgeschreven $u_r = f(x_r)$.

Dan geldt voor de oplossing $u - u_r =$
 $= x - x_r = y$ op de lijn $y = x - x_r$.

De oplossing is dus eenduidig bepaald in het gebied begrensd door de karakteristieken door $x = 0$ en $x = 1$ en wordt gegeven door

$$u(x,y) = f(x - y) + y.$$

Veronderstel dat we voor C een karakteristiek nemen, bijvoorbeeld de rechte

$y = x$ dan is er alleen een oplossing

mogelijk als op deze rechte geldt $u = x = y$. In dat geval is de oplossing buiten de rechte $y = x$ niet eenduidig.

We kunnen bijvoorbeeld nemen

$$u(x,y) = x + \alpha(x - y).$$

Voor een willekeurige waarde van α voldoet deze aan de vergelijking $p + q = 1$ en heeft de voorgeschreven waarde op de rechte $y = x$. We kunnen ook links en rechts van de rechte $y = x$ verschillende waarden van α nemen, dan worden de afgeleiden p en q discontinu op de rechte $y = x$.

We bekijken nog de mogelijkheid van het optreden van discontinuïteiten in de beginvoorwaarden en de voortplanting van deze discontinuïteiten in het veld. Daartoe nemen we weer de differentiaalvergelijking $p + q = 1$ met beginkromme C het interval $(0,1)$. We definiëren de beginwaarden van u op C als volgt:

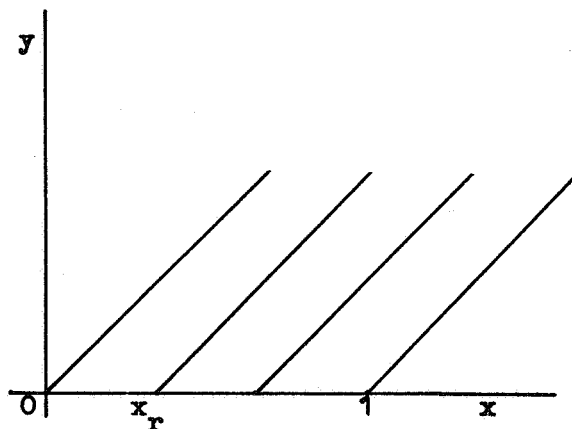
$$u(x,0) = \begin{cases} f_1(x) & 0 \leq x \leq \alpha \\ f_2(x) & \alpha < x \leq 1 \end{cases}$$

waarbij we veronderstellen dat $f_1(\alpha) \neq f_2(\alpha)$, zodat $u(x,0)$ discontinu is in $x = \alpha$. Deze discontinuïteit plant zich voort langs de karakteristiek door $x = \alpha$. We hebben eigenlijk twee verschillende problemen $p + q = 1$ met in het ene geval de beginwaarden $u = f_1(x)$ op $(0,\alpha)$, en in het andere geval $u = f_2(x)$ op $(\alpha,1)$.

Ook als een discontinuïteit in de afgeleide voorkomt wordt deze langs de karakteristiek voortgeplant. Neem bijvoorbeeld

$$u(x,0) = \begin{cases} 1 & 0 \leq x \leq \frac{1}{2} \\ x + \frac{1}{2} & \frac{1}{2} \leq x \leq 1 \end{cases}$$

De oplossing van dit probleem is



$$u(x,y) = \begin{cases} 1 + y & y \geq x - \frac{1}{2} \\ x + \frac{1}{2} & y < x - \frac{1}{2} \end{cases}$$

waarin we duidelijk zien dat nu de afgeleiden p en q discontinu zijn langs de karakteristiek $y = x - \frac{1}{2}$.

Opmerking. We zullen in het vervolg gebruik maken van het begrip differentiaal. Het voordeel van het gebruik van differentialen in een formule is dat we daarmee de introductie van een niet-essentiele parameter vermijden. Nemen we bijvoorbeeld de vergelijkingen (6.5). Voor de hierin voorkomende s mogen we een willekeurige parameter van de kromme (6.4) nemen. We schrijven daarom liever de vergelijkingen (6.5) in een vorm, waarin de parameter s niet voorkomt, nl

$$\frac{dx}{a} = \frac{dy}{b} = \frac{du}{c} . \quad (6.9)$$

Evenzo geldt dat de expliciete parametervoorstelling van C (blz NA-70, onderaan) onbelangrijk is. De vergelijkingen (6.8) worden met behulp van differentialen

$$\begin{aligned} p dx + q dy &= du \\ p a + q b &= c \end{aligned}$$

en de oplossing hiervan heeft dan de vorm

$$p = \frac{c dy - b du}{a dy - b dx} , \quad q = \frac{a du - c dx}{a dy - b dx} .$$

Ook in het geval van een quasi-lineaire vergelijking, dus als a en b afhankelijk zijn van u , noemen we de vergelijking

$$\frac{dy}{dx} = \frac{b}{a}$$

de karakteristieke vergelijking. De krommen $y = y(x)$ in het (x,y) -vlak die voldoen aan deze vergelijking noemen we weer de karakteristieken van de differentiaalvergelijking. Het verschil met het lineaire geval is, dat we hier de karakteristieken niet kunnen bepalen onafhankelijk van de oplossing $u = u(x,y)$ van de differentiaalvergelijking.

Voorbeeld

$$u \frac{\partial u}{\partial x} + y \frac{\partial u}{\partial y} = 1 .$$

Formule (6.9) wordt in dit geval $\frac{dx}{u} = \frac{dy}{y} = \frac{du}{1}$.

De karakteristieke vergelijking $\frac{dx}{u} = \frac{dy}{y}$ is niet direct te integreren. Daarom lossen we de twee andere vergelijkingen

$$dx = u du \quad \text{en} \quad \frac{dy}{y} = du$$

op, en vinden de oplossingen

$$x = \frac{1}{2}u^2 + \alpha \quad \text{en} \quad y = \beta e^u .$$

Hiermee hebben we een parametervoorstelling van de karakteristieken gevonden.

We nemen als beginkromme C het gedeelte van de rechte $y = 1$, waarvoor $0 \leq x \leq a$, en als beginvoorwaarde: $u = 0$ op C.

Uit deze beginvoorwaarde volgt $\beta = 1$ en $\alpha = x_r$ als x_r het punt op C is waar de karakteristiek doorgaat.

Voor de oplossing $u = u(x,y)$ van de differentiaalvergelijking die voldoet aan de beginvoorwaarde geldt dus

$$x - x_r = \frac{1}{2}u^2, \quad y = e^u \quad (6.10)$$

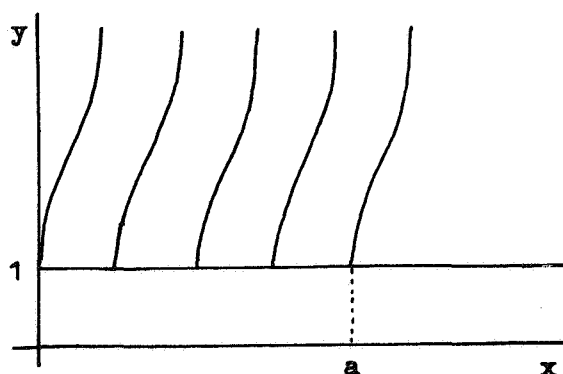
langs de karakteristiek die gaat door het punt $(x_r, 1)$.

Eliminatie van u uit deze twee betrekkingen geeft een expliciete uitdrukking voor de karakteristiek

$$y = e^{\sqrt{2(x-x_r)}} \quad y \geq 1.$$

Eliminatie van x_r uit bovenstaande betrekkingen (6.10) geeft de vergelijkingen van het oppervlak

$$y = e^u.$$



Als de differentiaalvergelijking niet analytisch op te lossen is moeten we een numerieke integratiemethode gebruiken. Een voor de hand liggende manier is discretisering van de vergelijking. Dat geeft een differentieformule op een rechthoekig rooster in het (x,y) -vlak. Deze methode kan echter ernstige moeilijkheden opleveren als de karakteristieken niet door de roosterpunten gaan, hetgeen speciaal het geval is wanneer de karakteristieken niet recht zijn, zoals in bovenstaand voorbeeld. De moeilijkheden ontstaan wanneer de beginvoorwaarde een discontinuïteit bevat, aangezien deze discontinuïteit zich langs een karakteristiek voortplant, en dus tussen de roosterpunten terecht komt. Het differentieschema is in dit geval een onnauwkeurige benadering van het probleem. In een dergelijk geval verdient het aanbeveling om gebruik te maken van een karakteristiekenmethode.

Een vraag die speciaal bij numerieke behandeling van het probleem belangrijk is, is de volgende. Is het mogelijk de gezochte oplossing te verkrijgen door integratie langs een kromme, zonder last te hebben van de functie in de omgeving van de kromme (hiermee bedoelen we dat in de integratieformule geen afgeleiden \neq richtingsafgeleide van de kromme voorkomen).

Beschouw daartoe een willekeurige kromme in het (x,y) -vlak. Langs deze kromme geldt

$$du = pdx + qdy.$$

Hiermee kunnen we uit de differentiaalvergelijking $ap + bq = c$ bijvoorbeeld p elimineren; we krijgen dan

$$adu - cdx + q(bdx - ady) = 0.$$

In deze formule hebben we geen last van de richtingsvreemde afgeleide q als $bdx - ady = 0$. Deze vergelijking voor de integratiekromme is juist de karakteristieke vergelijking, en de integratieformule langs deze kromme wordt dan $adu - cdx = 0$ hetgeen juist de karakteristieke voorwaarde is. Gezamenlijk vinden we dus weer de oorspronkelijke relaties.

We geven tenslotte nog een voorbeeld van numerieke integratie langs een karakteristiek. We nemen het bovenstaande voorbeeld $up + yq = 1$ waarvan de karakteristieke vergelijking en voorwaarde respectievelijk zijn

$$ydx = udy, \quad dx = udu.$$

Als benadering voor deze vergelijking nemen we

$$\frac{1}{2}(y + y_0)(x - x_0) = \frac{1}{2}(u + u_0)(y - y_0), \quad (x - x_0) = \frac{1}{2}(u + u_0)(u - u_0).$$

Als $u_0 = u(x_0, y_0)$ dan kunnen we uit deze twee niet-lineaire algebraïsche vergelijkingen bij gegeven waarde van x , de waarden van y en u berekenen waarvoor dan (bij benadering) geldt: (x, y) ligt op de karakteristiek door (x_0, y_0) en $u = u(x, y)$. In het algemeen zal de oplossing van een dergelijk stelsel iteratief bepaald worden. In dit geval is het oplossen zeer simpel omdat we u direct kunnen bepalen uit de tweede vergelijking en daarna y uit de eerste.

Bijvoorbeeld: als $x_0 = 0$, $y_0 = 1$, $u_0 = 0$ en we nemen $x = 0.2$ dan is $u = \sqrt{0.4} = 0.6325$ en daaruit volgt dat $y = 1.9249$ een benadering voor de exacte waarde van y die gelijk is aan $e^{0.6325} = 1.8823$.

We zien dus dat per stap gelijktijdig een volgend punt van de karakteristiek en de funktiewaarde in dit punt worden berekend. De benadering wordt natuurlijk beter als de stapgrootte in x kleiner genomen wordt. In de praktijk zal men de afbreekfout controleren door verschillende stapgrootten te gebruiken. De in ons voorbeeld gebruikte benaderingen zijn gelijkwaardig met de trapezium-regel en de afbreekfout is dan ook dienovereenkomstig.

7. Tweede-orde partiële differentiaalvergelijkingen

We beschouwen nu de quasi-lineaire partiële differentiaalvergelijking van de tweede orde voor een functie $u = u(x, y)$

$$a \frac{\partial^2 u}{\partial x^2} + b \frac{\partial^2 u}{\partial x \partial y} + c \frac{\partial^2 u}{\partial y^2} = e. \quad (7.1)$$

Omdat de differentiaalvergelijking quasi-lineair verondersteld is, kunnen de coëfficiënten a , b , c en e functies zijn van x , y , u , p en q , maar moeten onafhankelijk zijn van r , s en t waarbij

$$r = \frac{\partial^2 u}{\partial x^2}, \quad s = \frac{\partial^2 u}{\partial x \partial y}, \quad t = \frac{\partial^2 u}{\partial y^2}. \quad (7.2)$$

Met behulp van deze standaard notatie kunnen we in plaats van (7.1) ook schrijven

$$ar + bs + ct = e. \quad (7.3)$$

Om de gegevens te vinden die nodig zijn voor een éénduidig bepaalde oplossing van het probleem beschouwen we weer eerst de gewone differentiaalvergelijking. Bij een tweede-orde differentiaalvergelijking hebben we nodig de functiewaarde en de afgeleide in het beginpunt. Analooq hieraan nemen we nu een beginkromme C en op die kromme schrijven we de functiewaarde u en de eerste afgeleiden p en q voor. We mogen echter p en q niet willekeurig voorschrijven want langs de kromme C moet gelden

$$du = pdx + qdy.$$

(Dit is de voorwaarde dat $u(x, y)$ langs de kromme differentieerbaar is). We krijgen zo een beginwaarde-probleem of Cauchy-probleem. Voor dit probleem geldt (analooq als in het geval van de eerste orde partiële differentiaalvergelijking) dat de oplossing bestaat en éénduidig is als de tweede afgeleiden r , s en t éénduidig te bepalen zijn. Voor deze afgeleiden gelden de volgende relaties

$$\begin{aligned} dp &= rdx + sdy, \\ dq &= \quad \quad sdx + tdy, \\ e &= ra + sb + tc. \end{aligned} \quad (7.4)$$

Dit stelsel vergelijkingen heeft een éénduidig bepaalde oplossing als de coëfficiënten-determinant niet verdwijnt, d.w.z.

$$\begin{vmatrix} dx & dy & 0 \\ 0 & dx & dy \\ a & b & c \end{vmatrix} \neq 0$$

Is de coëfficiënten-determinant gelijk aan nul ofwel

$$c(dx)^2 - b \cdot dx \cdot dy + a(dy)^2 = 0 \quad (7.5)$$

dan heeft het stelsel vergelijkingen (7.4) geen éénduidig bepaalde oplossing. Er bestaat helemaal geen oplossing, tenzij alle andere determinanten van het stelsel eveneens verdwijnen. In dat geval zijn er oneindig veel oplossingen. Voor dit laatste geval is het voldoende dat één van de andere determinanten gelijk aan nul is, bijvoorbeeld

$$\begin{vmatrix} dp & dx & 0 \\ dq & 0 & dy \\ e & a & c \end{vmatrix} = 0$$

ofwel

$$e \cdot dx \cdot dy - a \cdot dp \cdot dy - c \cdot dq \cdot dx = 0. \quad (7.6)$$

Vergelijking (7.5) die een betrekking geeft tussen dx en dy noemen we weer de karakteristieke vergelijking, de oplossingskrommen heten weer karakteristieken. Vergelijking (7.6) die moet gelden langs de karakteristieken noemen we weer de karakteristieke voorwaarde.

Als de karakteristieken reëel zijn en de beginkromme geen karakteristiek dan kunnen we met (7.5) en (7.6) het beginwaarde-probleem numeriek oplossen. Analoog als bij de eerste-orde vergelijking bepalen we dan gelijktijdig nieuwe punten van het karakteristiekennet en de oplossing $u(x,y)$ in deze punten. We zullen dit nog expliciet behandelen.

De karakteristieken zijn ook in dit geval krommen waarlangs we de tweede-orde partiële differentiaalvergelijking kunnen schrijven als een gewone differentiaalvergelijking, m.a.w. bij integratie langs een karakteristiek hebben we geen last van afgeleiden in een andere richting dan de richting van de kromme.

Als de beginkromme C samenvalt met een karakteristiek van de differentiaalvergelijking dan heeft het beginwaarde-probleem geen oplossing als langs C niet voldaan is aan (7.6) en oneindig veel oplossingen als langs C wel voldaan is aan (7.6).

Voorbeeld

De vergelijking van de trillende snaar

$$\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = 0.$$

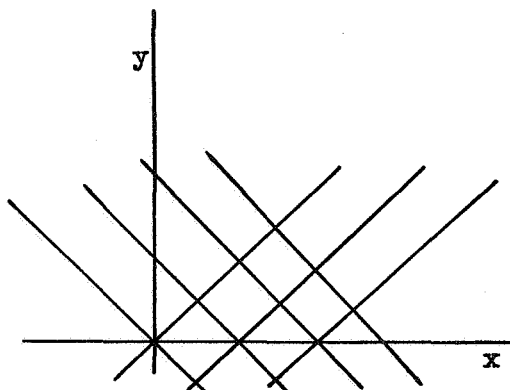
De karakteristieke vergelijking is $(dy)^2 - (dx)^2 = 0$ met de oplossingen

$\frac{dy}{dx} = \pm 1$. De karakteristieken zijn dus

$$y = x + \alpha, \quad y = -x + \alpha.$$

De karakteristieke voorwaarde is

$$-dp \cdot dy + dq \cdot dx = 0.$$



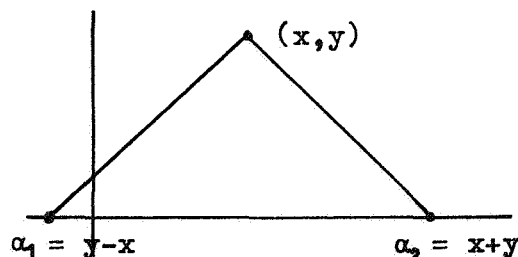
Langs de karakteristiek $y = x + \alpha_1$ geldt dus $-dp + dq = 0$ ofwel $p - q = \beta_1$ en langs de karakteristiek $y = -x + \alpha_2$ geldt $dp + dq = 0$ ofwel $p + q = \beta_2$.

Nemen we nu als beginkromme C de x-as en hierop de beginvoorwaarden $u(x,0) = 0$, $p(x,0) = 0$, $q(x,0) = \sin x$. In het punt (x,y) geldt dan

$$p + q = \beta_2 = \sin \alpha_2 = \sin (x+y)$$

$$p - q = \beta_1 = \sin \alpha_1 = \sin (y-x)$$

Hieruit volgt $p = \sin y \cdot \cos x$ en $q = \sin x \cdot \cos y$ en dus is de oplossing $u(x,y) = \sin x \cdot \sin y$.



Als we voor C de karakteristiek $y = x$ nemen dan moeten de beginvoorwaarden op deze kromme voldoen aan $p - q = \beta$, opdat het gestelde probleem een oplossing zal bezitten. In punten buiten de beginkromme is de oplossing dan niet eenduidig bepaald, terwijl bovendien de tweede afgeleiden discontinu kunnen zijn langs deze beginkromme $y = x$. Nemen we bijvoorbeeld de beginvoorwaarden

$$u(x,x) = 1, \quad p(x,x) = 1, \quad q(x,x) = -1$$

dan voldoet de functie

$$u(x,y) = 1 + (x-y) + \alpha(x-y)^2$$

voor iedere keuze van α aan de differentiaalvergelijking en de beginvoorwaarden. De discontinuïteit treedt op als we links en rechts van $y = x$ verschillende waarden voor α nemen.

In het geval van de partiële differentiaalvergelijkingen van de eerste orde waren de karakteristieken steeds reëel. Dit is bij de tweede-orde partiële differentiaalvergelijkingen niet altijd het geval. We beschouwen daartoe nog eens de karakteristieke vergelijking (7.5) die we schrijven in de vorm

$$a\left(\frac{dy}{dx}\right)^2 - b\frac{dy}{dx} + c = 0.$$

Deze vergelijking heeft in het algemeen twee oplossingen $\frac{dy}{dx} = \lambda_1$ en $\frac{dy}{dx} = \lambda_2$, en levert dan twee karakteristieken-richtingen in ieder punt (x,y) . Deze richtingen kunnen zijn toegevoegd complex, samenvallend reëel of reëel en verschillend. Op grond hiervan verdelen we de quasi-lineaire tweede-orde partiële differentiaalvergelijkingen in de volgende drie typen:

| | | |
|---------------------|-----------------|--|
| <u>hyperbolisch</u> | $b^2 - 4ac > 0$ | twee reële niet-samenvallende karakteristieken-richtingen, |
| <u>parabolisch</u> | $b^2 - 4ac = 0$ | reële samenvallende karakteristieken-richtingen, |
| <u>elliptisch</u> | $b^2 - 4ac < 0$ | twee complexe karakteristieken-richtingen. |

Het bestaan van reële niet samenvallende karakteristieken-richtingen is essentieel voor de mogelijkheid van een numerieke stap voor stap methode voor de oplossing van een beginwaarde-probleem. Deze methode is dus alleen mogelijk voor de hyperbolische differentiaalvergelijkingen. Voor het oplossen van de elliptische en parabolische differentiaalvergelijkingen zullen we andere oplossingsmethoden moeten gebruiken.

Opmerking. Als de coëfficiënten van (7.1) niet constant zijn is het mogelijk dat eenzelfde differentiaalvergelijking in verschillende gebieden van het (x,y) -vlak van verschillend type is. Zo is bijvoorbeeld de differentiaalvergelijking

$$\frac{\partial^2 u}{\partial x^2} + (x^2 - y^2) \frac{\partial^2 u}{\partial y^2} = F(u, p, q, x, y)$$

hyperbolisch als $x^2 < y^2$, elliptisch als $x^2 > y^2$ en parabolisch als $x^2 = y^2$ (als tenminste q in het rechterlid voorkomt).

Het is zelfs mogelijk, als het type van de vergelijking ook afhangt van de oplossing, dat in een bepaald gebied de vergelijking van type verandert bij verandering van de oplossing, dus van de beginvoorwaarden.

Als representant van de hyperbolische partiële differentiaalvergelijking neemt men meestal de vergelijking van de trillende snaar: $u_{xx} - u_{yy} = 0$, voor de parabolische vergelijking neemt men de warmtegeleidingsvergelijking: $u_{xx} = u_y$ en voor de elliptische vergelijking de potentiaalvergelijking: $u_{xx} + u_{yy} = 0$. We zullen nu laten zien dat deze vergelijkingen niet alleen speciale, eenvoudige, gevallen van de drie verschillende typen zijn, maar dat een willekeurige vergelijking zich ook essentieel gedraagt als een van de representanten. We zullen namelijk aantonen dat de algemene quasi-lineaire partiële differentiaalvergelijking van de tweede orde

$$au_{xx} + bu_{xy} + cu_{yy} = e \quad (7.7)$$

door een coördinatentransformatie te herleiden is tot een van de genoemde representanten.

1. hyperbolische geval: $b^2 - 4ac > 0$

In dit geval heeft de karakteristieke vergelijking (7.5) twee reële oplossingen

$$\frac{dy}{dx} = \frac{b + \sqrt{b^2 - 4ac}}{2a}, \quad \frac{dy}{dx} = \frac{b - \sqrt{b^2 - 4ac}}{2a}. \quad (7.8)$$

De coëfficiënten a, b en c zijn gegeven functies van x, y, u, p en q en mogen dus beschouwd worden als functies van x en y , waarbij we dan veronderstellen dat de oplossing $u(x, y)$ bekend is. De vergelijkingen (7.8) kunnen dan worden opgevat als gewone differentiaalvergelijkingen, waarvan de algemene oplossingen, in impliciete vorm geschreven, respectievelijk zijn

$$\varphi(x, y) = \lambda, \quad \phi(x, y) = \mu \quad (7.9)$$

met λ en μ willekeurige constanten. Ieder van deze relaties definieert in het (x, y) -vlak een schaar krommen (karakteristieken). Deze krommen nemen we nu als parameterkrommen in een nieuw coördinatensysteem.

We voeren dus nieuwe coördinaten ξ en η in, gedefinieerd door

$$\xi = \varphi(x, y), \quad \eta = \phi(x, y).$$

(7.10)

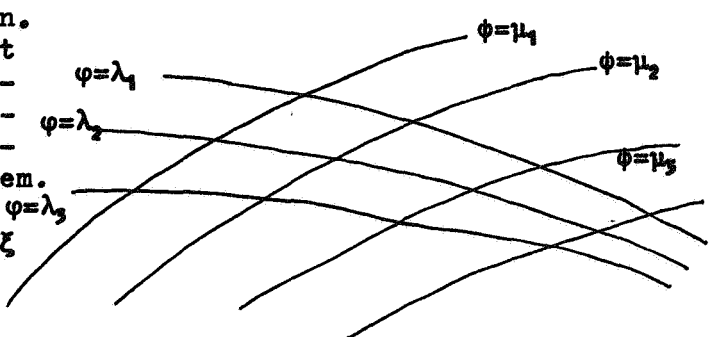
De eerste en tweede afgeleiden van u naar x en y worden in dit coördinatensysteem:

$$\begin{aligned} u_x &= u_\xi \xi_x + u_\eta \eta_x \\ u_y &= u_\xi \xi_y + u_\eta \eta_y \\ u_{xx} &= u_{\xi\xi} \xi_x^2 + 2u_{\xi\eta} \xi_x \eta_x + u_{\eta\eta} \eta_x^2 + u_\xi \xi_{xx} + u_\eta \eta_{xx} \\ u_{yy} &= u_{\xi\xi} \xi_y^2 + 2u_{\xi\eta} \xi_y \eta_y + u_{\eta\eta} \eta_y^2 + u_\xi \xi_{yy} + u_\eta \eta_{yy} \\ u_{xy} &= u_{\xi\xi} \xi_x \xi_y + u_{\xi\eta} (\xi_x \eta_y + \xi_y \eta_x) + u_{\eta\eta} \eta_x \eta_y + u_\xi \xi_{xy} + u_\eta \eta_{xy}. \end{aligned} \quad (7.11)$$

Invullen van deze uitdrukkingen in de differentiaalvergelijking (7.7) levert dan deze vergelijking in de nieuwe coördinaten:

$$\bar{a} u_{\xi\xi} + \bar{b} u_{\xi\eta} + \bar{c} u_{\eta\eta} = \bar{e}$$

waarin de coëfficiënten zijn:



$$\begin{aligned}
\bar{a} &= a\zeta_x^2 + b\zeta_x\zeta_y + c\zeta_y^2 \\
\bar{b} &= 2a\zeta_x\eta_x + b(\zeta_x\eta_y + \zeta_y\eta_x) + 2c\zeta_y\eta_y \\
\bar{c} &= a\eta_x^2 + b\eta_x\eta_y + c\eta_y^2 \\
\bar{e} &= e - a(u_\zeta\zeta_{xx} + u_\eta\eta_{xx}) - b(u_\zeta\zeta_{xy} + u_\eta\eta_{xy}) - c(u_\zeta\zeta_{yy} + u_\eta\eta_{yy}).
\end{aligned}
\tag{7.12}$$

Nu geldt in een willekeurig punt van de parameterkrommen $\varphi(x,y) = \lambda$

$$\frac{dy}{dx} = -\frac{\varphi_x}{\varphi_y} = -\frac{\zeta_x}{\zeta_y}$$

en dus volgens (7.8)

$$\frac{\zeta_x}{\zeta_y} = -\frac{b + \sqrt{b^2 - 4ac}}{2a}$$

en hieruit volgt dat

$$\bar{a} = a\zeta_x^2 + b\zeta_x\zeta_y + c\zeta_y^2 = 0.$$

Op geheel analoge wijze kunnen we afleiden dat $\bar{c} = 0$

Met de nieuwe coördinaten ζ en η wordt de differentiaalvergelijking (7.7) dus van de gedaante

$$\bar{b}u_{\zeta\eta} = \bar{e} \quad \text{ofwel} \quad u_{\zeta\eta} = \frac{\bar{e}}{\bar{b}}. \tag{7.13}$$

Het is gemakkelijk na te gaan dat in het rechterlid $\frac{\bar{e}}{\bar{b}}$ van (7.13) uitsluitend de grootheden ζ , η , u , u_ζ en u_η voorkomen.

De vergelijking (7.13) heet de normaalvorm van de tweede-orde partiële differentiaalvergelijking van het hyperbolische type. Dikwijls echter wordt een andere normaalvorm gebruikt. Deze krijgen we door over te gaan op nog andere coördinaten α en β door middel van de transformatie

$$\begin{aligned}
\zeta &= \frac{1}{2}(\alpha + \beta) & \alpha &= \zeta + \eta \\
\eta &= \frac{1}{2}(\alpha - \beta) & \beta &= \zeta - \eta.
\end{aligned}
\tag{7.14}$$

Dan geldt

$$u_\zeta = u_\alpha + u_\beta \quad u_\eta = u_\alpha - u_\beta \quad u_{\zeta\eta} = u_{\alpha\alpha} - u_{\beta\beta}$$

en de normaalvorm wordt dan

$$u_{\alpha\alpha} - u_{\beta\beta} = f \quad \text{waarbij } f = \frac{\bar{e}}{\bar{b}} \tag{7.15}$$

De meest elementaire vorm hiervan is de homogene vergelijking $u_{\alpha\alpha} - u_{\beta\beta} = 0$, welke juist de eerder genoemde vergelijking van de trillende snaar is.

2. parabolische geval: $b^2 - 4ac = 0$

In dit geval heeft de karakteristieke vergelijking (7.5) slechts één oplossing

$$\frac{dy}{dx} = \frac{b}{2a} \quad (7.16)$$

welke aanleiding geeft tot slechts één schaar parameterkrommen $\varphi(x,y) = \lambda$. Om de tweede schaar parameterkrommen te krijgen kiezen we een willekeurige, van $\varphi(x,y)$ onafhankelijke, functie $\phi(x,y)$. Dan kunnen we weer nieuwe coördinaten ξ en η invoeren door de relaties

$$\xi = \varphi(x,y) \quad \eta = \phi(x,y).$$

Omdat $\varphi(x,y) = \lambda$ ook nu oplossing is van de karakteristieke vergelijking geldt, evenals in het hyperbolische geval, dat $\bar{a} = 0$; maar nu is $c \neq 0$ daar $\phi(x,y)$ een willekeurige functie is.

Uit $b^2 - 4ac = 0$ volgt

$$\bar{a} = a\xi_x^2 + b\xi_x\xi_y + c\xi_y^2 = (\sqrt{a}\xi_x + \sqrt{c}\xi_y)^2 = 0$$

en hieruit vinden we voor de coefficient \bar{b} eenvoudig

$$\begin{aligned} \bar{b} &= a\xi_x\eta_x + b(\xi_x\eta_y + \xi_y\eta_x) + c\xi_y\eta_y = \\ &= (\sqrt{a}\xi_x + \sqrt{c}\xi_y)(\sqrt{a}\eta_x + \sqrt{c}\eta_y) = 0. \end{aligned}$$

Met de coördinaten ξ en η krijgt de differentiaalvergelijking (7.7) nu dus de normaalvorm

$$u_{\eta\eta} = \frac{\bar{e}}{c} = f(\xi, \eta, u, u_\xi, u_\eta). \quad (7.17)$$

Als in het rechterlid f de grootheid u_ξ niet voorkomt, dan is de differentiaalvergelijking een gewone differentiaalvergelijking, met daarin ξ als parameter.

De meest eenvoudige vorm van een parabolische differentiaalvergelijking is dus $u_{\eta\eta} = u_\xi$, hetgeen juist de warmtegeleidingsvergelijking is.

3. elliptische geval: $b^2 - 4ac < 0$

Dit geval kan op precies dezelfde wijze behandeld worden als het hyperbolische geval, met dit verschil dat de rechterleden van (7.8), en dus ook de functies $\varphi(x,y)$ en $\phi(x,y)$, toegevoegd complex zijn. Daardoor zijn de coördinaten ξ en η toegevoegd complex. We kunnen nu overgaan op

de reële coördinaten α en β door de transformatie

$$\alpha = \frac{\xi + \eta}{2} \quad \beta = \frac{\xi - \eta}{2i}$$

en vinden dan de normaalvorm

$$u_{\alpha\alpha} + u_{\beta\beta} = f(\alpha, \beta, u, u_{\alpha}, u_{\beta}) \quad (7.18)$$

Als eenvoudigste geval van een elliptische differentiaalvergelijking vinden we dus de potentiaalvergelijking: $u_{\alpha\alpha} + u_{\beta\beta} = 0$.

We behandelen nu het numeriek oplossen van een hyperbolische differentiaalvergelijking met behulp van de karakteristieken.

Als de differentiaalvergelijking (7.7) hyperbolisch is, dan heeft de karakteristieke vergelijking twee reële verschillende oplossingen die we nu schrijven in de vorm

$$dy = f dx \quad dy = g dx \quad (7.19)$$

waarbij f en g de wortels zijn van de vierkantsvergelijking

$$a \lambda^2 - b \lambda + c = 0.$$

Door ieder punt (x, y) gaan twee karakteristieken waarvan de richtingen gegeven worden door (7.19). We noemen deze karakteristieken respectievelijk: de f -karakteristiek en de g -karakteristiek. Omdat langs iedere karakteristiek de karakteristieke voorwaarde (7.6) moet gelden, geldt in het bijzonder langs de f -karakteristiek

$$e dy - a f dp - c dq = 0 \quad (7.20)$$

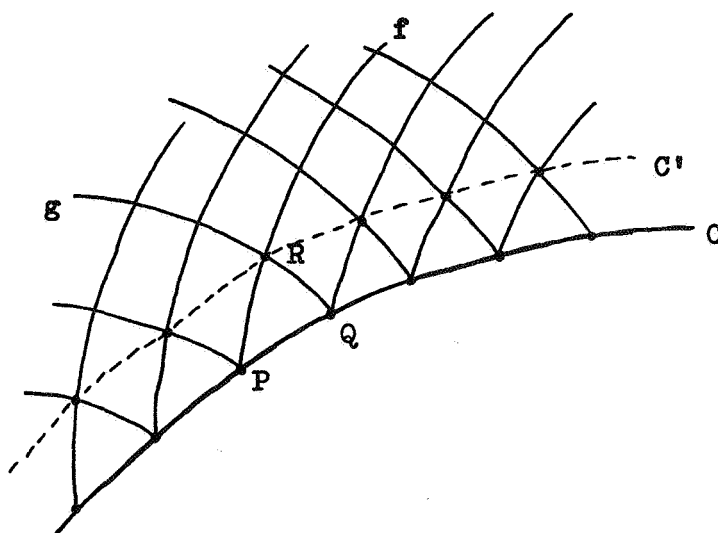
en langs de g -karakteristiek

$$e dy - a g dp - c dq = 0 \quad (7.21)$$

We veronderstellen nu dat op de beginkromme C de waarden van u , p en q gegeven zijn, waarbij voldaan is aan de relatie

$$du = p dx + q dy. \quad (7.22)$$

Nu beschouwen we op C twee naburige punten $P(x_0, y_0)$ en $Q(x_1, y_1)$. Het snijpunt van de f -karakteristiek door P en de g -karakteristiek door Q noemen we $R(x, y)$.



In P zijn bekend: x_0, y_0, u_0, p_0 en q_0 ; in Q zijn bekend x_1, y_1, u_1, p_1 en q_1 . Met deze gegevens moeten we in R x, y, u, p en q berekenen. Hiervan worden x en y bepaald door de karakteristieken (7.19); p en q worden bepaald door de karakteristieke voorwaarden (7.20) en (7.21) terwijl u bepaald wordt door (7.22).

We nemen voor deze vijf vergelijkingen de volgende benaderingsformules

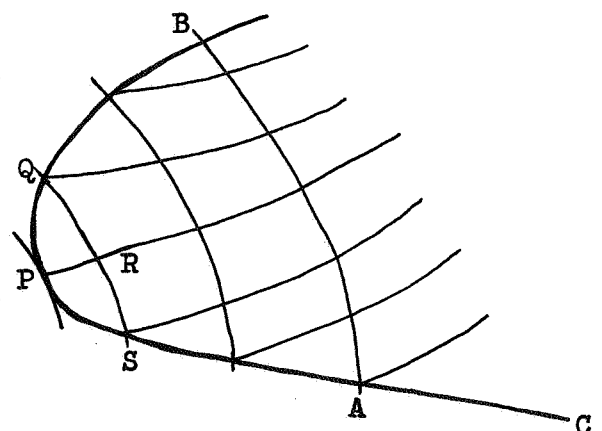
$$\begin{aligned} y - y_0 &= \frac{1}{2}(f + f_0)(x - x_0) \\ y - y_1 &= \frac{1}{2}(g + g_1)(x - x_1) \\ \frac{1}{2}(e + e_0)(y - y_0) - \frac{1}{2}(af + a_0 f_0)(p - p_0) - \frac{1}{2}(c + c_0)(q - q_0) &= 0 \\ \frac{1}{2}(e + e_1)(y - y_1) - \frac{1}{2}(ag + a_1 g_1)(p - p_1) - \frac{1}{2}(c + c_1)(q - q_1) &= 0 \\ u - u_0 &= \frac{1}{2}(p + p_0)(x - x_0) + \frac{1}{2}(q + q_0)(y - y_0). \end{aligned} \quad (7.23)$$

We hebben nu een stelsel van vijf niet-lineaire vergelijkingen voor de vijf onbekenden x, y, u, p en q . Dit stelsel kunnen we oplossen met het volgende iteratieproces. Als beginschatting voor x, y, u, p en q kunnen we bijvoorbeeld de overeenkomstige waarden in P nemen. Hiermee berekenen we de coëfficiënten a, b, c en e en vervolgens f en g . Deze waarden vullen we in het stelsel (7.23) in en lossen dan het (nu lineaire) stelsel op. De oplossing is een nieuwe schatting voor x, y, u, p en q waarmee we het proces herhalen. Dit iteratieproces zal in het algemeen snel convergeren als de afstand tussen P en Q klein is.

We zien dus dat uit de gegeven waarden van u, p en q op de beginkromme C, de waarden van deze grootheden op een naburige kromme C' berekend kunnen worden. We kunnen dan C' als nieuwe beginkromme nemen en het proces herhalen. Het is nu ook direct duidelijk dat, als de beginkromme eindig is, de oplossing van het beginwaardeprobleem alleen bepaald is in het gebied begrensd door de karakteristieken die gaan door de eindpunten van C.

We onderzoeken nu nog het speciale geval dat de beginkromme C eenzelfde schaar karakteristieken tweemaal snijdt.

Het gedeelte PSA van C kan gewoon als beginkromme beschouwd worden. Geven we hierop u, p en q die voldoen aan (7.22) dan kunnen we de waarden van u, p en q vinden in het gebied dat begrensd wordt door de karakteristiek door P en R. Om de waarden van u, p en q te vinden het overige gebied binnen C hebben we voldoende aan een relatie tussen p en q of de waarde van u (waarmee in feite met (7.22) ook een relatie tussen p en q gegeven is) op PQB. Want de karakteristieke voorwaarde langs de karak-



teristiek QR geeft een relatie tussen p en q in Q en de bekende p en q in R . Een tweede relatie tussen p en q in Q is dus voldoende om hun waarde eenduidig te bepalen. We noemen het gedeelte PQB dan ook liever geen beginkromme maar rand met een randvoorwaarde.

Voorbeeld

De half oneindige trillende snaar.

De vergelijking is $u_{xx} - u_{tt} = 0$.

De beginvoorwaarden zijn u en $q = \frac{\partial u}{\partial t}$ gegeven voor $t = 0$ en $x \geq 0$.

Deze leggen de beginstand en de beginsnelheid van de snaar vast. Als randvoorwaarde geven we u op de lijn $x = 0$ en $t \geq 0$.

Hiermee is de uitwijking van het eindpunt op ieder tijdstip vastgelegd.

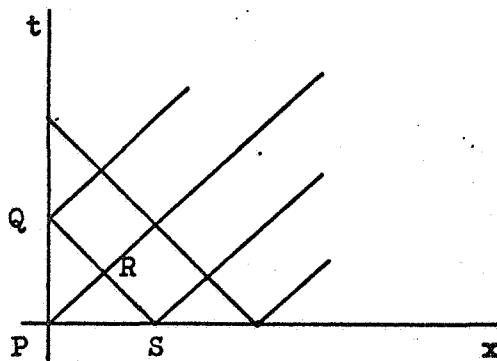
De karakteristieken met bijbehorende relaties zijn

$$t - x = \alpha_1 \quad \text{met } p - q = \beta_1$$

$$t + x = \alpha_2 \quad \text{met } p + q = \beta_2$$

In het punt R bepalen we de grootheden u, p en q op de normale wijze uit de beginvoorwaarden op de x -as. Langs de karakteristiek RQ geldt nu de relatie $p + q = \beta$, waarbij β gelijk is aan $p + q$ in R .

We kennen u en daarmee ook q langs de t -as, in het bijzonder in het punt Q . Dan kunnen we uit $p + q = \beta$ de waarde van p in Q direct berekenen. We zijn dus in staat de oplossing $u = u(x, t)$ te bepalen in het gehele eerste kwadrant.



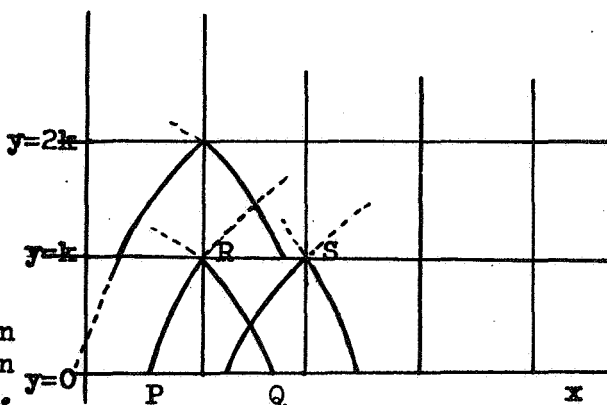
Als de karakteristieke richtingen (7.19) afhangen van de coördinaten x en y en eventueel ook nog van de onbekende functie u , dan zullen de karakteristieken gekromd zijn. Het gevolg is dat de snijpunten van de f - en g -karakteristieken niet aequidistant in x en y zullen liggen. Het kan echter zijn dat de oplossing van het probleem gewenst wordt op een rechthoekig net met constante maaswijdte. Het is duidelijk dat dan achteraf nog veel rekenwerk nodig is om de oplossing, met behulp van interpolatieformules, in de gewenste punten te bepalen.

We kunnen deze moeilijkheid op de volgende manier vermijden.

Veronderstel, bijvoorbeeld, dat de beginkromme juist de x -as is, en dat de oplossing u moet worden bepaald in aequidistant gelegen punten op de rechten $y = n.k$ voor $n = 1, 2, 3, \dots$.

We veronderstellen nu dat de punten RS... aequidistant gelegen zijn op de rechte $y = k$. De twee karakteristieken door R snijden de rechte $y = 0$ in de punten P en Q. Het verband tussen de grootheden x, y, u, p en q in de punten P, Q en R wordt gegeven door het stelsel vergelijkingen (7.23).

In dit geval zijn de coördinaten x en y van R bekend, maar moeten we x_0 van P en x_1 van Q bepalen.



Samen met u, p en q zijn dit dus weer vijf onbekenden, die we door iteratie uit het stelsel (7.23) kunnen bepalen. De enige extra complicatie is dat we u_0, p_0 en q_0 in P en u_1, p_1 en q_1 in Q bij iedere iteratiestap opnieuw door interpolatie moeten bepalen. We zien dus dat de interpolatie achteraf vervangen is door interpolatie tijdens het rekenproces.

Hebben we de waarden van u, p en q bepaald op de rechte $y = k$ dan kunnen we daarmee het proces voortzetten op de rechte $y = 2k$, etc. Het is natuurlijk noodzakelijk, dat alle punten van het rechthoekige net liggen binnen het gebied begrensd door de eindkarakteristieken. Het is ook duidelijk dat de maaswijdte in de y -richting niet noodzakelijk constant behoeft te zijn.

We beschouwen nu nog het geval van een simultaan systeem van twee eerste-orde quasi-lineaire vergelijkingen:

$$a \frac{\partial u}{\partial x} + b \frac{\partial u}{\partial y} + e \frac{\partial v}{\partial x} + f \frac{\partial v}{\partial y} = g, \quad (7.24)$$

$$A \frac{\partial u}{\partial x} + B \frac{\partial u}{\partial y} + E \frac{\partial v}{\partial x} + F \frac{\partial v}{\partial y} = G.$$

waarin de coëfficiënten functies zijn van x, y, u en v , doch niet afhangen van de afgeleiden van u en v .

We veronderstellen nu dat op de beginkromme C de functies $u(x, y)$ en $v(x, y)$ gegeven zijn.

Langs een willekeurige kromme moeten u en v ook voldoen aan de betrekkingen:

$$\begin{aligned} du &= \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy \\ dv &= \frac{\partial v}{\partial x} dx + \frac{\partial v}{\partial y} dy. \end{aligned} \quad (7.25)$$

Op analoge wijze als voorheen volgt dat de oplossing van het beginwaardeprobleem éénduidig bepaald is, tenzij langs de beginkromme C , de coëfficiëntendeterminant van het stelsel (7.24) en (7.25) verdwijnt, m.a.w.

$$\begin{vmatrix} dx & dy & 0 & 0 \\ 0 & 0 & dx & dy \\ a & b & e & f \\ A & B & E & F \end{vmatrix} = 0$$

ofwel

$$(eA - aE) \cdot (dy)^2 + (bE - eB + aF - fA) \cdot dx \cdot dy + (fB - bF) \cdot (dx)^2 = 0.$$

We noemen deze vergelijking weer de karakteristieke vergelijking. We vinden in het algemeen dus weer twee scharen karakteristieken. Als de karakteristieken reëel zijn noemen we het stelsel hyperbolisch. Ook nu vinden we karakteristieke voorwaarden waaraan langs de karakteristieken moet zijn voldaan. Bijvoorbeeld

$$\begin{vmatrix} du & dx & dy & 0 \\ dv & 0 & 0 & dx \\ g & a & b & e \\ G & A & B & E \end{vmatrix} = 0.$$

Deze karakteristieke voorwaarde kan men ook nu weer gebruiken om de oplossing te bepalen langs de karakteristieken. Door ieder punt gaan twee karakteristieken, en de numerieke oplossingsmethode is bijna identiek met die van de tweede-orde hyperbolische vergelijking.

Opmerking.

Analoog als bij de gewone differentiaalvergelijkingen is een quasi-lineaire tweede-orde partiële differentiaalvergelijking altijd te schrijven als een stelsel van twee quasi-lineaire eerste-orde partiële differentiaalvergelijkingen.

Stellen we namelijk

$$U = \frac{\partial u}{\partial x} \quad V = \frac{\partial u}{\partial y}$$

dan gaat de vergelijking (7.7) over in het stelsel

$$aU_x + bU_y + cV_y = e$$

$$U_y - V_x = 0.$$

8. Differentiemethoden voor hyperbolische differentiaalvergelijkingen

Bij een hyperbolische differentiaalvergelijking kunnen we naast de behandelde numerieke oplossingsmethode met behulp van de karakteristieken, ook gebruik maken van eindige differentiemethoden. Met een dergelijke methode vinden we de oplossing van het begin- of randwaardeprobleem dan in de punten van een rechthoekig net.

We beschouwen het standaardtype voor de hyperbolische differentiaalvergelijkingen, namelijk de snaarvergelijking

$$u_{xx} - u_{tt} = 0 \quad t \geq 0 \quad (8.1)$$

met de beginvoorwaarden

$$u(x,0) = f(x) \quad u_t(x,0) = g(x). \quad (8.2)$$

Men kan gemakkelijk verifiëren dat voor iedere tweemaal differentieerbare functie $F(x)$ geldt dat zowel $u(x,t) = F(x+t)$ als $u(x,t) = F(x-t)$ voldoet aan de differentiaalvergelijking (8.1). De algemene oplossing van de snaarvergelijking wordt dan ook gegeven door

$$u(x,t) = F(x+t) + G(x-t) \quad (8.3)$$

waarin $F(x)$ en $G(x)$ willekeurige functies mogen zijn. Deze oplossing voldoet aan de beginvoorwaarden (8.2) als

$$\begin{aligned} u(x,0) &= F(x) + G(x) = f(x), \\ u_t(x,0) &= F'(x) - G'(x) = g(x). \end{aligned}$$

Hieruit volgt dat de oplossing van de differentiaalvergelijking die voldoet aan de beginvoorwaarden, is:

$$u(x,t) = \frac{1}{2} \{ f(x+t) + f(x-t) + \int_{x-t}^{x+t} g(\xi) d\xi \} \quad (8.4)$$

(formule van d'Alembert).

Opmerking

Deze formule laat nog eens duidelijk zien dat de oplossing in het punt (x,t) alleen afhangt van de beginwaarden op de x -as tussen de punten $x-t$ en $x+t$; dit zijn juist de snijpunten van de beginkromme met de twee karakteristieken door (x,t) .

We nemen een constante stapgrootte h in de x -richting, en een constante stapgrootte k in de t -richting. Dan geldt

$$\frac{\partial^2 u}{\partial x^2} = \frac{u(x-h,t) - 2u(x,t) + u(x+h,t)}{h^2} + \mathcal{O}(h^2) \quad (8.5)$$

$$\frac{\partial^2 u}{\partial t^2} = \frac{u(x,t-k) - 2u(x,t) + u(x,t+k)}{k^2} + \mathcal{O}(k^2).$$

We vervangen nu in de vergelijking (8.1) de partiële afgeleiden door de centrale differenties, en krijgen dan als benadering de differentieformule

$$\frac{U(x-h,t) - 2U(x,t) + U(x+h,t)}{h^2} = \frac{U(x,t-k) - 2U(x,t) + U(x,t+k)}{k^2}. \quad (8.6)$$

Voor de benadering van de afgeleide in de tweede beginvoorwaarde moeten we gebruik maken van de voorwaartse differentie, waarvoor geldt

$$\frac{\partial u}{\partial t} = \frac{u(x,t+k) - u(x,t)}{k} + \sigma(k). \quad (8.7)$$

De beginvoorwaarden voor de differentieformule (8.6) worden dan

$$U(x,0) = f(x), \quad \frac{U(x,k) - U(x,0)}{k} = g(x). \quad (8.8)$$

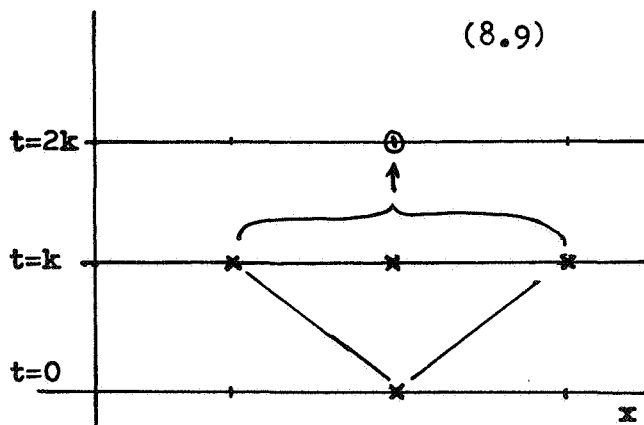
De numerieke oplossing van de differentievergelijking (8.6) is in dit geval erg eenvoudig. De beginvoorwaarden (8.8) die we ook kunnen schrijven als

$$U(x,0) = f(x), \quad U(x,k) = kg(x) + f(x)$$

geven de waarden van $U(x,t)$ op de beginrechten $t = 0$ en $t = k$, en door de differentieformule (8.6) te schrijven in de expliciete vorm

$$U(x,t+k) = 2U(x,t) - U(x,t-k) + \lambda^2 \{U(x+h,t) - 2U(x,t) + U(x-h,t)\} \quad (8.9)$$

waarbij $\lambda = \frac{k}{h}$, zien we dat uit de waarden van $U(x,t)$ op twee opeenvolgende rechten: $t = (m-1)k$ en $t = mk$, de waarden van $U(x,t)$ op de daaropvolgende rechte $t = (m+1)k$ berekend kunnen worden.



Uitgaande van de beginvoorwaarden op een interval (a,b) op de x -as kunnen we de waarden van $U(x,t)$ berekenen in de roosterpunten op en binnen de driehoek begrensd door het interval (a,b) en de twee zijden met helling $\pm \lambda$, door de eindpunten van (a,b) . In het bijzonder hangt de waarde van $U(x,t)$ in de top van de driehoek af van alle punten op de basis (a,b) .

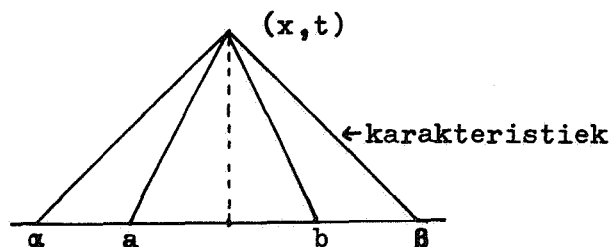
Opmerkelijk is hier de analogie met de karakteristieken-driehoek (zie opmerking blz. NA-88). Er is echter een belangrijk verschil. De karakteristieken-driehoek ligt vast en daarmee ook de punten op de begin-kromme die de waarde van $u(x,t)$ bepalen. De invloedsdriehoek bij het differentieprobleem, dus ook welke punten op de beginkromme de waarde van $U(x,t)$ bepalen, hangt af van de keuze van $\lambda = \frac{k}{h}$. In het differentieprobleem is deze λ nog vrij te kiezen.

Nu geeft de differentiemethode alleen dan een bruikbare benadering van de oplossing van de differentiaalvergelijking als voor $h,k \rightarrow 0$ de oplossing $U(x,t)$ convergeert naar $u(x,t)$.

Voor $h,k \rightarrow 0$ gaat de differentievergelijking (8.6) met bijbehorende beginvoorwaarden (8.8) over in de differentiaalvergelijking (8.1) met beginvoorwaarden (8.2). Dit betekent echter in het geheel niet dat $U(x,t)$ dan ook tot $u(x,t)$ nadert.

We zullen bewijzen dat de convergentie van $U(x,t)$ naar $u(x,t)$ gegarandeerd is als $\lambda \leq 1$, en dat dit niet het geval is als $\lambda > 1$.

Kiezen we $\lambda_0 > 1$, dan ligt het afhankelijkheidsinterval (a,b) van de differentievergelijking binnen het afhankelijkheidsinterval (α,β) van de differentiaalvergelijking.



We laten h en k naar nul naderen, zodanig dat $\lambda \geq \lambda_0$. Veronderstel nu dat voor zekere beginfuncties $f(x)$

en $g(x)$ de rij oplossingen $U(x,t)$ convergeert naar $u(x,t)$. Veranderen we dan de beginfuncties op de intervallen (α,a) en (b,β) , dan verandert wel de waarde van $u(x,t)$, maar niet de waarde van de oplossingen $U(x,t)$. Voor het aldus gewijzigde probleem zal de oplossing $u(x,t)$ dan verschillend zijn van de limietwaarde van $U(x,t)$ voor h en $k \rightarrow 0$.

De convergentie van $U(x,t)$ naar $u(x,t)$ moet dus beschouwd worden als een uitzondering indien de verhouding $\lambda = \frac{k}{h} > 1$.

Convergentie-onderzoek voor $\lambda < 1$

Teneinde voor het geval $\lambda < 1$ de convergentie van $U(x,t)$ naar $u(x,t)$ voor $h,k \rightarrow 0$ te onderzoeken, zullen we de oplossing $U(x,t)$ in gesloten vorm trachten te bepalen. We doen dit met de methode van de scheiding van variabelen (separatie).

We zullen met deze methode eerst de oplossing van de differentiaalvergelijking bepalen.

Van de snaarvergelijking $u_{xx} - u_{tt} = 0$ zoeken we een oplossing die te schrijven is als

$$u(x,t) = \varphi(x) \cdot \phi(t). \quad (8.10)$$

Substitutie van deze functie in de differentiaalvergelijking geeft,

$$\frac{\varphi''(x)}{\varphi(x)} = \frac{\phi''(t)}{\phi(t)}. \quad (8.11)$$

Het linkerlid van deze gelijkheid is een functie van x alleen, terwijl het rechterlid uitsluitend van t afhangt. Dit is alleen mogelijk als beide leden gelijk zijn aan een constante. Noemen we deze constante $-\alpha^2$, waarbij α willekeurig is, dan krijgen we twee gewone differentiaalvergelijkingen

$$\begin{aligned} \varphi''(x) + \alpha^2 \varphi(x) &= 0 \\ \phi''(t) + \alpha^2 \phi(t) &= 0. \end{aligned} \quad (8.12)$$

Deze twee vergelijkingen kunnen we algemeen oplossen. Substitueren we deze algemene oplossingen in (8.10) krijgen we bij iedere keuze van α een speciale oplossing van de snaarvergelijking, nl.

$$u_\alpha(x, t) = (A \sin \alpha x + B \cos \alpha x)(C \sin \alpha t + D \cos \alpha t). \quad (8.13)$$

Met behulp van deze oplossingen zoeken we de oplossing die voldoet aan de begin- en randvoorwaarden van de snaarvergelijking. We beschouwen hier het geval dat de snaar wel een beginuitwijking heeft, maar dat de beginsnelheid nul is. Nemen we voor het gemak aan dat de eindpunten van de snaar $x = 0$ en $x = \pi$ zijn, dan zijn de beginvoorwaarden

$$\begin{aligned} u(x, 0) &= f(x) & 0 \leq x \leq \pi \\ u_t(x, 0) &= 0 & 0 \leq x \leq \pi \end{aligned} \quad (8.14)$$

en de randvoorwaarden

$$u(0, t) = u(\pi, t) = 0 \quad t \geq 0. \quad (8.15)$$

We ontwikkelen nu de functie $f(x)$ in een Fourierreeks

$$f(x) = \sum_{n=1}^{\infty} a_n \sin nx \quad (8.16)$$

met

$$a_n = \frac{2}{\pi} \int_0^\pi f(x) \sin nx \, dx. \quad (8.17)$$

We nemen aan dat de beginvoorwaarde $f(x)$ zodanig is, dat de reeks $\sum_{n=1}^{\infty} |a_n|$ convergent is (m.a.w. voldoende sterke afname van de coëfficiënten van de Fourierreeks).

Bij iedere term $a_n \sin nx$ van de Fourierreeks van $f(x)$ zoeken we een oplossing (8.13), die voldoet aan $U(x, 0) = a_n \sin nx$, en verder aan de begin- en randvoorwaarden.

Nu is $u(0, t) = 0$ als $B = 0$, en $u(\pi, t) = 0$ als we voor α een natuurlijk getal kiezen. Het ligt voor de hand dat we hier $\alpha = n$ kiezen.

$u(x, 0) = a_n \sin nx$ als we nemen $A = a_n$ en $D = 1$. Tenslotte volgt uit $u_t(x, 0) = nC \sin nx = 0$ dat $C = 0$.

De gezochte oplossing (8.13) is dus

$$u_n(x,t) = a_n \sin nx \cos nt. \quad (8.18)$$

De oplossing $u(x,t)$ van de snaarvergelijking die voldoet aan de voorwaarden (8.14) en (8.15) kan dan geschreven worden als

$$u(x,t) = \sum_{n=1}^{\infty} u_n(x,t) = \sum_{n=1}^{\infty} a_n \sin nx \cos nt. \quad (8.19)$$

Deze oplossing in de vorm van superpositie van staande golven moet overeenkomen met de oplossing

$$u(x,t) = \frac{1}{2}\{f(x+t) + f(x-t)\}$$

in de vorm van lopende golven, welke men verkrijgt door substitutie van $g(x) = 0$ in de formule van d'Alembert (8.4). Nu geldt

$$\frac{1}{2}\{f(x+t) + f(x-t)\} = \frac{1}{2} \sum_{n=1}^{\infty} a_n \{\sin n(x+t) + \sin n(x-t)\}$$

waaruit gemakkelijk volgt

$$\frac{1}{2}\{f(x+t) + f(x-t)\} = \sum_{n=1}^{\infty} a_n \sin nx \cos nt.$$

We pakken nu het overeenkomstige probleem voor de differentievergelijking op dezelfde wijze aan. Analooq als in het bovenstaande geval zoeken we nu naar oplossingen van (8.6) die te schrijven zijn als

$$U(x,t) = \Phi(x) \cdot \Psi(t). \quad (8.20)$$

Substitutie hiervan in de differentievergelijking levert

$$\Psi(t) \frac{\Phi(x+h) - 2\Phi(x) + \Phi(x-h)}{h^2} = \Phi(x) \frac{\Psi(t+k) - 2\Psi(t) + \Psi(t-k)}{k^2}$$

ofwel

$$\frac{\Phi(x+h) - 2\Phi(x) + \Phi(x-h)}{h^2 \Phi(x)} = \frac{\Psi(t+k) - 2\Psi(t) + \Psi(t-k)}{k^2 \Psi(t)}. \quad (8.21)$$

Het linkerlid hangt weer alleen van x af en het rechterlid alleen van t , dus beide leden zijn gelijk aan een constante, die we weer $-\alpha^2$ noemen. We krijgen nu twee gewone differentievergelijkingen

$$\begin{aligned} \Phi(x+h) - 2\Phi(x) + \Phi(x-h) + \alpha^2 h^2 \Phi(x) &= 0 \\ \Psi(t+k) - 2\Psi(t) + \Psi(t-k) + \alpha^2 k^2 \Psi(t) &= 0. \end{aligned} \quad (8.22)$$

Om aan de beginvoorwaarden te voldoen, kiezen we α weer zo dat $\sin nx$ oplossing is van de vergelijking voor $\Phi(x)$. Dit levert de relatie

$$\sin n(x+h) - 2\sin nx + \sin n(x-h) + \alpha^2 h^2 \sin nx = 0$$

ofwel

$$2\sin nx \cos nh - 2\sin nx + \alpha^2 h^2 \sin nx = 0$$

waaruit dan gemakkelijk volgt

$$\alpha = \frac{2}{h} \sin \frac{nh}{2}. \quad (8.23)$$

We substitueren deze waarde voor α in de differentievergelijking voor $\Psi(t)$ en krijgen dan:

$$\Psi(t+k) - 2\Psi(t) + \Psi(t-k) + 4 \frac{k^2}{h^2} \sin^2 \frac{nh}{2} \Psi(t) = 0. \quad (8.24)$$

De beginvoorwaarden van de differentievergelijking zijn

$$U(x,0) = f(x) \quad U(x,k) = k g(x) + f(x) = f(x). \quad (8.25)$$

Willen we voldoen aan de n -de term van de Fourierreeks (8.16) van de beginvoorwaarden, dan moet gelden

$$\Phi(x) \cdot \Psi(0) = a_n \sin nx$$

$$\Phi(x) \cdot \Psi(k) = a_n \sin nx.$$

Dus de beginvoorwaarden behorende bij (8.24) zijn in dit geval

$$\Psi(0) = \Psi(k) = 1. \quad (8.26)$$

We moeten nu dus een differentievergelijking met constante coëfficiënten oplossen. In analogie met het overeenkomstige geval van de differentiaalvergelijking stellen we

$$\Psi(t) = e^{i\mu t},$$

dan geldt

$$e^{i\mu(t+k)} - 2e^{i\mu t} + e^{i\mu(t-k)} + 4 \lambda^2 \sin^2 \frac{nh}{2} e^{i\mu t} = 0$$

ofwel

$$2 \cos \mu k - 2 + 4 \lambda \sin^2 \frac{nh}{2} = 0$$

ofwel

$$\sin^2 \frac{\mu k}{2} = \lambda^2 \sin^2 \frac{nh}{2}$$

ofwel

$$\sin \frac{\mu k}{2} = \pm \lambda \sin \frac{nh}{2}.$$

Voor $\lambda < 1$ levert deze vergelijking reële oplossingen voor μ . De bij n behorende oplossingen noemen we $\mu = \pm \mu_n$.

De algemene oplossing $\Psi(t)$ van de differentievergelijking (8.24) kunnen we dan schrijven in de vorm

$$\Psi(t) = C_1 e^{i\mu_n t} + C_2 e^{-i\mu_n t} \quad (8.27)$$

De beginvoorwaarden (8.26) leveren de vergelijkingen voor C_1 en C_2 :

$$\begin{aligned} C_1 + C_2 &= 1 \\ C_1 e^{i\mu_n k} + C_2 e^{-i\mu_n k} &= 1 \end{aligned}$$

met als oplossing

$$C_1 = \frac{1 - e^{-i\mu_n k}}{e^{i\mu_n k} - e^{-i\mu_n k}} \quad C_2 = \frac{e^{i\mu_n k} - 1}{e^{i\mu_n k} - e^{-i\mu_n k}}$$

Door dit in (8.27) in te vullen kan men gemakkelijk vinden dat

$$\Psi(t) = \frac{\cos \mu_n (t - \frac{k}{2})}{\cos \frac{\mu_n k}{2}} \quad (8.28)$$

Omdat $\lambda < 1$ is $\sin \frac{\mu_n k}{2} < 1$ (zie boven) en dus $\cos \frac{\mu_n k}{2} \neq 0$.

We krijgen de oplossing van de differentievergelijking, welke voldoet aan de beginvoorwaarden (8.25) door sommatie over alle n van de gevonden speciale oplossingen, dus

$$U(x, t) = \sum_{n=1}^{\infty} a_n \frac{\cos \mu_n (t - \frac{k}{2})}{\cos \frac{\mu_n k}{2}} \cdot \sin nx$$

Dat deze reeks convergeert, kunnen we op de volgende manier inzien.

$$\left| \frac{\cos \mu_n (t - \frac{k}{2})}{\cos \frac{\mu_n k}{2}} \right| \leq \frac{1}{\left| \cos \frac{\mu_n k}{2} \right|} = \frac{1}{\sqrt{1 - \sin^2 \frac{\mu_n k}{2}}} = \frac{1}{\sqrt{1 - \lambda^2 \sin^2 \frac{nh}{2}}} \leq \frac{1}{\sqrt{1 - \lambda^2}}$$

Voor de algemene term van de reeks geldt dus

$$\left| a_n \frac{\cos \mu_n (t - \frac{k}{2})}{\cos \frac{\mu_n k}{2}} \sin nx \right| \leq \frac{|a_n|}{\sqrt{1 - \lambda^2}} \quad (8.29)$$

Omdat we veronderstellen dat de reeks $\sum_{n=1}^{\infty} |a_n|$ convergent is, is de reeks voor $U(x, t)$ uniform convergent (criterium van Weierstrass).

We moeten nu nog bewijzen dat $U(x,t)$ nadert tot $u(x,t)$ als $h \rightarrow 0$, in de veronderstelling dat $\lambda = \frac{k}{h}$ een constante is, die kleiner dan 1 is. Nu was μ_n oplossing van de vergelijking

$$\sin \frac{\mu k}{2} = \lambda \sin \frac{nh}{2}$$

die we ook kunnen schrijven in de vorm

$$\frac{\sin \frac{\mu k}{2}}{\frac{nk}{2}} = \frac{\sin \frac{nh}{2}}{\frac{nh}{2}} .$$

Als $h \rightarrow 0$ gaat het rechterlid naar 1, dus ook het linkerlid gaat naar 1 als $k \rightarrow 0$. Hieruit volgt dan dat $\mu_n \rightarrow n$ voor $k \rightarrow 0$. Voor de oplossing $\Psi(t)$ (8.28) geldt dan

$$\lim_{k \rightarrow 0} \Psi(t) = \cos nt .$$

We zien dus dat iedere term uit de reeks van de oplossing $U(x,t)$ van het differentieprobleem nadert tot de overeenkomstige term uit de reeks van de oplossing $u(x,t)$ van de snaarvergelijking, indien $h, k \rightarrow 0$ bij constante $\lambda < 1$.

Volgens een stelling over uniform convergente reeksen *) geldt dan ook dat

$$\lim_{h, k \rightarrow 0} U(x,t) = u(x,t) .$$

Hiermee hebben we bewezen dat voor $\lambda < 1$ de oplossing van de differentievergelijking convergeert naar de oplossing van de differentiaalvergelijking als de stapgrootten h en $k \rightarrow 0$.

*) De gebruikte stelling luidt:

Als de reeks van functies $\sum_{n=1}^{\infty} a_n(x)$ uniform convergeert in een omgeving van $x = a$ en voor iedere n bestaat $\lim_{x \rightarrow a} a_n(x)$ dan geldt:

$$\lim_{x \rightarrow a} \sum_{n=1}^{\infty} a_n(x) = \sum_{n=1}^{\infty} \lim_{x \rightarrow a} a_n(x) .$$

Voor definitie en stellingen over uniforme convergentie zie bijv. R.Courant: Differential and Integral Calculus, Vol 1. pg 386 ev.

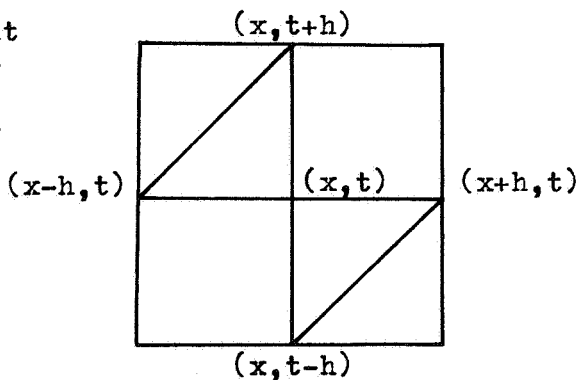
Convergentie onderzoek voor $\lambda = 1$

In het vorige geval hebben we expliciet gebruik gemaakt van het feit dat $\lambda < 1$, nl. om te bewijzen dat $\cos \frac{\mu_n^k}{2} \neq 0$. Daarom kunnen we het gegeven bewijs niet gebruiken voor het geval $\lambda = 1$. We zullen nu laten zien dat we ook in dit geval convergentie hebben.

Daar $h = k$ kunnen we de differentievergelijking vereenvoudigen tot

$$U(x+h,t) - U(x,t-h) = U(x,t+h) - U(x-h,t) \quad (8.30)$$

Volgens nevenstaande figuur betekent dit dat het verschil van twee waarden van U in een diagonaalrichting constant is in de andere diagonaalrichting.



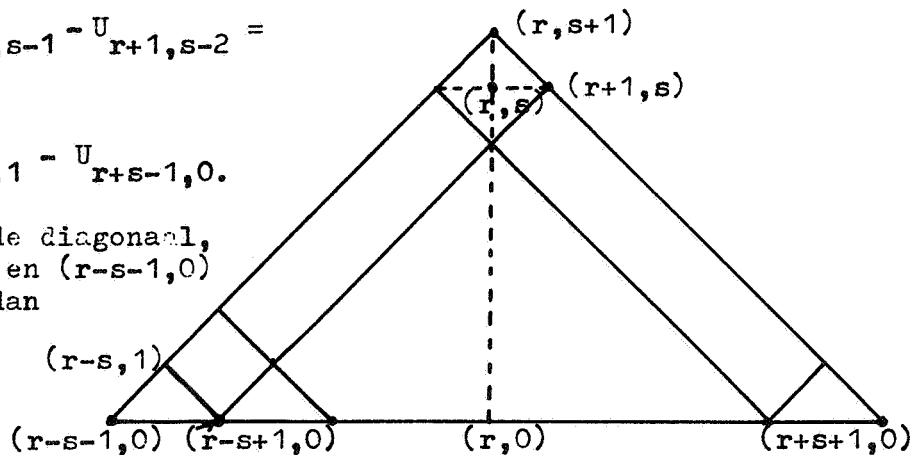
We nemen nu $x = rh$ en $t = sh$ en voeren nog de afkorting

$$U(rh,sh) = U_{r,s}$$

in, dan geldt achtereenvolgens

$$\begin{aligned} U_{r,s+1} - U_{r-1,s} &= U_{r+1,s} - U_{r,s-1} = \\ &= U_{r+2,s-1} - U_{r+1,s-2} = \\ &\dots \\ &= U_{r+s,1} - U_{r+s-1,0}. \end{aligned}$$

We sommeren nu langs de diagonaal, die de punten $(r,s+1)$ en $(r-s-1,0)$ verbindt. We krijgen dan



$$\sum_{\sigma=0}^s (U_{r-s+2\sigma,1} - U_{r-s-1+2\sigma,0}) = U_{r,s+1} - U_{r-s-1,0}.$$

In deze formule vullen we de beginvoorwaarden

$$U(x,0) = f(x) \quad U(x,h) = hg(x) + f(x)$$

in; we krijgen dan de formule

$$U(x, t+h) - f(x-t-h) = \sum_{\sigma=0}^S \{hg(x-t+2\sigma h) + f(x-t+2\sigma h) + \\ - f(x-t-h+2\sigma h)\}. \quad (8.31)$$

Nu wordt verondersteld dat $g(x)$ continu is en $f(x)$ continu differentieerbaar. Volgens de middelwaardstelling geldt dan

$$f(x-t+2\sigma h) - f(x-t+2\sigma h-h) = h f'(x-t+2\sigma h-h+\theta)$$

met $0 < \theta < h$.

Nu is

$$\lim_{h \rightarrow 0} \sum_{\sigma=0}^S h g(x-t+2\sigma h) = \frac{1}{2} \int_0^{2t} g(x-t+\tau) d\tau$$

$$\lim_{h \rightarrow 0} \sum_{\sigma=0}^S h f'(x-t+2\sigma h-h+\theta) = \frac{1}{2} \int_0^{2t} f'(x-t+\tau) d\tau \quad (8.32)$$

$$= \frac{1}{2} \{f(x+t) - f(x-t)\}.$$

Dus geldt

$$\lim_{h \rightarrow 0} [U(x, t+h) - f(x-t-h)] = \frac{1}{2} \int_{x-t}^{x+t} g(\tau) d\tau + \frac{1}{2} \{f(x+t) - f(x-t)\}$$

ofwel

$$\lim_{h \rightarrow 0} U(x, t+h) = \frac{1}{2} \{f(x+t) + f(x-t)\} + \frac{1}{2} \int_{x-t}^{x+t} g(\tau) d\tau. \quad (8.33)$$

Volgens de formule van d'Alembert is het rechterlid van (8.33) echter juist de oplossing $u(x, t)$ van de differentiaalvergelijking met de gegeven beginvoorwaarden.

Hieruit kunnen we dus besluiten dat

$$\lim_{h \rightarrow 0} U(x, t+h) + u(x, t)$$

m.a.w. de oplossing van het differentieelprobleem nadert tot de oplossing van het differentiaalprobleem voor $h, k \rightarrow 0$ ook in het geval $\lambda = 1$.

9. Stabiliteitsonderzoek voor differentiemethoden

Zoals op blz NA-90 werd aangetoond is de differentiemethode met formule (8.9) in het algemeen niet convergent als $\lambda > 1$. Als echter aan de beginfuncties speciale eisen worden opgelegd, bijvoorbeeld de eis dat $f(x)$ en $g(x)$ beide analytische functies zijn, dan kan men bewijzen dat ook voor $\lambda > 1$ geldt: $U(x,t)$ nadert tot $u(x,t)$ als $h, k \rightarrow 0$.

Dit is wel begrijpelijk want een analytische functie is volledig bepaald door de functiewaarden in een willekeurig klein interval. Het is in dit geval dus niet mogelijk de beginfuncties alleen in een gedeelte van het begininterval te veranderen. Daarmee vervalt de redenering op blz NA-90.

Men zou hieruit de conclusie kunnen trekken dat in het geval van analytische beginfuncties voor iedere waarde van λ de differentieformule (8.9) bruikbaar is voor de oplossing van de snaarvergelijking. Dit is echter niet het geval. Er is nog een ander effect dat de bruikbaarheid van een methode bepaalt, namelijk de voortplanting van reeds gemaakt fouten. Bij iedere numerieke berekening worden afrondingsfouten geïntroduceerd. We zullen laten zien dat juist voor $\lambda > 1$ het effect van deze afrondingsfouten catastrofaal is voor de oplossing van de differentievergelijking.

Om het effect van een afrondingsfout duidelijk te kunnen zien, beschouwen we het geval van de snaarvergelijking met beginfuncties $f(x) \equiv 0$ en $g(x) \equiv 0$. De oplossing van de differentiaalvergelijking is dan $u(x,t) \equiv 0$; evenzo is de oplossing van (8.9) dan $U(x,t) \equiv 0$. Nu veronderstellen we dat er één afrondingsfout geïntroduceerd is. We kunnen dit doen door bijvoorbeeld $U(0,k) = \epsilon$ te nemen, in plaats van nul. Onder invloed van deze afrondingsfout zal de oplossing van de differentievergelijking van nul verschillen. Deze van nul afwijkende oplossing vinden we uit de differentievergelijking (8.9) met als beginvoorwaarden

$$f(x) = 0 \quad g(x) = \begin{cases} 0 & \text{als } x \neq 0 \\ \epsilon/k & \text{als } x = 0 \end{cases}$$

ofwel

$$U(x,0) = 0 \quad U(x,k) = \begin{cases} 0 & \text{als } x \neq 0 \\ \epsilon & \text{als } x = 0 \end{cases} \quad (9.1)$$

In onderstaande tabel wordt de oplossing gegeven voor het geval $\lambda = 2$ en $\epsilon = 1$.

| t | | | | | | | | | |
|----|-----|-------|------|-------|------|-------|------|-------|-----|
| 5k | 256 | -1536 | 4432 | -7920 | 9541 | -7920 | 4432 | -1536 | 256 |
| 4k | 0 | 64 | -288 | 616 | -780 | 616 | -288 | 64 | 0 |
| 3k | 0 | 0 | 16 | -48 | 67 | -48 | 16 | 0 | 0 |
| 2k | 0 | 0 | 0 | 4 | -6 | 4 | 0 | 0 | 0 |
| k | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | -4h | -3h | -2h | -h | 0 | h | 2h | 3h | 4h |

Uit deze tabel zien we dat de fouten in de oplossing snel toenemen en van de gewenste nauwkeurigheid van de oplossing is na enige stappen niet veel meer over.

We zullen dit verschijnsel van aangroeiende fouten weer instabiliteit noemen.

Opmerking. In dit eenvoudige geval van de snaarvergelijking kunnen we het effect van één enkele afrondingsfout analytisch bepalen.

Bij meer gecompliceerde vergelijkingen is dit dikwijls niet mogelijk. In zo'n geval bepalen we het effect van een afrondingsfout experimenteel door de differentievergelijking tweemaal op te lossen, waarbij de tweede keer de beginvoorwaarde in één enkel punt gewijzigd wordt. Het verschil van de beide oplossingen geeft dan een indruk van de stabiliteit van de methode.

We behandelen nu twee methoden om de stabiliteit van een differentieformule te onderzoeken: de Fourier-methode en de matrix-methode.

a) Methode met behulp van Fourierreeksen

We veronderstellen dat de functiewaarden $U(x,t)$ op de lijn $t = 0$ afrondingsfouten bevatten, en zullen nagaan hoe deze fouten zich voortplanten. Een willekeurige verdeling van de fouten kan voorgesteld worden door

een eindige Fourierreeks $\sum_{(n)} A_n e^{i\beta_n x}$. Het aantal termen van deze reeks is juist gelijk aan het aantal basispunten op de rechte $t = 0$. De coëfficiënten β_n mogen willekeurig gekozen worden, $\beta_n = n$ geeft de gewone Fourierreeks. De waarde van A_n wordt dan bepaald door de eis dat de reeks in ieder basispunt de waarde van de fout in dat punt moet hebben. Omdat het probleem lineair is kunnen we ons beperken tot het onderzoek van de invloed van één term van de Fourierreeks. We nemen dus een foutenverdeling van de vorm $e^{i\beta x}$, en gaan na onder welke voorwaarden de invloed van deze fout beperkt blijft.

De voortplanting van de fouten vinden we evenals hierboven in de van nul verschillende oplossing van de differentievergelijking (8.9), nu met de beginvoorwaarde

$$U(x,0) = e^{i\beta x}. \quad (9.2)$$

We kunnen deze oplossing vinden met de methode van de scheiding van variabelen. Daartoe stellen we

$$U(x,t) = e^{i\beta x} e^{at}. \quad (9.3)$$

De groei van de fout hangt dan af van de grootte van de factor e^{at} .

Als $|e^{at}| \leq 1$ dan blijven de voortgeplante fouten beperkt, de methode heet dan stabiel. Maar als $|e^{at}| > 1$ dan groeien de fouten in de oplossing exponentieel, en hebben we een instabiele methode.

Substitutie van (9.3) in de differentieformule (8.9) geeft

$$e^{i\beta x} e^{\alpha(t+k)} = -e^{i\beta x} e^{\alpha(t-k)} + 2(1-\lambda^2)e^{i\beta x} e^{\alpha t} + \lambda^2 [e^{i\beta(x+h)} e^{\alpha t} + e^{i\beta(x-h)} e^{\alpha t}]$$

ofwel

$$e^{\alpha k} + e^{-\alpha k} = 2(1-\lambda^2) + \lambda^2(e^{i\beta h} + e^{-i\beta h})$$

$$e^{\alpha k} + e^{-\alpha k} = 2 - 4\lambda^2 \sin^2 \frac{\beta h}{2} \quad (9.4)$$

Deze vergelijking heeft twee oplossingen α en $-\alpha$. Welke van de twee genomen moet worden hangt af van de tweede (hier niet gegeven) beginvoorwaarde $U(x,k)$.

De methode is alleen stabiel als zowel $|e^{\alpha k}| \leq 1$ als $|e^{-\alpha k}| \leq 1$ dus

$|e^{\alpha k}| = 1$, waaruit volgt dat α zuiver imaginair is dus $\alpha = i\varphi$. Vergelijking (9.4) wordt dan

$$\cos \varphi = 1 - 2\lambda^2 \sin^2 \frac{\beta h}{2}.$$

Deze vergelijking bezit alleen een oplossing φ als $|1 - 2\lambda^2 \sin^2 \frac{\beta h}{2}| \leq 1$.

Daar dit moet gelden voor alle β volgt hieruit direct dat $\lambda^2 \leq 1$ en dus $\lambda \leq 1$. We zien dus dat de voorwaarde voor stabiliteit van de differentieformule (8.9) is $\lambda \leq 1$. Hoewel dus soms voor $\lambda > 1$ de convergentie in orde is zal de instabiliteit van de differentieformule tot gevolg hebben, dat we geen betrouwbare resultaten vinden.

We hebben vroeger bij de gewone differentiaalvergelijkingen gezien (blz NA-62,63) dat een impliciete differentieformule beter was, vanwege de stabiliteit, dan enkele expliciete formules.

We zullen nu aan een voorbeeld laten zien dat zo iets ook geldt bij hyperbolische partiële differentiaalvergelijkingen.

We krijgen een impliciete differentiaalvergelijking voor het snaarprobleem door in (8.1) de tweede afgeleide $\frac{\partial^2 u}{\partial x^2}$ in het punt (x,t) niet te

benaderen door de centrale differentie in (x,t) zoals vroeger, maar door de gemiddelde van de centrale differenties in de punten $(x,t+k)$ en $(x,t-k)$. Men kan bewijzen dat de afbreekfout van dezelfde orde is dus

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{2h^2} [U(x+h,t+k) - 2U(x,t+k) + U(x-h,t+k) + U(x+h,t-k) - 2U(x,t-k) + U(x-h,t-k)] + O(h^2).$$

De differentievergelijking wordt in dit geval

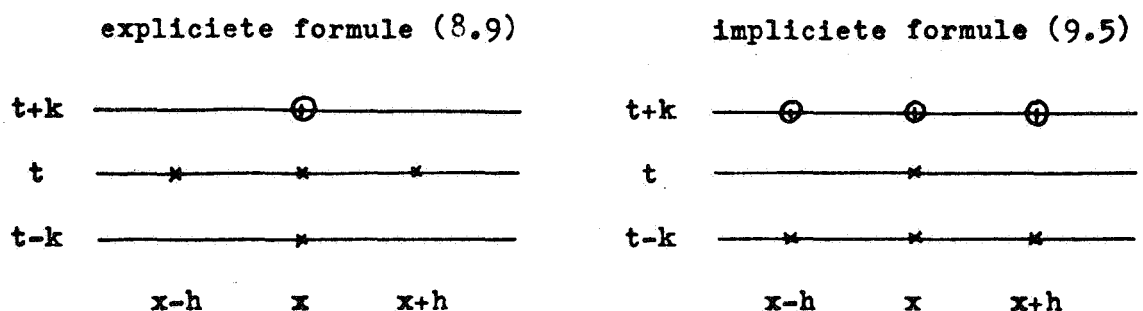
$$\frac{1}{2h^2} [U(x+h,t+k) - 2U(x,t+k) + U(x-h,t+k) + U(x+h,t-k) - 2U(x,t-k) + U(x-h,t-k)] = \frac{1}{k^2} [U(x,t+k) - 2U(x,t) + U(x,t-k)]$$

ofwel

$$U(x,t+k) - 2U(x,t) + U(x,t-k) = \frac{1}{2} \lambda^2 [U(x+h,t+k) - 2U(x,t+k) + U(x-h,t+k) + U(x+h,t-k) - 2U(x,t-k) + U(x-h,t-k)] \quad (9.5)$$

In deze differentieformule zijn $U(x+h,t+k)$, $U(x,t+k)$ en $U(x-h,t+k)$ onbekend.

Het rekenpatroon ziet er als volgt uit



Uit één differentievergelijking kunnen we niet de drie onbekende functiewaarden vinden. Daarom beschouwen we de serie vergelijkingen (9.5) van het niveau $t+k$ gelijktijdig. Nemen we aan dat in de x -richting $N+1$ basispunten zijn en voeren we nog de afkorting

$$U(x_0 + rh, sk) = U_{r,s} \quad (9.6)$$

in, dan krijgen we het stelsel van $N-1$ lineaire vergelijkingen

$$\frac{1}{2} \lambda^2 U_{r+1,s+1} - (1 + \lambda^2) U_{r,s+1} + \frac{1}{2} \lambda^2 U_{r-1,s+1} = R \quad (9.7)$$

$$r = 1, 2, \dots, N-1$$

waarbij het rechterlid R bestaat uit bekende functiewaarden op het niveau $t = (s-1)k$ en $t = sk$. Nemen we nu ook nog aan dat $U_{0,s+1}$ en $U_{N,s+1}$ bekend zijn (randvoorwaarden), dan is het aantal onbekenden in het stelsel vergelijkingen (9.7) gelijk aan $N-1$, en dus is dit stelsel oplosbaar.

We onderzoeken eerst de stabiliteit van deze differentiemethode. Analoog als in het voorgaande stellen we

$$U(x,t) = e^{i\beta x} e^{\alpha t} .$$

Substitutie hiervan in (9.5) geeft

$$e^{\alpha k} - 2 + e^{-\alpha k} = \frac{1}{2} \lambda^2 [e^{i\beta h} e^{\alpha k} - 2e^{\alpha k} + e^{-i\beta h} e^{\alpha k} + e^{i\beta h} e^{-\alpha k} - 2e^{-\alpha k} + e^{-i\beta h} e^{-\alpha k}]$$

waaruit gemakkelijk volgt

$$e^{\alpha k} + e^{-\alpha k} = \frac{2}{1 + 2\lambda^2 \sin^2 \frac{\beta h}{2}} \quad (9.8)$$

Een analoge redenering als op blz NA-100 geeft als voorwaarde voor stabiliteit

$$\left| \frac{1}{1 + 2\lambda^2 \sin^2 \frac{\beta h}{2}} \right| \ll 1$$

en dit geldt voor iedere waarde van λ . Dus de impliciete methode (9.5) is onvoorwaardelijk stabiel; een soortgelijk resultaat als op blz NA-63 voor gewone differentiaalvergelijkingen werd gevonden.

We kunnen nu de vraag stellen of het oplossen van een, meestal nogal groot, stelsel lineaire vergelijkingen niet een hoge prijs is voor de onvoorwaardelijke stabiliteit van de methode.

Het oplossen van het stelsel vergelijkingen is echter in dit geval eenvoudig door de speciale gedaante van de coëfficiënten-matrix.

Het stelsel lineaire vergelijkingen ziet eruit als volgt

$$\begin{aligned} a_1 x_1 + b_1 x_2 &= d_1 \\ c_2 x_1 + a_2 x_2 + b_2 x_3 &= d_2 \\ &\vdots \\ c_r x_{r-1} + a_r x_r + b_r x_{r+1} &= d_r \\ &\vdots \\ c_n x_{n-1} + a_n x_n &= d_n \end{aligned} \quad (9.9)$$

ofwel in matrix-notatie $Bx = d$, waarbij B een tridiagonale matrix is, d.w.z. een matrix die alleen in de hoofddiagonaal en de beide neven-diagonalen van nul verschillende elementen heeft.

$$B = \begin{pmatrix} a_1 & b_1 & & & \theta \\ c_2 & a_2 & b_2 & & \\ & \ddots & \ddots & \ddots & \\ & & c_{n-1} & a_{n-1} & b_{n-1} \\ \theta & & & c_n & a_n \end{pmatrix} \quad (9.10)$$

Hierbij is $n = N-1$, $a_1 = -(1 + \lambda^2)$, $b_1 = c_1 = \frac{1}{2} \lambda^2$.

Het oplossen van $Bx = d$ gaat nu als volgt.

We zoeken een onderdiagonaalmatrix W en een bovendiagonaalmatrix Q zodanig dat $B = W.Q$.

$$W = \begin{pmatrix} w_1 & & & \theta \\ c_2 & \ddots & & \\ & \ddots & \ddots & \\ \theta & & c_n & w_n \end{pmatrix} \quad Q = \begin{pmatrix} 1 & q_1 & & \theta \\ & \ddots & \ddots & \\ & & \ddots & q_{n-1} \\ \theta & & & 1 \end{pmatrix} \quad (9.11)$$

De elementen w_i en q_i vinden we op de volgende manier.

Uit $B = W \cdot Q$ volgt voor de hoofddiagonaal

$$\begin{array}{ll} a_1 = w_1 & w_1 = a_1 \\ a_2 = c_2 q_1 + w_2 & w_2 = a_2 - c_2 q_1 \\ \vdots & \vdots \\ a_n = c_n q_{n-1} + w_n & w_n = a_n - c_n q_{n-1} \end{array} \quad \text{dus} \quad (9.12)$$

en voor de nevendiaagonaal

$$\begin{array}{ll} b_1 = w_1 q_1 & q_1 = \frac{b_1}{w_1} \\ b_2 = w_2 q_2 & q_2 = \frac{b_2}{w_2} \\ \vdots & \vdots \\ b_{n-1} = w_{n-1} q_{n-1} & q_{n-1} = \frac{b_{n-1}}{w_{n-1}} \end{array} \quad \text{dus} \quad (9.13)$$

Uit (9.12) en (9.13) kunnen we de elementen w_i en q_i als volgt berekenen

$$w_1 = a_1 \Rightarrow q_1 = \frac{b_1}{w_1} \Rightarrow w_2 = a_2 - c_2 q_1 \Rightarrow q_2 = \frac{b_2}{w_2} \quad \text{etc.}$$

Als we op deze manier de matrices W en Q hebben gevonden, dan lossen we $B\mathbf{x} = \mathbf{d}$ op de volgende manier op. Eerst bepalen we de vector \mathbf{y} uit $W\mathbf{y} = \mathbf{d}$ als volgt

$$y_1 = \frac{d_1}{w_1} \Rightarrow y_2 = \frac{d_2 - c_2 y_1}{w_2} \quad \text{etc}$$

daarna bepalen we de gezochte oplossing \mathbf{x} uit $Q\mathbf{x} = \mathbf{y}$ als volgt

$$x_n = y_n \Rightarrow x_{n-1} = y_{n-1} - q_{n-1} y_n \quad \text{etc.}$$

Is bij eliminatiemethode voor het oplossen van een stelsel lineaire vergelijkingen voor een willekeurige matrix het aantal vermenigvuldigingen in de orde van n^3 ; voor een tridiagonale matrix is met bovenstaande rekenwijze het aantal vermenigvuldigingen in de orde van n .

b). Stabiliteitsonderzoek met behulp van de matrixmethode

De impliciete differentiemethode vereist randvoorwaarden. Bijvoorbeeld moeten bij het probleem van de trillende snaar de functiewaarden $U(x,t)$ voorgeschreven zijn op twee rechten $x = \text{constant}$. Voor het stabiliteitsonderzoek kan men nu het effect van deze randvoorwaarden direct in rekening brengen door inplaats van de Fouriermethode de zgn. matrix-methode te gebruiken.

We veronderstellen in het volgende dat de functie $U(x,t) = 0$ op de twee rechten $x = x_0$ en $x = x_0 + Nh$.

Met de afkorting (9.6) wordt de impliciete differentie-vergelijking (9.5):

$$U_{r,s+1} - 2U_{r,s} + U_{r,s-1} = \frac{1}{2}\lambda^2 [U_{r+1,s+1} - 2U_{r,s+1} + U_{r-1,s+1} + U_{r+1,s-1} - 2U_{r,s-1} + U_{r-1,s-1}]$$

ofwel

$$\frac{1}{2}\lambda^2 U_{r+1,s+1} - (1+\lambda^2)U_{r,s+1} + \frac{1}{2}\lambda^2 U_{r-1,s+1} = -2U_{r,s} + -[\frac{1}{2}\lambda^2 U_{r+1,s-1} - (1+\lambda^2)U_{r,s-1} + \frac{1}{2}\lambda^2 U_{r-1,s-1}] \quad (9.14)$$

We beschouwen de onbekende functiewaarden $U(x,t)$ in de $N-1$ roosterpunten $x_r = x_0 + rh$ op de rechte $t = sk$ als een vector die we zullen aanduiden met \underline{U}_s , dus

$$\underline{U}_s = \begin{pmatrix} U_{1,s} \\ U_{2,s} \\ \vdots \\ U_{N-1,s} \end{pmatrix}.$$

Dan kunnen we het stelsel vergelijkingen (9.14) voor $r = 1, 2, \dots, N-1$ als volgt in matrixnotatie schrijven

$$A\underline{U}_{s+1} = B\underline{U}_s - C\underline{U}_{s-1} \quad \text{ofwel} \quad \underline{U}_{s+1} = A^{-1}B\underline{U}_s - A^{-1}C\underline{U}_{s-1} \quad (9.15)$$

waarbij $B = -2I$ met $I =$ eenheidsmatrix en $A = C$ een tridiagonaalmatrix is (zie (9.10)).

De vergelijking (9.15) voor de vector \underline{U}_{s+1} kunnen we herleiden tot een zgn. "two-level" formule door de vectoren \underline{U}_s en \underline{U}_{s-1} te combineren tot een enkele vector \underline{V}_s

$$\underline{V}_s = \begin{pmatrix} \underline{U}_s \\ \underline{U}_{s-1} \end{pmatrix} = \begin{pmatrix} U_{1,s} \\ \vdots \\ U_{N-1,s} \\ U_{1,s-1} \\ \vdots \\ U_{N-1,s-1} \end{pmatrix}.$$

Uit de twee formules

$$\begin{aligned} \underline{U}_{s+1} &= A^{-1}B\underline{U}_s - A^{-1}C\underline{U}_{s-1} \\ \underline{U}_s &= \underline{U}_s \end{aligned} \tag{9.16}$$

Volgt in matrixnotatie de vergelijking voor de vector \underline{V}_{s+1}

$$\underline{V}_{s+1} = P\underline{V}_s$$

waarbij P een matrix is die samengesteld is uit de matrices van (9.16)

$$P = \begin{pmatrix} A^{-1}B & -A^{-1}C \\ I & 0 \end{pmatrix}. \tag{9.17}$$

We veronderstellen nu dat de matrix P allemaal verschillende eigenwaarden heeft. Er zijn dus $2(N-1)$ verschillende eigenwaarden t_i , de bijbehorende eigenvectoren noemen we \underline{v}_i . Deze vectoren vormen dan een basis van de lineaire ruimte der \underline{V} 's, dus we kunnen de beginvector \underline{V}_1 van ons probleem schrijven als een lineaire combinatie van de eigenvectoren

$$\underline{V}_1 = \sum_{i=1}^{2(N-1)} \alpha_i \underline{v}_i.$$

Dan geldt echter

$$\underline{V}_{s+1} = P^s \underline{V}_1 = \sum_{i=1}^{2(N-1)} \alpha_i t_i^s \underline{v}_i.$$

Beschouwen we nu de elementen van \underline{V}_1 als de fouten op het beginniveau, dan geeft de vector \underline{V}_{s+1} de voortgeplante fouten in de oplossing op het niveau $t = (s+1)k$. Voor de stabiliteit van de methode is vereist dat deze fouten begrensd zijn, dus $|t_j| < 1$.

De eigenwaarden t_j van de matrix P bepalen we uit de vergelijking $\det(P-tI) = 0$, ofwel

$$\det \begin{pmatrix} A^{-1}B-tI & -A^{-1}C \\ I & -tI \end{pmatrix} = 0.$$

We tellen t keer de eerste "kolom" bij de tweede op (deze kolommen bestaan uit twee matrices, men kan gemakkelijk nagaan dat de rekenwijze juist is).

We vinden dan

$$\det \begin{pmatrix} A^{-1}B-tI & -A^{-1}C+t(A^{-1}B-tI) \\ I & 0 \end{pmatrix} = 0.$$

Uitwerken geeft

$$\det I \cdot \det(A^{-1}C+t(A^{-1}B-tI)) = 0$$

en daar $\det I \neq 0$, moet dus gelden

$$\det(-A^{-1}C+t(A^{-1}B-tI)) = 0.$$

Omdat $A^{-1}C = I$ is $\det A^{-1} \neq 0$ en $t = 0$ geen eigenwaarde, dus

$$\det(t \cdot A^{-1}) \cdot \det(-\frac{1}{t}C + B - tA) = 0$$

geeft

$$\det(B - (t + \frac{1}{t})C) = 0. \quad (9.18)$$

Hiermee hebben we de eigenwaardenvergelijking voor t teruggebracht tot determinantvergelijking van de orde $N-1$.

Stellen we $t + \frac{1}{t} = \mu$ dan wordt deze vergelijking in μ :

$$\begin{vmatrix} -2+(1+\lambda^2)\mu & -\frac{1}{2}\lambda^2\mu & & & & & \ominus \\ -\frac{1}{2}\lambda^2\mu & -2+(1+\lambda^2)\mu & & & & & \\ & & \cdot & & & & \\ & & & \cdot & & & \\ & & & & \cdot & & \\ \ominus & & & & & \cdot & \\ & & & & & & -\frac{1}{2}\lambda^2\mu & -2+(1+\lambda^2)\mu \end{vmatrix} = 0 \quad (9.19)$$

Dit is geen normaal eigenwaardenprobleem.
We beschouwen nu de differentievergelijking

$$-\frac{1}{2}\lambda^2\mu Y_{r+1} + (-2+(1+\lambda^2)\mu) \cdot Y_r - \frac{1}{2}\lambda^2\mu Y_{r-1} = 0 \quad (9.20)$$

met de randvoorwaarden

$$Y_0 = Y_N = 0.$$

Dan is de matrix van (9.19) gelijk aan de coëfficiënten-matrix van het stelsel vergelijkingen (9.20) voor $r = 1, 2, \dots, N-1$. De oplossingen van (9.19) zijn dus de waarden van μ waarvoor (9.20) een niet-triviale oplossing heeft. We bepalen daarom de niet-triviale oplossingen van (9.20), en stellen daartoe

$$Y_r = e^{i\beta r}.$$

Substitutie hiervan in (9.20) levert

$$-\frac{1}{2}\lambda^2\mu e^{i\beta} + (-2+(1+\lambda^2)\mu) - \frac{1}{2}\lambda^2\mu e^{-i\beta} = 0$$

en hieruit volgt

$$\cos \beta = \frac{-2+(1+\lambda^2)\mu}{\lambda^2\mu}. \quad (9.21)$$

Deze vergelijking heeft twee oplossingen namelijk $\pm \beta$. De algemene oplossing van de differentievergelijking (9.20) kan dus geschreven worden in de gedaante

$$Y_r = c_1 e^{i\beta r} + c_2 e^{-i\beta r}.$$

De coëfficiënten c_1 en c_2 volgen uit de randvoorwaarden

$$\begin{aligned} Y_0 &= c_1 + c_2 = 0 \\ Y_N &= c_1 e^{i\beta N} + c_2 e^{-i\beta N} = 0 \end{aligned}$$

waaruit volgt $c_1 = -c_2$, en $2ic_1 \sin \beta N = 0$. $c_1 = 0$ geeft de triviale oplossing. Voor de niet-triviale oplossingen moet dus

$$\sin \beta N = 0 \quad \text{ofwel} \quad \beta = \frac{k\pi}{N} \quad \text{met } k = 1, 2, \dots, N-1.$$

Dit invullen in (9.21) geeft

$$\cos \frac{k\pi}{N} = \frac{-2+(1+\lambda^2)\mu}{\lambda^2\mu} .$$

Hieruit kunnen we tenslotte μ oplossen. We vinden dan

$$\mu = \frac{2}{1+2\lambda^2 \sin^2 \frac{k\pi}{2N}} . \quad (9.22)$$

We zien hieruit dat voor iedere waarde van λ geldt $\mu \leq 2$.

Nu is $\mu = t + \frac{1}{t}$, dus geldt dat $t + \frac{1}{t}$ reëel is en $0 \leq t + \frac{1}{t} \leq 2$ waaruit we gemakkelijk kunnen afleiden dat $|t| = 1$.

Hiermee is aangetoond dat alle eigenwaarden t_j van P in absolute waarde gelijk zijn aan 1. Dus hebben we bewezen dat de methode met differentieformule (9.14) onvoorwaardelijk stabiel is.

We onderzoeken nu nog de stabiliteit van differentiemethoden bij systemen van eerste orde partiële differentiaalvergelijkingen.

De eindige differentiemethoden kunnen ook gebruikt worden voor eerste orde partiële differentiaalvergelijkingen of systemen van eerste orde vergelijkingen.

De convergentie van zo'n methode is in het algemeen in orde, evenals bij de tweede orde partiële differentiaalvergelijkingen, indien de verhouding van de stapgrootten h en k zo is dat het invloedsgebied van de differentiaalvergelijking (karakteristieken-driehoek) geheel valt binnen de invloedsdriehoek van de eindige differentiemethode (zie blz NA-90 ev)

Het onderzoek naar de stabiliteit kan weer geschieden met bijvoorbeeld de Fourier-methode. Als voorbeeld beschouwen we het volgende systeem

$$\frac{\partial p}{\partial x} = \frac{\partial q}{\partial t} \quad \frac{\partial q}{\partial x} = \frac{\partial p}{\partial t} . \quad (9.23)$$

Dit systeem is equivalent met de vergelijking voor de trillende snaar

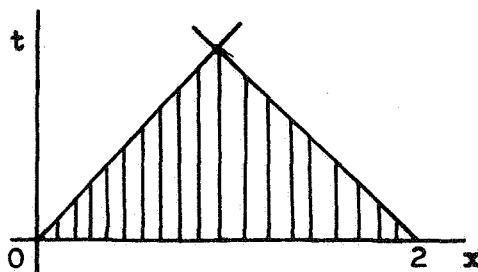
hetgeen direct duidelijk is als we $p = \frac{\partial u}{\partial x}$ en $q = \frac{\partial u}{\partial t}$ stellen.

We veronderstellen nu dat als beginvoorwaarden zijn gegeven: $p(x, t)$ en $q(x, t)$ op het interval $0 \leq x \leq 2$ van de rechte $t = 0$.

De karakteristieken zijn de rechten $x-t = \alpha$ en $x+t = \alpha$, en het invloedsgebied voor $t > 0$ wordt begrensd door de rechten:

$$t = 0 \quad x-t = 0 \quad x+t = 2.$$

Indien we nu met de eenvoudigste eindige differentieformules willen werken, dan benaderen we de afgeleiden naar t door de voorwaartse differentie. In de x -richting kunnen we dan nog kiezen tussen voorwaartse, centrale en achterwaartse differenties.



Om binnen het gebied te blijven moeten we bij de rechte $x-t=0$ geen achterwaartse differenties en bij $x+t=2$ geen voorwaartse differenties gebruiken. Centrale differenties zijn overal in het invloedsgebied te gebruiken.

We veronderstellen nu dat op de rechten $x=0$ en $x=2$ als randvoorwaarde de waarde van $q(x,t)$ gegeven is. Dan is daarmee de oplossing van het systeem (9.23) eenduidig bepaald in de strook tussen de rechten $x=0$ en $x=2$.

We beschouwen nu de differentiemethode die ontstaat door in de x -richting centrale differenties en in de t -richting voorwaartse differenties te nemen.

De differentieformules worden dan:

$$\frac{p_{r+1,s} - p_{r-1,s}}{2h} = \frac{q_{r,s+1} - q_{r,s}}{k} \quad (9.24)$$

$$\frac{q_{r+1,s} - q_{r-1,s}}{2h} = \frac{p_{r,s+1} - p_{r,s}}{k}$$

Dit zijn expliciete differentieformules. De waarden van p en q op het niveau $t = (s+1)k$ worden bepaald met behulp van de waarden van p en q op het niveau $t = sk$.

Om de stabiliteit te onderzoeken veronderstellen we een fout $\epsilon_r = A e^{i\beta rh}$ in de beginwaarden $p_{r,0}$ en een fout $\eta_r = B e^{i\beta rh}$ in de beginwaarden $q_{r,0}$ en bepalen weer de oplossing die bestaat uit de voortgeplante fouten. Dit is de oplossing van (9.24) met de beginvoorwaarden

$$p_{r,0} = A e^{i\beta rh} \quad q_{r,0} = B e^{i\beta rh} \quad (9.25)$$

We vinden deze oplossing weer met de methode van de scheiding van variabelen. We stellen

$$p_{r,s} = A e^{i\beta rh} e^{\alpha sk} \quad q_{r,s} = B e^{i\beta rh} e^{\alpha sk} \quad (9.26)$$

Substitutie hiervan in (9.24) levert dan

$$\begin{aligned} A e^{\alpha sk} \frac{e^{i\beta(r+1)h} - e^{i\beta(r-1)h}}{2h} &= B e^{i\beta rh} \frac{e^{\alpha(s+1)k} - e^{\alpha sk}}{k} \\ B e^{\alpha sk} \frac{e^{i\beta(r+1)h} - e^{i\beta(r-1)h}}{2h} &= A e^{i\beta rh} \frac{e^{\alpha(s+1)k} - e^{\alpha sk}}{k} \end{aligned} \quad (9.27)$$

ofwel

$$A \frac{e^{i\beta h} - e^{-i\beta h}}{2h} = B \frac{e^{\alpha k} - 1}{k}$$

$$B \frac{e^{i\beta h} - e^{-i\beta h}}{2h} = A \frac{e^{\alpha k} - 1}{k}$$

ofwel

$$i\lambda A \sin \beta h - B(e^{\alpha k} - 1) = 0 \quad (9.28)$$

$$-A(e^{\alpha k} - 1) + i\lambda B \sin \beta h = 0.$$

Dit stelsel vergelijkingen voor A en B heeft alleen dan een oplossing als α zodanig gekozen wordt dat

$$\begin{vmatrix} i\lambda \sin \beta h & -(e^{\alpha k} - 1) \\ -(e^{\alpha k} - 1) & i\lambda \sin \beta h \end{vmatrix} = 0 \quad (9.29)$$

ofwel $(e^{\alpha k} - 1)^2 = -\lambda^2 \sin^2 \beta h$.

Dan is echter $e^{\alpha k} - 1$ zuiver imaginair: $e^{\alpha k} - 1 = i\gamma$

dus $e^{\alpha k} = 1 + i\gamma$, waaruit volgt dat $|e^{\alpha k}| > 1$.

We zien dus dat voor geen enkele waarde van $\lambda \neq 0$ de methode stabiel is. De differentieformules (9.24) zijn dus volkomen onbruikbaar.

Een andere differentiemethode krijgen we door zowel in de x-richting als in de t-richting centrale differenties te nemen. We krijgen dan de differentieformules

$$\frac{p_{r+1,s} - p_{r-1,s}}{2h} = \frac{q_{r,s+1} - q_{r,s-1}}{2k} \quad (9.30)$$

$$\frac{q_{r+1,s} - q_{r-1,s}}{2h} = \frac{p_{r,s+1} - p_{r,s-1}}{2k} .$$

Dit zijn eveneens expliciete formules; de waarden van p en q op een nieuwe lijn worden nu bepaald uit de waarden op de twee voorgaande lijnen.

Opmerking. Voor deze methode is de beginvoorwaarde: p en q gegeven op de lijn $t = 0$ niet voldoende. We moeten eerst p en q ook berekenen, bijvoorbeeld met voorwaartse differentieformules, op de lijn $t = k$, m.a.w. de methode heeft een aparte startprocedure.

De stabiliteit van deze methode vinden we, analoog als in het vorige geval, uit de oplossing van (9.30) met de beginvoorwaarden (9.25). We substitueren (9.26) in (9.30) en krijgen dan

$$\begin{aligned} A e^{\alpha k s} \frac{e^{i\beta(r+1)h} - e^{i\beta(r-1)h}}{2h} &= B e^{i\beta r h} \frac{e^{\alpha(s+1)k} - e^{\alpha(s-1)k}}{2k} \\ B e^{\alpha k s} \frac{e^{i\beta(r+1)h} - e^{i\beta(r-1)h}}{2h} &= A e^{i\beta r h} \frac{e^{\alpha(s+1)k} - e^{\alpha(s-1)k}}{2k} \end{aligned} \quad (9.31)$$

Hieruit volgt

$$\begin{aligned} 2i\lambda A \sin \beta h - B(e^{\alpha k} - e^{-\alpha k}) &= 0 \\ A(e^{\alpha k} - e^{-\alpha k}) - 2i\lambda B \sin \beta h &= 0 \end{aligned}$$

en dit stelsel heeft een oplossing A en B als de coëfficiëntendeterminant gelijk aan nul is, dus

$$\begin{aligned} (e^{\alpha k} - e^{-\alpha k})^2 &= -4\lambda^2 \sin^2 \beta h \\ e^{\alpha k} - e^{-\alpha k} &= \pm 2i\lambda \sin \beta h. \end{aligned} \quad (9.32)$$

Analoog als op blz NA-100 vinden we dat de methode alleen stabiel is als $|\lambda \sin \beta h| \leq 1$ voor alle β . De voorwaarde voor stabiliteit is dus $\lambda \leq 1$.

Als derde methode voor het systeem (9.23) behandelen we tenslotte nog een impliciete methode. We krijgen de impliciete formules door de tweede afgeleide naar x weer te vervangen door het gemiddelde van twee centrale differenties (zie blz NA-100).

De differentieformules worden dan

$$\begin{aligned} \frac{p_{r+1,s+1} - p_{r-1,s+1} + p_{r+1,s} - p_{r-1,s}}{4h} &= \frac{q_{r,s+1} - q_{r,s}}{k} \\ \frac{q_{r+1,s+1} - q_{r-1,s+1} + q_{r+1,s} - q_{r-1,s}}{4h} &= \frac{p_{r,s+1} - p_{r,s}}{k} \end{aligned} \quad (9.33)$$

Na substitutie van (9.26) in deze formules vinden we

$$\begin{aligned} \frac{1}{2}i\lambda A(e^{\alpha k} + 1)\sin \beta h - B(e^{\alpha k} - 1) &= 0 \\ A(e^{\alpha k} - 1) - \frac{1}{2}i\lambda B(e^{\alpha k} + 1)\sin \beta h &= 0 \end{aligned}$$

waaruit we nu als vergelijking voor α vinden

$$(e^{\alpha k} - 1)^2 = -\frac{1}{4}\lambda^2 (e^{\alpha k} + 1)^2 \sin^2 \beta h$$

ofwel

$$\frac{e^{ak} - 1}{e^{ak} + 1} = \pm \frac{i\lambda}{2} \sin \beta h. \quad (9.34)$$

Voor stabiliteit is weer vereist $|e^{ak}| \leq 1$ en $|e^{-ak}| \leq 1$, dus $\alpha = i\varphi$. Vullen we dit in (9.34) in dan krijgen we

$$\tan \frac{\varphi}{2} = \pm \frac{\lambda}{2} \sin \beta h$$

en deze vergelijking heeft voor alle waarden van λ een oplossing φ . We zien dus dat ook in dit geval de impliciete methode onvoorwaardelijk stabiel is.

10. Parabolische differentiaalvergelijkingen

Physische problemen welke mathematisch beschreven worden door een parabolische differentiaalvergelijking van de tweede orde hebben veelal betrekking op diffusieverschijnselen (o.a. warmtegeleiding). De normaalvorm voor een parabolische differentiaalvergelijking is

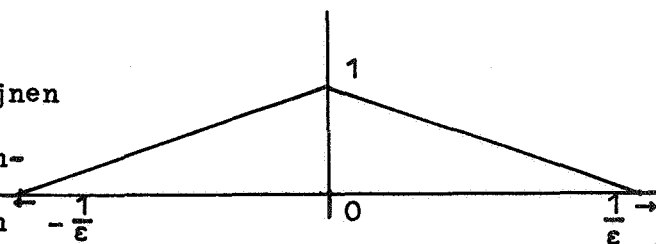
$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \quad (10.1)$$

waarbij x de plaatscoördinaat en t de tijd voorstelt. Deze differentiaalvergelijking heeft slechts één schaar karakteristieken, nl. de lijnen $t = \text{constant}$. Dit kunnen we bijvoorbeeld inzien door de vergelijking $u_t = u_{xx}$ op te vatten als de limiet van de hyperbolische vergelijking

$$u_{xx} - \epsilon^2 u_{tt} = u_t \text{ voor } \epsilon \rightarrow 0.$$

De karakteristieken-richtingen hiervan zijn $\frac{dt}{dx} = \pm \epsilon$. Beide scharen karakteristieken naderen dus tot de lijnen $t = \text{constant}$ als $\epsilon \rightarrow 0$.

Daarmee verdwijnt de karakteristieken-driehoek; het invloedsgebied voor de waarde van de oplossing van (10.1) in een willekeurig punt is de gehele beginlijn. Omdat de diffusievergelijking van de eerste orde in t is, mogen we verwachten dat één beginvoorwaarde $u(x,0) = f(x)$ voldoende is voor een zinvol probleem. Is de beginvoorwaarde gegeven op een eindig interval dan zijn randvoorwaarden aan beide einden noodzakelijk voor een oplossing.



Uit bovenstaande blijkt dat voor de numerieke oplossing van een parabolische differentiaalvergelijking geen gebruik kan worden gemaakt van de karakteristieken. Daarom zijn we thans geheel aangewezen op de eindige differentiemethoden.

We beschouwen nu het klassieke warmtegeleidingsprobleem

$$\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} \quad (10.2)$$

met de beginvoorwaarde

$$u(x,0) = g(x) \quad 0 \leq x \leq L \quad (10.3)$$

en de randvoorwaarden

$$\begin{aligned} u(0,t) &= f_1(t) & t \geq 0 \\ u(L,t) &= f_2(t) & t \geq 0 \end{aligned} \quad (10.4)$$

De oplossing $u(x,t)$ van dit probleem geeft de temperatuur in een dunne staaf waarvan L de lengte, en de constante $a > 0$ de warmtegeleidingscoëfficiënt voorstelt.

De voorwaarden (10.3) en (10.4) geven de functie $u(x,t)$ op drie zijden van een gebied in het (x,t) -vlak. Aan één zijde is het gebied onbegrensd.

Het probleem is nu de functie $u(x,t)$ te bepalen in het gebied begrensd door deze drie zijden. Daartoe nemen we een rechthoekig net over het gebied. De stapgrootte in de x -richting noemen we weer h en in de t -richting k . Als benadering voor u_t nemen we de voorwaartse differentie, waarvoor geldt

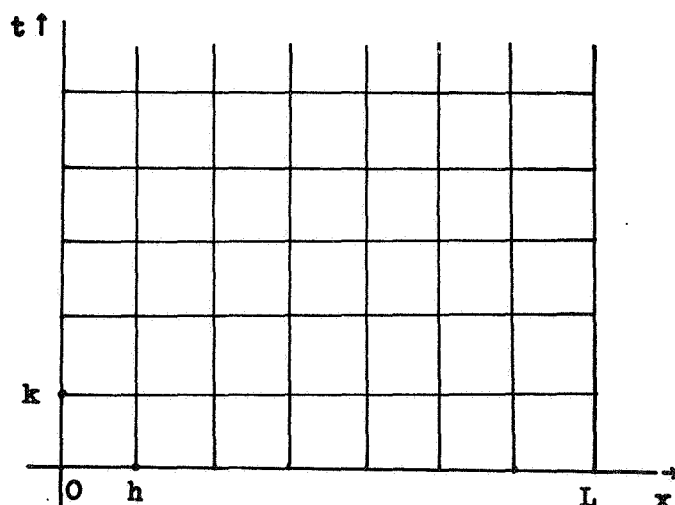
$$\frac{\partial u}{\partial t} = \frac{u(x,t+k) - u(x,t)}{k} - \mathcal{O}(k)$$

en als benadering voor u_{xx} nemen we de centrale differentie, waarvoor geldt

$$\frac{\partial^2 u}{\partial x^2} = \frac{u(x+h,t) - 2u(x,t) + u(x-h,t)}{h^2} + \mathcal{O}(h^2).$$

Op deze manier krijgen we een differentieformule die er als volgt uitziet

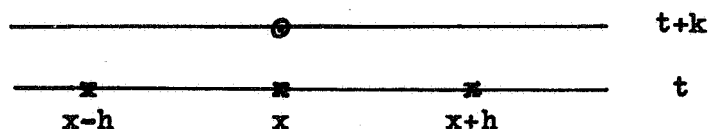
$$\frac{U(x,t+k) - U(x,t)}{k} = a \frac{U(x+h,t) - 2U(x,t) + U(x-h,t)}{h^2}$$



ofwel met $\lambda = \frac{ak}{h^2}$ (dan is λ een dimensieloze grootheid)

$$U(x, t+k) = (1 - 2\lambda)U(x, t) + \lambda\{U(x+h, t) + U(x-h, t)\}. \quad (10.5)$$

Het rekenpatroon ziet er met deze formule als volgt uit



We zullen de afbreekfout nu eens nader bekijken. We maken daarbij gebruik van de Taylor-ontwikkelingen van $u(x, t)$ naar x alleen, en naar t alleen. Er geldt

$$u(x+h, t) = u(x, t) + h u'_x + \frac{h^2}{2} u''_{xx} + \frac{h^3}{3!} u'''_{xxx} + \frac{h^4}{4!} u^{(4)}_{xxxx} + \frac{h^5}{5!} u^{(5)}_{xxxxx} + \frac{h^6}{6!} u^{(6)}_{xxxxxx}(\xi_1, t)$$

$$u(x-h, t) = u(x, t) - h u'_x + \frac{h^2}{2} u''_{xx} - \frac{h^3}{3!} u'''_{xxx} + \frac{h^4}{4!} u^{(4)}_{xxxx} - \frac{h^5}{5!} u^{(5)}_{xxxxx} + \frac{h^6}{6!} u^{(6)}_{xxxxxx}(\xi_2, t)$$

en

$$u(x, t+k) = u(x, t) + k u'_t + \frac{k^2}{2!} u''_{tt} + \frac{k^3}{3!} u^{(3)}_{ttt}(x, \tau).$$

Vullen we deze reeksen in de differentiaalvergelijking (10.2) in dan krijgen we

$$u'_t - a u_{xx} = \frac{u(x, t+k) - u(x, t)}{k} - a \frac{u(x+h, t) - 2u(x, t) + u(x-h, t)}{h^2} + R.$$

Hierin is R de afbreekfout van de differentieformule (10.5) waarvoor geldt

$$R = -\frac{1}{k} \left\{ \frac{k^2}{2!} u''_{tt}(x, t) + \frac{k^3}{3!} u^{(3)}_{ttt}(x, \tau) \right\} + \frac{a}{h^2} \left\{ \frac{2h^4}{4!} u^{(4)}_{xxxx}(x, t) + \frac{2h^6}{6!} u^{(6)}_{xxxxxx}(\xi, t) \right\}.$$

Nu geldt volgens de differentiaalvergelijking

$$u'_t = a u_{xx} \quad \text{en dus} \quad u''_{tt} = a^2 u^{(4)}_{xxxx}$$

$$\text{en ook} \quad u^{(3)}_{ttt} = a^3 u^{(6)}_{xxxxxx}.$$

Met behulp hiervan kunnen we de fout R schrijven in de vorm

$$R = \left(\frac{2ah^2}{4!} - \frac{a^2 k}{2!} \right) u^{(4)}_{xxxx}(x, t) + \left(\frac{2ah^4}{6!} - \frac{a^3 k^2}{3!} \right) u^{(6)}_{xxxxxx}(\rho, \sigma).$$

De afbreekfout wordt in eerste instantie gegeven door de eerste term van het rechterlid van deze formule. We zien hieraan nog eens dat de fout van de orde h^2 en k is.

Kiezen we nu h en k zodanig dat

$$\frac{2ah^2}{4!} - \frac{a^2k}{2!} = 0 \quad \text{ofwel} \quad \lambda = \frac{1}{6}$$

dan verdwijnt de eerste term van R . De afbreekfout is dan van hogere orde, namelijk van de orde h^4 en k^2 . De speciale keuze $\lambda = \frac{1}{6}$ kan dus beschouwd worden als een optimale keuze voor wat betreft de afbreekfout.

Convergentie-onderzoek voor de expliciete formule (10.5).

We onderzoeken de convergentie van deze differentiemethode alleen voor het geval van homogene randvoorwaarden, m.a.w. de randvoorwaarden (10.4) worden in dit geval

$$u(0,t) = u(L,t) = 0 \quad t > 0. \quad (10.6)$$

Eerst bepalen we de oplossing van de differentiaalvergelijking, met de methode van scheiding van de variabelen. Daartoe stellen we

$$u(x,t) = \sum_{n=1}^{\infty} b_n(t) \cdot \sin \frac{n\pi x}{L}. \quad (10.7)$$

Dan is aan de randvoorwaarden (10.6) voldaan.

Wegens de lineariteit van de differentiaalvergelijking kunnen we ons, evenals vroeger, beperken tot één term, $u_n(x,t) = b_n(t) \cdot \sin \frac{n\pi x}{L}$, van de Fourierreeks (10.7). Substitutie van $u_n(x,t)$ in de differentiaalvergelijking geeft een vergelijking voor $b_n(t)$:

$$b_n'(t) \sin \frac{n\pi x}{L} = -a \frac{n^2 \pi^2}{L^2} b_n(t) \sin \frac{n\pi x}{L}$$

ofwel

$$b_n'(t) + \frac{an^2\pi^2}{L^2} b_n(t) = 0. \quad (10.8)$$

De algemene oplossing van deze vergelijking is

$$b_n(t) = B_n e^{-\frac{an^2\pi^2}{L^2} t} \quad (10.9)$$

met B_n een willekeurige constante.

Met behulp hiervan kunnen we de oplossing $u(x,t)$ dan schrijven in de vorm

$$u(x,t) = \sum_{n=1}^{\infty} B_n e^{-\frac{a n^2 \pi^2}{L^2} t} \sin \frac{n\pi x}{L}. \quad (10.10)$$

De constanten B_n volgen dan uit de beginvoorwaarde, immers

$$u(x,0) = g(x) = \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{L} \quad (10.11)$$

dus de constanten B_n zijn de Fouriercoëfficiënten van $g(x)$

$$B_n = \frac{2}{L} \int_0^L g(x) \sin \frac{n\pi x}{L} dx. \quad (10.12)$$

Hiermee hebben we de oplossing van de differentiaalvergelijking (10.2) die voldoet aan de beginvoorwaarde (10.3) en de randvoorwaarden (10.6) gevonden.

Nu bepalen we de oplossing van de differentievergelijking, ook met separatie. We stellen

$$U(x,t) = \Phi(x) \cdot \Psi(t) \quad (10.13)$$

en substitueren dit in de differentievergelijking (10.5). Dan krijgen we

$$\Phi(x) \Psi(t+k) = \Phi(x) \Psi(t) + \lambda \{ \Phi(x+h) - 2\Phi(x) + \Phi(x-h) \} \Psi(t)$$

ofwel

$$\frac{\Psi(t+k)}{\Psi(t)} = 1 + \lambda \frac{\Phi(x+h) - 2\Phi(x) + \Phi(x-h)}{\Phi(x)}. \quad (10.14)$$

Het linkerlid is uitsluitend een functie van t , het rechterlid hangt alleen van x af. Beide leden zijn dus gelijk aan een constante C . Hiermee krijgen we de volgende twee gewone differentievergelijkingen

$$\Psi(t+k) = C \Psi(t) \quad (10.15)$$

$$\lambda \Phi(x+h) + (1 - 2\lambda) \Phi(x) + \lambda \Phi(x-h) = C \Phi(x).$$

Om aan de begin- en randvoorwaarden te kunnen voldoen zorgen we weer, evenals bij de hyperbolische vergelijkingen, dat $\sin \frac{n\pi x}{L}$ oplossing is van de vergelijking voor $\Phi(x)$. Dit levert voor C de vergelijking

$$\lambda \sin \frac{n\pi(x+h)}{L} + (1-2\lambda) \sin \frac{n\pi x}{L} + \lambda \sin \frac{n\pi(x-h)}{L} = C \sin \frac{n\pi x}{L}$$

waaruit gemakkelijk gevonden kan worden dat

$$C = 1 - 4\lambda \sin^2 \frac{n\pi h}{2L}. \quad (10.16)$$

De differentievergelijking voor $\Psi(t)$ wordt nu

$$\Psi(t+k) = (1 - 4\lambda \sin^2 \frac{n\pi h}{2L}) \Psi(t).$$

Om deze vergelijking op te lossen stellen we

$$\Psi(t) = e^{\alpha_n t}.$$

Door substitutie hiervan in de differentievergelijking vinden we voor α_n de waarde

$$\alpha_n = \frac{1}{k} \log(1 - 4\lambda \sin^2 \frac{n\pi h}{2L}). \quad (10.17)$$

Nu geldt dus dat voor iedere waarde van A de functie

$$U_n(x,t) = A e^{\alpha_n t} \sin \frac{n\pi x}{L} \quad (10.18)$$

waarin α_n gegeven is door (10.17), een oplossing is van de differentievergelijking (10.5) welke voldoet aan de homogene randvoorwaarden (10.6). Kiezen we nu speciaal $A = B_n$ (10.12) dan voldoet $U_n(x,t)$ eveneens aan de n-de term van de Fourierreeks (10.11) van de beginvoorwaarde. Door nu deze speciale oplossingen $U_n(x,t)$ te sommeren over alle n krijgen we de oplossing $U(x,t)$ van de differentievergelijking welke voldoet aan de beginvoorwaarde (10.11), dus

$$U(x,t) = \sum_{n=1}^{\infty} B_n e^{\alpha_n t} \sin \frac{n\pi x}{L}. \quad (10.19)$$

Het probleem van de convergentie van de methode is nu weer: nadert $U(x,t)$ tot $u(x,t)$ als $h,k \rightarrow 0$ bij constante $\lambda = \frac{ak}{h^2}$.

De oplossing van de differentievergelijking verschilt van de oplossing van de differentiaalvergelijking alleen in de exponenten van de e-machten.

Nu is

$$\begin{aligned}\alpha_n &= \frac{1}{k} \log\left(1 - 4\lambda \sin^2 \frac{n\pi h}{2L}\right) = \\ &= \frac{1}{k} \frac{\log\left(1 - 4\lambda \sin^2 \frac{n\pi h}{2L}\right)}{4\lambda \sin^2 \frac{n\pi h}{2L}} \cdot \frac{4\lambda \sin^2 \frac{n\pi h}{2L}}{\frac{n^2 \pi^2 h^2}{4L^2}} \cdot \frac{n^2 \pi^2 h^2}{4L^2}\end{aligned}$$

waaruit met behulp van de standaardlimieten

$$\lim_{x \rightarrow 0} \frac{\log(1-x)}{x} = -1 \quad \text{en} \quad \lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$$

gemakkelijk volgt dat

$$\lim_{h \rightarrow 0} \alpha_n = -\frac{an^2 \pi^2}{L^2} \quad \text{als} \quad \lambda = \frac{ak}{h^2} \quad \text{constant is.}$$

We zien dus dat voor $h, k \rightarrow 0$ bij constante λ iedere term uit de reeks van de oplossing (10.19) van de differentievergelijking convergeert naar de overeenkomstige term in de oplossing (10.10) van de differentiaalvergelijking.

Als we nu veronderstellen dat de beginvoorwaarde

$$g(x) = \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{L} \quad \text{zodanig is dat} \quad \sum_{n=1}^{\infty} |B_n| \quad \text{convergeert, dan conver-}$$

geert de reeks (10.19) uniform in h en k als $\text{Re } \alpha_n \leq 0$. In dat geval mogen we volgens de stelling op blz. NA-95 de limietovergang termsgewijs uitvoeren. Daaruit volgt dat voor $\text{Re } \alpha_n \leq 0$ geldt: $U(x,t) \rightarrow u(x,t)$ voor $h, k \rightarrow 0$.

Nu is

$$\text{Re } \alpha_n = \frac{1}{k} \log \left| 1 - 4\lambda \sin^2 \frac{n\pi h}{2L} \right|$$

dus $\text{Re } \alpha_n \leq 0$ als $\left| 1 - 4\lambda \sin^2 \frac{n\pi h}{2L} \right| \leq 1$. Aan deze ongelijkheid is voldaan voor iedere waarde van h als $\lambda \leq \frac{1}{4}$. Hiermee hebben we dus gevonden dat de methode convergent is als $\lambda \leq \frac{1}{4}$.

Nemen we nu het geval $\lambda > \frac{1}{4}$. Voor de waarden van n , waarvoor $\frac{nh}{L}$ een on-even geheel getal is, geldt dan

$$1 - 4\lambda \sin^2 \frac{n\pi h}{2L} = 1 - 4\lambda < -1$$

en dus is

$$\alpha_n = \frac{1}{k} \log(1-4\lambda)$$

een complex getal waarvan het reële deel positief is. Weliswaar convergeert ook in dit geval de reeks (10.19) voor iedere h en k , daar $|U_n(x,t)| \leq C \cdot B_n$ waarbij $C = (4\lambda - 1)^{t/k}$, maar deze convergentie is niet uniform in h en k , omdat C niet begrensd is voor k in de buurt van nul. Daarmee vervalt de mogelijkheid om termsgewijs de limietovergang uit te voeren.

Dientengevolge kunnen we de convergentie voor $\lambda > \frac{1}{4}$ dan ook slechts verwachten bij zeer speciale eisen voor de beginvoorwaarde $g(x)$; bijvoorbeeld zeer sterke afname van de coëfficiënten B_n , in het bijzonder $B_n = 0$ voor $n > n_0$.

Stabiliteitsonderzoek voor verschillende differentieformules.

Allereerst onderzoeken we met de Fourier-methode de stabiliteit van de expliciete formule (10.5). Met de afkorting (9.6) wordt deze formule

$$U_{r,s+1} = (1-2\lambda)U_{r,s} + \lambda(U_{r+1,s} + U_{r-1,s}) \quad (10.20)$$

Evenals in het geval van de hyperbolische vergelijking stellen we

$$U_{r,s} = e^{\alpha sk} e^{i\beta rh}$$

Substitutie hiervan in de differentieformule geeft

$$e^{\alpha k} = 1 - 2\lambda + \lambda(e^{i\beta h} + e^{-i\beta h})$$

ofwel

$$e^{\alpha k} = 1 - 4\lambda \sin^2 \frac{\beta h}{2} \quad (10.21)$$

De methode is alleen stabiel als $|e^{\alpha k}| \leq 1$, dus als $\lambda \leq \frac{1}{4}$. Ook in dit geval vinden we dus dat, hoewel soms voor $\lambda > \frac{1}{4}$ de methode wel convergent is, we altijd $\lambda \leq \frac{1}{4}$ moeten nemen om betrouwbare resultaten te krijgen, vanwege de instabiliteit van de methode voor $\lambda > \frac{1}{4}$.

We onderzoeken nu ook weer de mogelijkheid om met impliciete differentieformules grotere stabiliteit te krijgen. In dit geval benaderen we de tweede afgeleide u_{xx} in het punt (x,t) door een gewogen gemiddelde van centrale differenties in de punten $(x,t+k)$ en (x,t) . De afbreekfout is dan weer van dezelfde orde, dus

$$\frac{\partial^2 u}{\partial x^2} = \frac{p\{u(x+h,t+k) - 2u(x,t+k) + u(x-h,t+k)\} + (1-p)\{u(x+h,t) - 2u(x,t) + u(x-h,t)\}}{h^2} + \mathcal{O}(h^2).$$

voor iedere $0 \leq p \leq 1$. Nemen we dan in de t -richting de voorwaartse differentie, dan krijgen we de formule

$$U_{r,s+1} - U_{r,s} = \lambda [p(U_{r+1,s+1} - 2U_{r,s+1} + U_{r-1,s+1}) + (1-p)(U_{r+1,s} - 2U_{r,s} + U_{r-1,s})]$$

ofwel

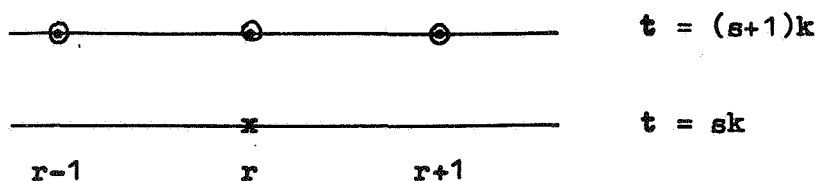
$$\lambda p U_{r+1,s+1} - (2\lambda p + 1) U_{r,s+1} + \lambda p U_{r-1,s+1} = -\lambda(1-p) U_{r+1,s} + (2\lambda - 2\lambda p - 1) U_{r,s} - \lambda(1-p) U_{r-1,s} \quad (10.22)$$

We beschouwen van deze formule enkele speciale gevallen.

Voor $p = 1$ krijgen we de formule

$$\lambda U_{r+1,s+1} - (2\lambda + 1) U_{r,s+1} + \lambda U_{r-1,s+1} = -U_{r,s} \quad (10.23)$$

Het rekenpatroon ziet er met deze formule als volgt uit

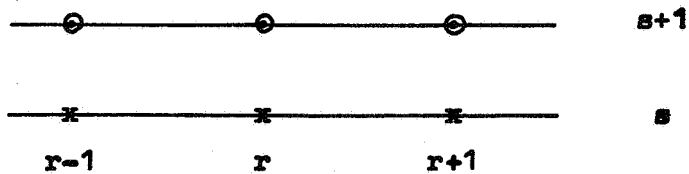


Deze formule staat bekend onder de naam: vierpunts impliciete formule van Laasonen.

Voor $p = \frac{1}{2}$ nemen we juist het gemiddelde van de centrale differenties op de lijnen t en $t+k$. We krijgen in dit geval de formule

$$\begin{aligned} \frac{\lambda}{2} U_{r+1,s+1} - (\lambda + 1) U_{r,s+1} + \frac{\lambda}{2} U_{r-1,s+1} &= \\ &= -\frac{\lambda}{2} U_{r+1,s} + (\lambda - 1) U_{r,s} - \frac{\lambda}{2} U_{r-1,s} \end{aligned} \quad (10.24)$$

Het rekenpatroon is nu



Deze formule heet: zespunts impliciete formule van Crank-Nicolson.

Opmerking. Ook hier geven de impliciete differentieformules aanleiding tot een stelsel lineaire vergelijkingen waarvan de matrix de gedaante van een bandmatrix heeft.

De algemene impliciete differentieformule (10.22) levert bij het stabiliteitsonderzoek met de Fourier-methode de vergelijking

$$2\lambda p \cos \beta h e^{\alpha k} - (2\lambda p + 1)e^{\alpha k} = -2\lambda(1-p)\cos \beta h + (2\lambda - 2\lambda p - 1)$$

waaruit volgt

$$e^{\alpha k} = \frac{-4\lambda(1-p)\sin^2 \frac{\beta h}{2} + 1}{4\lambda p \sin^2 \frac{\beta h}{2} + 1} \quad (10.25)$$

Voor $p = 0$ hebben we het geval van de expliciete formule (10.20). Dan is (10.25) gelijk aan (10.21) en de stabiliteitsvoorwaarde is $\lambda \leq \frac{1}{2}$.

Voor $p = 1$ wordt de vergelijking (10.25)

$$e^{\alpha k} = \frac{1}{1 + 4\lambda \sin^2 \frac{\beta h}{2}} \quad .$$

Voor iedere waarde van λ is $|e^{\alpha k}| \leq 1$, dus de formule van Laasonen is onvoorwaardelijk stabiel.

Voor $p = \frac{1}{2}$ krijgen we

$$e^{\alpha k} = \frac{1 - 2\lambda \sin^2 \frac{\beta h}{2}}{1 + 2\lambda \sin^2 \frac{\beta h}{2}} \quad .$$

Dus ook bij de formule van Crank-Nicolson hebben we onvoorwaardelijke stabiliteit.

De formule van Laasonen kan ook worden opgevat als een differentieformule opgebouwd uit een achterwaartse differentie in de t-richting en een normale centrale differentie in de x-richting.

Men kan zich afvragen of het gebruik van een nauwkeurigere benadering voor de afgeleide naar t bv. de centrale differentie, waarvoor geldt

$$\frac{\partial u}{\partial t} = \frac{u(t+k) - u(t-k)}{2k} + \mathcal{O}(k^2),$$

een bruikbare formule oplevert. De differentieformule kan in dit geval geschreven worden in de vorm

$$U_{r,s+1} - U_{r,s-1} = 2\lambda(U_{r+1,s} - 2U_{r,s} + U_{r-1,s}). \quad (10.26)$$

Het onderzoek naar de stabiliteit van deze formule levert de vergelijking

$$e^{\alpha k} - e^{-\alpha k} = -8\lambda \sin^2 \frac{\beta h}{2}.$$

Voor stabiliteit is vereist dat $\alpha = i\varphi$ (zie blz. NA-100), dan wordt de vergelijking voor φ

$$\sin \varphi = 4 i \lambda \sin^2 \frac{\beta h}{2}$$

en deze vergelijking heeft voor geen enkele waarde van λ een oplossing φ . We zien dus dat, hoewel de differentieformule (10.26) nauwkeuriger is dan alle voorgaande formules, d.w.z. een afbreekfout van hogere orde in k heeft, deze formule totaal ongeschikt is vanwege de instabiliteit voor alle waarden van λ .

Evenals bij de hyperbolische vergelijkingen kan ook hier voor het stabiliteitsonderzoek gebruik gemaakt worden van de matrix-methode. We zullen dit demonstreren aan de differentieformule van Crank-Nicolson, die we schrijven in de vorm

$$U_{r-1,s+1} - 2\left(1 + \frac{1}{\lambda}\right)U_{r,s+1} + U_{r+1,s+1} = -\left\{U_{r-1,s} - 2\left(1 - \frac{1}{\lambda}\right)U_{r,s} + U_{r+1,s}\right\}. \quad (10.27)$$

We nemen aan dat de randvoorwaarden homogeen zijn, immers het gaat om het foutenonderzoek, waarbij we aannemen dat de echte randvoorwaarde goed in rekening worden gebracht.

Analoog als op blz. NA-104 kunnen we het stelsel vergelijkingen (10.27) als volgt in matrix-notatie schrijven

$$A \underline{U}_{s+1} = B \underline{U}_s \quad \text{ofwel} \quad \underline{U}_{s+1} = A^{-1} B \underline{U}_s \quad (10.28)$$

waarbij $A = C - \frac{2}{\lambda} I$ en $B = -(C + \frac{2}{\lambda} I)$ met $I =$ eenheidsmatrix en

en hieruit volgt de relatie tussen β en ν

$$\cos \beta = \frac{2 + \nu}{\nu} . \quad (10.34)$$

De algemene oplossing van (10.33) wordt dan

$$Y_r = c_1 e^{i\beta r} + c_2 e^{-i\beta r}$$

en deze oplossing voldoet aan de randvoorwaarden als $c_1 = -c_2$ en $2ic_1 \sin \beta N = 0$. We krijgen de niet-triviale oplossingen voor $\sin \beta N = 0$ ofwel $\beta = \frac{m\pi}{N}$ met $m = 1, 2, \dots, N-1$.

Invullen in (10.34) geeft

$$\cos \frac{m\pi}{N} = \frac{2 + \nu}{\nu}$$

en hieruit vinden we voor de eigenwaarden ν van C (omdat $Nh = L$)

$$\nu_m = -4 \sin^2 \frac{m\pi h}{2L} . \quad (10.35)$$

De eigenwaarden μ van $A^{-1}B$ volgen dan uit (10.32)

$$\mu_m = \frac{2 - 4\lambda \sin^2 \frac{m\pi h}{2L}}{2 + 4\lambda \sin^2 \frac{m\pi h}{2L}} . \quad (10.36)$$

Hieruit zien we direct dat alle eigenwaarden μ_m voldoen aan $|\mu_m| \leq 1$ voor iedere waarde van $\lambda > 0$. Daarom hebben we nogmaals aangetoond dat de formule van Crank-Nicolson onvoorwaardelijk stabiel is.

We zullen nu ook nog bij Crank-Nicolson de convergentie van de methode onderzoeken, met behulp van de matrix-methode.

We bepalen daartoe eerst de eigenvectoren van de matrix $A^{-1}B$. Zoals we in het voorgaande zagen zijn dit eveneens de eigenvectoren van C . Een eigenvector van C is de niet-triviale oplossing van (10.33) behorende bij een van de eigenwaarden van C . Zoals we hierboven zagen kunnen we de oplossing welke bij ν_m behoort schrijven in de vorm

$$Y_r = c(e^{i\beta r} - e^{-i\beta r}) = 2ic \sin \beta r$$

met $\beta = \frac{m\pi}{N} . \quad (10.37)$

We hebben volgens (10.35) $N-1$ verschillende eigenwaarden. De $N-1$ bijbehorende eigenvectoren zijn dus

$$\underline{w}_m = \left\{ \sin \frac{m\pi hr}{L} \right\} \quad m = 1, 2, \dots, N-1. \quad (10.38)$$

Deze eigenvectoren zijn lineair onafhankelijk. Ze vormen een orthogonale basis in de ruimte van de vectoren \underline{U}_s . De orthogonaliteitsrelaties zijn

$$(\underline{w}_m, \underline{w}_n) = \begin{cases} 0 & m \neq n \\ \frac{N}{2} & m = n \end{cases} .$$

Een willekeurige vector \underline{U} is dan te schrijven als een lineaire combinatie van de basisvectoren \underline{w}_m , op de volgende wijze

$$\underline{U} = \frac{2}{N} \sum_{m=1}^{N-1} (\underline{U}, \underline{w}_m) \underline{w}_m. \quad (10.39)$$

Nu nemen we aan dat als beginvoorwaarde op de lijn $t = 0$ de waarden

$$U_{r,0} = f_r \quad r = 1, 2, \dots, N-1 \quad (10.40)$$

gegeven zijn. Deze beginwaarden vormen voor de vergelijking (10.28) de beginvector $\underline{U}_0 = \{U_{r,0}\}$, welke te schrijven is als lineaire combinatie van de eigenvectoren \underline{w}_m (10.39). Dus

$$\underline{U}_0 = \sum_{m=1}^{N-1} a_m \underline{w}_m$$

waarbij

$$a_m = \frac{2}{N} \sum_{r=1}^{N-1} f_r \sin \frac{m\pi hr}{L}. \quad (10.41)$$

In coördinaten uitgeschreven krijgen we dan

$$U_{r,0} = \sum_{m=1}^{N-1} a_m \sin \frac{m\pi hr}{L} .$$

Dit is de discrete Fourier-ontwikkeling van de beginvoorwaarde.

Voor de vector \underline{U}_s , d.i. de oplossing van (10.27) op het niveau $t = sk$, geldt dan

$$\underline{U}_s = (A^{-1}B)^s \underline{U}_0 = \sum_{m=1}^{N-1} a_m \mu_m^s \underline{w}_m. \quad (10.42)$$

Dus de oplossing van de differentievergelijking van Crank-Nicolson met homogene randvoorwaarden en beginvoorwaarde (10.40) is van de gedaante

$$U_{r,s} = \sum_{m=1}^{N-1} a_m \left(\frac{1 - 2\lambda \sin^2 \frac{m\pi h}{2L}}{1 + 2\lambda \sin^2 \frac{m\pi h}{2L}} \right)^s \sin \frac{m\pi r}{L}$$

waarin de coëfficiënten a_m gegeven zijn door (10.41). Vullen we hierin $x = rh$ en $t = sk$ in, dan krijgen we

$$U(x,t) = \sum_{m=1}^{N-1} a_m \left(\frac{1 - 2\lambda \sin^2 \frac{m\pi h}{2L}}{1 + 2\lambda \sin^2 \frac{m\pi h}{2L}} \right)^{t/k} \sin \frac{m\pi x}{L}. \quad (10.43)$$

Nu geldt

$$\left(1 - 2\lambda \sin^2 \frac{m\pi h}{2L} \right)^{\frac{t}{k}} = e^{\frac{t}{k} \log \left(1 - 2\lambda \sin^2 \frac{m\pi h}{2L} \right)}$$

waaruit met behulp van het resultaat op blz. NA-118 direct volgt

$$\lim_{h,k \rightarrow 0} \left(1 - 2\lambda \sin^2 \frac{m\pi h}{2L} \right)^{\frac{t}{k}} = e^{-a \frac{m^2 \pi^2}{2L^2} t} \quad \text{bij constante } \lambda.$$

Op analoge wijze krijgen we

$$\lim_{h,k \rightarrow 0} \left(1 + 2\lambda \sin^2 \frac{m\pi h}{2L} \right)^{\frac{t}{k}} = e^{a \frac{m^2 \pi^2}{2L^2} t}$$

waaruit we tenslotte vinden

$$\lim_{h,k \rightarrow 0} \left(\frac{1 - 2\lambda \sin^2 \frac{m\pi h}{2L}}{1 + 2\lambda \sin^2 \frac{m\pi h}{2L}} \right)^{\frac{t}{k}} = e^{-a \frac{m^2 \pi^2}{L^2} t}. \quad (10.44)$$

λ constant

Voor de Fourier-coëfficiënten van $U(x,t)$ geldt

$$a_m = \frac{2}{L} \sum_{r=1}^{N-1} h \cdot f_r \sin \frac{m\pi r h}{L} = \frac{2}{L} \sum_{r=0}^N f_r \sin \frac{m\pi r h}{L} \cdot h$$

waaruit volgt

$$\lim_{h \rightarrow 0} a_m = B_m = \frac{2}{L} \int_0^L f(x) \sin \frac{m\pi x}{L} dx. \quad (10.45)$$

We zien dus dat de m-de term uit de (eindige) Fourierreeks (10.43) van $U(x,t)$ convergeert naar de overeenkomstige term uit de Fourierreeks (10.10) van $u(x,t)$. Verder is $N = \frac{L}{h}$ dus als $h \rightarrow 0$ dan $N \rightarrow \infty$, dus de eindige reeks (10.43) gaat voor $h \rightarrow 0$ over in een oneindige reeks. Onder dezelfde aanname als vroeger, nl. $\sum_{n=1}^{\infty} |B_n|$ convergeert, kan men dan bewijzen dat $U(x,t) \rightarrow u(x,t)$ als $h,k \rightarrow 0$, bij constante λ , waarmee dan bewezen is dat de methode van Crank-Nicolson convergent is voor alle waarden van λ .

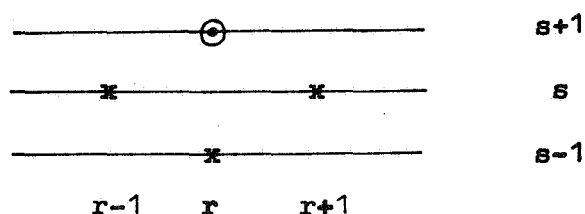
We zullen nu nog een expliciete differentieformule behandelen, welke het voordeel van een impliciete formule heeft, nl. onbeperkte stabiliteit. Om deze formule te krijgen gaan we uit van de totaal instabiele formule (10.26), ook wel formule van Richardson genoemd. We vervangen nu in het rechterlid de term $2U_{r,s}$ door de som van $U_{r,s-1}$ en $U_{r,s+1}$ en krijgen zo de vierpunts expliciete formule van Du Fort and Frankel.

$$U_{r,s+1} - U_{r,s-1} = 2\lambda(U_{r+1,s} - U_{r,s+1} - U_{r,s-1} + U_{r-1,s})$$

ofwel

$$(1 + 2\lambda) U_{r,s+1} = 2\lambda(U_{r+1,s} + U_{r-1,s}) + (1 - 2\lambda) U_{r,s-1} \quad (10.46)$$

Het rekenpatroon van deze formule is



Opmerking. Een moeilijkheid bij de methode van Du Fort and Frankel is de start. Omdat de functiewaarden op een nieuwe lijn worden berekend met behulp van de functiewaarden op de twee voorgaande lijnen, heeft Du Fort and Frankel een aparte startprocedure nodig om de functiewaarden op de lijn $t = k$ te bepalen.

In matrix-vorm ziet de formule (10.46) er als volgt uit

$$A \underline{U}_{s+1} = B \underline{U}_s + C \underline{U}_{s-1} \quad (10.47)$$

Hierin is $A = (1 + 2\lambda)I$ en $C = (1 - 2\lambda)I$ met $I =$ eenheidsmatrix, terwijl B een tridiagonaalmatrix is met nul op de hoofddiagonaal

$$B = \begin{pmatrix} 0 & 2\lambda & & \ominus \\ 2\lambda & 0 & \cdot & \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ \ominus & \cdot & \cdot & 2\lambda \\ & & 2\lambda & 0 \end{pmatrix}$$

Op dezelfde manier als op blz. NA-105 voeren we in

$$\underline{V}_s = \begin{pmatrix} \underline{U}_s \\ \underline{U}_{s-1} \end{pmatrix}$$

en krijgen dan de vergelijking $\underline{V}_{s+1} = P \underline{V}_s$ met

$$P = \begin{pmatrix} A^{-1}B & A^{-1}C \\ I & 0 \end{pmatrix}.$$

Voor stabiliteit is weer vereist dat voor de eigenwaarden μ van P geldt $|\mu| \leq 1$. De eigenwaarden-vergelijking is

$$\begin{vmatrix} A^{-1}B - \mu I & A^{-1}C \\ I & -\mu I \end{vmatrix} = 0$$

welke we gemakkelijk kunnen herleiden tot de vergelijking

$$\det(A\mu^2 - B\mu - C) = 0. \quad (10.48)$$

De matrix $A\mu^2 - B\mu - C$ heeft de gedaante

$$\begin{pmatrix} (1+2\lambda)\mu^2 - (1-2\lambda) & -2\lambda\mu & & \ominus \\ -2\lambda\mu & (1+2\lambda)\mu^2 - (1-2\lambda) & -2\lambda\mu & \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ \ominus & \cdot & \cdot & -2\lambda\mu \\ & -2\lambda\mu & & (1+2\lambda)\mu^2 - (1-2\lambda) \end{pmatrix} \quad (10.49)$$

De hierbijbehorende homogene differentievergelijking is

$$-2\lambda\mu Y_{r+1} + \{ (1+2\lambda)\mu^2 - (1-2\lambda) \} Y_r - 2\lambda\mu Y_{r-1} = 0 \quad (10.50)$$

met de randvoorwaarden

$$Y_0 = Y_N = 0.$$

Stellen we hierin de oplossing $Y_r = e^{i\beta r}$ dan krijgen we de betrekking

$$\cos \beta = \frac{(1+2\lambda)\mu^2 - (1-2\lambda)}{4\lambda\mu} \quad (10.51)$$

en de algemene oplossing

$$Y_r = c(e^{i\beta r} - e^{-i\beta r}).$$

Met behulp van de homogene randvoorwaarden vinden we de voorwaarde voor niet-triviale oplossingen

$$\sin \beta N = 0 \quad \text{ofwel} \quad \beta = \frac{m\pi}{N} \quad m = 1, 2, \dots, N-1.$$

Dit ingevuld in de betrekking (10.51) geeft een vierkantsvergelijking voor μ

$$(1+2\lambda)\mu^2 - 4\lambda\mu \cos \frac{m\pi}{N} - (1-2\lambda) = 0$$

welke voor iedere m twee wortels μ_1 en μ_2 heeft

$$\mu_{1,2} = \frac{2\lambda \cos \frac{m\pi}{N} \pm \sqrt{1 - 4\lambda^2 \sin^2 \frac{m\pi}{N}}}{1 + 2\lambda}.$$

Dan geldt

$$|\mu| \leq \frac{2\lambda |\cos \frac{m\pi}{N}| + \sqrt{1 - 4\lambda^2 \sin^2 \frac{m\pi}{N}}}{1 + 2\lambda} \leq \frac{2\lambda |\cos \frac{m\pi}{N}| + 1}{1 + 2\lambda} \leq 1$$

als $4\lambda^2 \sin^2 \frac{m\pi}{N} \leq 1$, en

$$|\mu| = \frac{\sqrt{4\lambda^2 \cos^2 \frac{m\pi}{N} + (4\lambda^2 \sin^2 \frac{m\pi}{N} - 1)}}{1 + 2\lambda} = \frac{\sqrt{4\lambda^2 - 1}}{1 + 2\lambda} \leq 1$$

als $4\lambda^2 \sin^2 \frac{m\pi}{N} > 1$.

Dus voor alle eigenwaarden μ van P geldt $|\mu| \leq 1$. M.a.w. de methode van Du Fort and Frankel is stabiel voor alle waarden van $\lambda > 0$.

Tot nu toe hebben we als vanzelfsprekend aangenomen dat de differentievergelijking overgaat in de differentiaalvergelijking als de stapgrootten h en k tot nul naderen. Bij een partiële differentiaalvergelijking kan zich echter het merkwaardige verschijnsel voordoen (dat niet voorkomt bij gewone differentiaalvergelijkingen) dat de oplossing van de gebruikte differentievergelijking wel stabiel is, doch niet convergeert naar de oplossing van de bedoelde differentiaalvergelijking, maar naar een andere. We zeggen in zo'n geval dat de differentieformule niet consistent of compatibel is met de differentiaalvergelijking.

We zullen nu de consistentie of compatibiliteit van verschillende behandelde differentieformules onderzoeken.

Het verschil tussen differentievergelijking en differentiaalvergelijking hebben we de afbreekfout (truncation error) genoemd, en het is duidelijk dat voor een compatibele formule de afbreekfout naar nul gaat als $h, k \rightarrow 0$.

De differentievergelijkingen die we hebben afgeleid voor het warmtegeleidingsprobleem kunnen we schrijven in de algemene gedaante

$$L U(x, t) = 0 \quad (10.52)$$

waarin L een lineaire operator voorstelt, de zg. differentie-operator. De exacte oplossing van de differentiaalvergelijking voldoet dan aan de vergelijking

$$L u(x, t) = T \quad (10.53)$$

waarbij T de lokale afbreekfout voorstelt. De vorm van T kunnen we bepalen door gebruik te maken van Taylorreeksontwikkelingen van $u(x, t)$.

Bij de expliciete formule

$$L U(x, t) = \frac{U(x, t+k) - U(x, t)}{k} - a \frac{U(x+h, t) - 2U(x, t) + U(x-h, t)}{h^2}$$

vinden we als we de Taylorreeksen (zie blz. NA-114) van $u(x, t)$ invullen

$$L u(x, t) = T = \left(\frac{\partial u}{\partial t} + \frac{k}{2} \frac{\partial^2 u}{\partial t^2} + \frac{k^2}{6} \frac{\partial^3 u}{\partial t^3} + \dots \right) - a \left(\frac{\partial^2 u}{\partial x^2} + \frac{h^2}{12} \frac{\partial^4 u}{\partial x^4} + \frac{h^4}{360} \frac{\partial^6 u}{\partial x^6} + \dots \right).$$

Op dezelfde manier vinden we voor de formule van Crank-Nicolson

$$T = \left(\frac{\partial u}{\partial t} + \frac{k}{2} \frac{\partial^2 u}{\partial t^2} + \frac{k^2}{6} \frac{\partial^3 u}{\partial t^3} + \dots \right) + \\ - a \left(1 + \frac{1}{2} k \frac{\partial}{\partial t} + \frac{1}{4} k^2 \frac{\partial^2}{\partial t^2} + \dots \right) \left(\frac{\partial^2 u}{\partial x^2} + \frac{h^2}{12} \frac{\partial^4 u}{\partial x^4} + \frac{h^4}{360} \frac{\partial^6 u}{\partial x^6} + \dots \right). \quad (10.55)$$

De differentiaal-operator $\left(1 + \frac{1}{2} k \frac{\partial}{\partial t} + \frac{1}{4} k^2 \frac{\partial^2}{\partial t^2} + \dots \right)$ komt hierbij overeen

met het middelen in de t -richting van de centrale differenties.

De afbreekfout bij Du Fort and Frankel tenslotte is

$$T = \left(\frac{\partial u}{\partial t} + \frac{k^2}{6} \frac{\partial^3 u}{\partial t^3} + \dots \right) - a \left(\frac{\partial^2 u}{\partial x^2} + \frac{h^2}{12} \frac{\partial^4 u}{\partial x^4} + \dots - \frac{k^2}{h^2} \frac{\partial^2 u}{\partial t^2} - \frac{k^4}{12h^2} \frac{\partial^4 u}{\partial t^4} - \dots \right).$$

We hebben nu dus de volgende formules gevonden

Expliciete formule

$$L u - \left(\frac{\partial u}{\partial t} - a \frac{\partial^2 u}{\partial x^2} \right) = \left(\frac{k}{2} \frac{\partial^2 u}{\partial t^2} - \frac{ah^2}{12} \frac{\partial^4 u}{\partial x^2} \right) + \dots$$

Crank-Nicolson

$$L u - \left(\frac{\partial u}{\partial t} - a \frac{\partial^2 u}{\partial x^2} \right) = \frac{k}{2} \frac{\partial}{\partial t} \left(\frac{\partial u}{\partial t} - a \frac{\partial^2 u}{\partial x^2} \right) - \frac{ah^2}{12} \frac{\partial^4 u}{\partial x^4} + \\ + \frac{k^2}{6} \left(\frac{\partial^3 u}{\partial t^3} - \frac{3a}{2} \frac{\partial^4 u}{\partial x^2 \partial t^2} \right) + \dots$$

Du Fort and Frankel

$$L u - \left(\frac{\partial u}{\partial t} - a \frac{\partial^2 u}{\partial x^2} \right) = \frac{k^2}{6} \frac{\partial^3 u}{\partial t^3} - \frac{ah^2}{12} \frac{\partial^4 u}{\partial x^4} + \frac{ak^2}{h^2} \frac{\partial^2 u}{\partial t^2} + \dots$$

In de formule van Crank-Nicolson zien we dat de term $O(k)$ wegvalt, althans indien de hogere afgeleiden van $u(x,t)$ bestaan. Daardoor wordt de afbreekfout $O(h^2)$ en $O(k^2)$, m.a.w. in k een orde beter in vergelijking met de formule van Laasonen.

Verder zien we dat zowel voor de expliciete formule als voor Crank-Nicolson geldt dat de afbreekfout tot nul nadert als $h, k \rightarrow 0$. De bijbehorende differentievergelijking gaat dus in de limiet over in de bedoelde differentiaalvergelijking voor de warmtegeleiding. We spreken in dit geval van consistente of compatibele formules.

Bij Du Fort and Frankel is de situatie enigszins anders. In dit geval gaat de differentievergelijking alleen over in de differentiaalvergelijking als k sneller naar nul gaat dan h . Indien h en k even snel naar nul gaan, dus als k/h constant is, dan lossen we in feite een andere differentiaalvergelijking op, nl. de hyperbolische vergelijking

$$\frac{\partial u}{\partial t} - a \frac{\partial^2 u}{\partial x^2} + ac^2 \frac{\partial^2 u}{\partial t^2} = 0.$$

Het is gemakkelijk in te zien, dat de begrippen convergentie, consistentie en stabiliteit niet als geheel onafhankelijk te beschouwen zijn. Door Richtmyer (en John) is bewezen dat de consistentie en stabiliteit veelal convergentie insluiten.

11. Elliptische partiële differentiaalvergelijkingen

We hebben in de voorafgaande paragrafen gezien dat bij hyperbolische vergelijkingen de reële karakteristieken een overheersende invloed hebben op de methode om deze vergelijkingen op te lossen. Bij de parabolische vergelijkingen waren de karakteristieken wel reëel, maar doordat ze samenvielen, konden we er geen gebruik van maken. Nu in het geval van de elliptische differentiaalvergelijkingen in twee dimensies hebben we geen reële karakteristieken. De karakteristieken zijn toegevoegd complex, en dus is er nu ook geen mogelijkheid om ze te gebruiken bij een numerieke oplossingsmethode.

Ook in een ander opzicht onderscheiden de elliptische problemen zich van de hyperbolische en parabolische, namelijk dat beginwaarde-problemen ofwel Cauchy-problemen niet zinvol zijn voor elliptische differentiaalvergelijkingen.

We spreken, in navolging van Hadamard, van een goed-gesteld probleem (well-posed problem) als kleine veranderingen in de gegeven voorwaarden geen catastrofale verandering van de oplossing ten gevolge hebben. Anders gezegd, een probleem is goed-gesteld, als de oplossing éénduidig en continu van de nevenvoorwaarden afhangt.

We zullen nu een voorbeeld geven van een probleem dat niet goed-gesteld is. Het is een Cauchyprobleem voor de vergelijking van Laplace namelijk

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad y > 0$$

met de beginvoorwaarden

$$\begin{aligned} u(x, 0) &= f(x) \\ u_y(x, 0) &= g(x). \end{aligned}$$

Nemen we hierin speciaal

$$f(x) = e^{-\sqrt{n}} \sin nx \quad g(x) = 0$$

dan is de oplossing van dit probleem

$$u(x, y) = e^{-\sqrt{n}} \cosh ny \sin nx.$$

Laten we nu n naar oneindig gaan, dan gaat $f(x)$ naar nul voor iedere x , evenals alle afgeleiden van $f(x)$. Was het probleem goed-gesteld, dan zou ook $u(x, y) \rightarrow 0$ als $n \rightarrow \infty$; echter $u(x, y) \rightarrow \infty$ voor alle $y \neq 0$.

Een beginwaarde-probleem voor de vergelijking van Laplace is kennelijk een niet goed-gesteld probleem en kan daarom ook geen reële fysische betekenis hebben. Bij elliptische vergelijkingen zullen we dan ook altijd te maken hebben met randwaarde-problemen.

Daarbij spreken we van een eerste randwaarde-probleem of Dirichlet-probleem als op de rand de functiewaarden gegeven zijn, en van een tweede randwaarde-probleem of Neumann-probleem als op de rand de normale afgeleiden gegeven zijn. Een probleem waarbij op de rand een relatie tussen functiewaarden en afgeleiden gegeven is zullen we een derde randwaarde-probleem noemen.

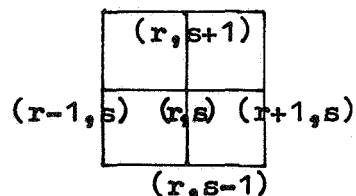
Als voorbeeld van een elliptische vergelijking beschouwen we nu de potentiaalvergelijking of vergelijking van Laplace

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \quad (11.1)$$

Bij deze differentiaalvergelijking behoort een gebied G begrensd door een gesloten kromme C . De oplossing van het randwaarde-probleem is de functie $u(x,y)$ die in het gebied G voldoet aan de vergelijking (11.1) en op de rand C aan de gegeven randvoorwaarde.

De algemene methode om een dergelijk randwaarde-probleem op te lossen bestaat hierin dat we een rooster over het gebied G leggen, de differentiaalvergelijking vervangen door een geschikte differentievergelijking op de roosterpunten en het stelsel algebraïsche vergelijkingen, dat we op deze manier krijgen, oplossen. Dus het numeriek oplossen van een elliptisch probleem komt meestal neer op het oplossen van een groot stelsel lineaire vergelijkingen; een onderwerp dat apart behandeld zal worden.

De eenvoudigste differentievergelijking krijgen we door het rooster vierkant te nemen, d.w.z. $h = k$, en de partiële afgeleiden in (11.1) te vervangen door centrale differenties. Op deze manier krijgen we dan



$$U_{r+1,s} + U_{r-1,s} + U_{r,s+1} + U_{r,s-1} - 4U_{r,s} = 0. \quad (11.2)$$

Volgens deze differentievergelijking is de functiewaarde in een punt het gemiddelde van de functiewaarden in de vier omliggende punten.

De oplossingen van (11.2) hebben eigenschappen welke de oplossingen van (11.1), dit zijn de harmonische functies, eveneens hebben.

1). Uit het feit dat de waarde $U_{r,s}$ het gemiddelde is van de vier omliggende functiewaarden volgt dat de extrema van de oplossing van (11.2) op de rand liggen. Voor harmonische functies geldt eveneens: de extreme waarden van een op een gesloten gebied harmonische functie worden op de rand aangenomen.

Ook analytische functies hebben deze eigenschap (maximum-modulus theorema).

2). Door herhaald toepassen van de differentievergelijking is het mogelijk de waarde van $U_{r,s}$ uit te drukken in de functiewaarden van de rand.

Hieruit volgt dat de functiewaarden op de rand voldoende zijn voor het bepalen van de oplossing van (11.2) in het gehele gebied.

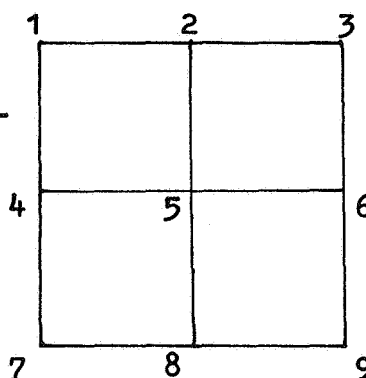
Een harmonische functie is in ieder punt van een gebied uit te drukken in de functiewaarden op de rand met behulp van de functie van Green. De overeenkomstige eigenschap voor een analytische functie wordt gegeven door de integraalformule van Cauchy.

3). Uit de eerste stelling van Green volgt dat voor een harmonische functie $\varphi(x,y)$ geldt

$$\int_C \frac{\partial \varphi}{\partial \underline{n}} ds = 0 \quad (11.3)$$

voor een gesloten kromme C. Hieruit zien we dat bij een Neumann-probleem de waarden van $\frac{\partial \varphi}{\partial \underline{n}}$ op de rand niet willekeurig voorgeschreven kunnen worden, maar dat deze waarden moeten voldoen aan (11.3).

Voor de oplossing van de differentievergelijking (11.2) geldt iets analoogs. We beschouwen daartoe het speciale geval dat we de oplossing van (11.2) willen bepalen in het inwendige van nevenstaande gebied, als op de rand de normale afgeleide $\frac{\partial u}{\partial \underline{n}}$, waarbij \underline{n} de naar buiten gerichte normaal is, gegeven is.



$$\frac{\partial u}{\partial \underline{n}} = -g(x,y).$$

We veronderstellen nog dat in de hoekpunten de beide mogelijke normale afgeleiden samenvallen, m.a.w. dat daar geen discontinuïteit optreedt. In ieder van de negen punten kunnen we de differentievergelijking opstellen. Er zijn drie soorten punten gerepresenteerd door 5, 2 en 1. De differentievergelijking in ieder van deze punten luidt als volgt. Voor punt 5

$$U_2 + U_4 + U_6 + U_8 - 4U_5 = 0.$$

Voor punt 2

$$U_1 + U_3 + U_5 + U_7 - 4U_2 = 0.$$

Met $U_{5^*} - U_5 = -2hg_2$, dit is de discrete voorstelling van de randvoorwaarde, wordt deze vergelijking

$$U_1 + U_3 + 2U_5 - 4U_2 = 2hg_2.$$

Voor punt 1

$$U_2 + U_{2^*} + U_4 + U_{4^*} - 4U_1 = 0.$$

Met $U_{2^*} - U_2 = -2hg_1$ en $U_{4^*} - U_4 = -2hg_1$ krijgen we

$$2U_2 + 2U_4 - 4U_1 = 4hg_1.$$

De coëfficiënten-matrix van het stelsel vergelijkingen, dat we op deze manier krijgen, is singulier. Er is dus geen sprake van een éénduidige oplossing. In het algemeen zal er helemaal geen oplossing zijn, tenzij voor de rechterleden dezelfde afhankelijkheid geldt als voor de linkerleden. Deze voorwaarde is

$$h. \sum_{\substack{m=1 \\ m \neq 5}}^9 g_m = 0 \quad (11.4)$$

en dit is een simpele differentie-benadering voor (11.3). Als de randvoorwaarde g zodanig is dat (11.4) geldt, dan is de oplossing van het differentieprobleem bepaald, op een constante na. Dit is een bekend feit bij een Neumann-probleem.

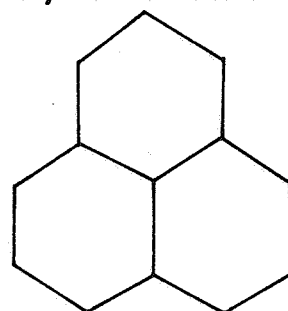
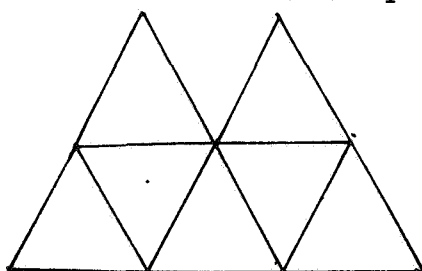
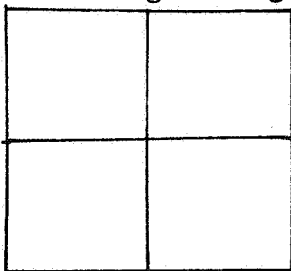
We zullen ons nu gaan bezighouden met het opstellen van eindige-differentie formules voor het geval van de lineaire elliptische differentiaalvergelijking,

$$a \frac{\partial^2 u}{\partial x^2} + b \frac{\partial^2 u}{\partial x \partial y} + c \frac{\partial^2 u}{\partial y^2} + e \frac{\partial u}{\partial x} + f \frac{\partial u}{\partial y} + gu = p \quad (11.5)$$

waarbij $b^2 - 4ac < 0$ en de coëfficiënten a, b, c, e, f, g en p wel mogen afhangen van x en y maar niet van de oplossing u .

De eerste vraag die zich voordoet betreft de keuze van het netwerk.

Indien we eisen dat het veld opgevuld wordt door regelmatige veelhoeken, dan hebben we drie typen van roosters. We kunnen namelijk bewijzen dat van de regelmatige veelhoeken alleen de gelijkzijdige driehoek, het vierkant en de regelmatige zeshoek een vlak kunnen opvullen



Dit bewijs gaat als volgt. Veronderstel dat de regelmatige N -hoek het vlak kan opvullen. Als θ de hoek van de N -hoek is, dan is er een natuurlijk getal n zodanig dat $n\theta = 2\pi$. Verder geldt de relatie $N\theta = (N-2)\pi$. Combinatie van deze twee relaties geeft

$$\frac{2\pi}{n} = \frac{N-2}{N} \pi$$

ofwel $\frac{2N}{N-2}$ is een natuurlijk getal.

Door invullen van $N = 3, 4, 5, 6$ zien we dat alleen $N = 3, 4$ en 6 voldoen. Verder zien we dat voor $N > 6$ het getal

$$\frac{2N}{N-2} = 2 + \frac{4}{N-2}$$

niet geheel is.

Intermezzo

Symbolische operatorenmethode.

Een lineaire operator L is een transformatie van een verzameling functies in een andere verzameling functies, zodanig dat geldt

$$\begin{aligned} 1) L(f + g) &= Lf + Lg \\ 2) L(\alpha f) &= \alpha Lf . \end{aligned} \quad (11.6)$$

Bij allerlei differentieformules in de numerieke analyse spelen de volgende lineaire operatoren een rol.

$$\begin{aligned} \Delta \text{ voorwaartse differentie} & \quad \Delta f(x) = f(x+h) - f(x) \\ \nabla \text{ achterwaartse differentie} & \quad \nabla f(x) = f(x) - f(x-h) \\ \delta \text{ centrale differentie} & \quad \delta f(x) = f(x+\frac{h}{2}) - f(x-\frac{h}{2}) \\ E \text{ verplaatsingsoperator} & \quad Ef(x) = f(x+h) \\ \mu \text{ middelingsoperator} & \quad \mu f(x) = \frac{1}{2}[f(x+\frac{h}{2}) + f(x-\frac{h}{2})] \end{aligned}$$

Met deze vijf lineaire operatoren kunnen we op eenvoudige wijze interpolatieformules afleiden. Daartoe bepalen we eerst een aantal relaties welke tussen deze operatoren gelden.

De n -de macht van een operator wordt gedefinieerd op de voor de hand liggende wijze. Bijvoorbeeld

$$E^n f(x) = E^{n-1} f(x+h) = \dots = f(x+nh).$$

De inverse van een operator vermenigvuldigd met deze operator levert de identiteit. Daarom geldt

$$E^{-1} f(x) = E^{-1} E f(x-h) = f(x-h).$$

Algemeen definiëren we nu voor willekeurige reële α

$$E^{\alpha} f(x) = f(x+\alpha h).$$

Nu gelden de volgende relaties

$$\begin{aligned} \Delta &= E - 1 & \delta &= E^{\frac{1}{2}} - E^{-\frac{1}{2}} \\ \nabla &= 1 - E^{-1} & \mu &= \frac{1}{2}(E^{\frac{1}{2}} + E^{-\frac{1}{2}}). \end{aligned} \quad (11.7)$$

De geldigheid van de eerste relatie volgt uit

$$\Delta f(x) = f(x+h) - f(x) = Ef(x) - f(x) = (E-1)f(x).$$

Voor de overige relaties gaat het bewijs analoog.

Uit de relaties (11.7) kunnen we nog de volgende afleiden

$$\begin{aligned} \Delta &= \nabla E = \delta E^{\frac{1}{2}} & \delta^2 &= E + E^{-1} - 2 \\ \mu \delta &= \frac{1}{2}(E - E^{-1}) & \mu^2 &= \frac{1}{4}(E + E^{-1} + 2) \\ & & \mu^2 &= 1 + \frac{1}{4}\delta^2 \end{aligned} \quad (11.8)$$

Als voorbeeld leiden we de interpolatieformule van Newton af.

$$\begin{aligned} y(x+\alpha h) &= E^\alpha y(x_0) = (1+\Delta)^\alpha y(x_0) \\ &= [1 + \alpha\Delta + \binom{\alpha}{2}\Delta^2 + \binom{\alpha}{3}\Delta^3 + \dots] y(x_0). \end{aligned}$$

Dus

$$y(x_0 + \alpha h) = y(x_0) + \alpha\Delta y(x_0) + \binom{\alpha}{2}\Delta^2 y(x_0) + \dots + \binom{\alpha}{n}\Delta^n y(x_0) + R(x).$$

Op soortgelijke wijze kunnen we de interpolatieformules van Gauss, Everett, Bessel etc afleiden.

We voeren nu nog de volgende operator in

$$D \quad \text{differentiaal operator} \quad D f(x) = \frac{df}{dx}$$

waarmee we formules voor de afgeleiden kunnen maken.

Volgens de formule van Taylor geldt

$$\begin{aligned} E y(x) = y(x+h) &= y(x) + hy'(x) + \frac{h^2}{2!} y''(x) + \frac{h^3}{3!} y'''(x) + \dots \\ &= [1 + hD + \frac{h^2 D^2}{2!} + \frac{h^3 D^3}{3!} + \dots] y(x) \\ &= e^{hD} y(x). \end{aligned}$$

We definiëren dus de operator e^{hD} door middel van de machtreeksontwikkeling van de functie e^x . Op deze manier vinden we de relatie

$$E = e^{hD}.$$

We kunnen nu ook nog andere relaties afleiden

$$\delta = E^{\frac{1}{2}} - E^{-\frac{1}{2}} = e^{\frac{hD}{2}} - e^{-\frac{hD}{2}} = 2 \sinh \frac{hD}{2}.$$

Op analoge wijze geldt

$$\mu = \cosh \frac{hD}{2}.$$

Omgekeerd kunnen we ook schrijven

$$hD = 2 \sinh^{-1} \frac{1}{2} \delta \quad \text{of} \quad hD = \log E = \log (1+\Delta).$$

Toepassing

$$\begin{aligned} hy'_0 &= hDy_0 = \log (1+\Delta)y_0 \\ &= (\Delta - \frac{1}{2}\Delta^2 + \frac{1}{3}\Delta^3 \dots)y_0. \end{aligned}$$

De coëfficiënten nemen zeer langzaam af, daarom zijn centrale differentieformules beter. Bijvoorbeeld

$$\begin{aligned} h^2 y''_0 &= h^2 D^2 y_0 = (2 \sinh^{-1} \frac{\delta}{2})^2 y_0 \\ &= (\delta^2 - \frac{1}{12} \delta^4 + \frac{1}{90} \delta^6 \dots) y_0, \end{aligned}$$

want de reeksontwikkeling van $\sinh^{-1} x$, d.i. de inverse functie van $\sinh x$, is

$$\sinh^{-1} x = x - \frac{1}{6} x^3 + \frac{3}{40} x^5 \dots$$

Als we centrale differenties willen gebruiken voor de eerste afgeleide y'_0 , dan komen we in moeilijkheden als we alleen beschikken over een tabel van de functie $y(x)$.

$$h y'_0 = h D y_0 = (2 \sinh^{-1} \frac{\delta}{2}) y_0$$

geeft een ontwikkeling in oneven centrale differenties van y_0 , en deze komen niet voor in de differentietabel. Daarom passen we een kunstgreep toe. In plaats van centrale differenties nemen we gemiddelde centrale differenties van oneven orde. Deze kunnen we wel uit de differentietabel halen, want

$$\mu \delta^{2n+1} = \mu \delta \cdot \delta^{2n} = \frac{1}{2} (E - E^{-1}) \delta^{2n}.$$

We krijgen de gemiddelde centrale differenties op de volgende manier

$$\begin{aligned} h y'_0 &= (2 \sinh^{-1} \frac{\delta}{2}) y_0 = (1 + \frac{1}{3} \delta^2)^{-\frac{1}{2}} (2 \sinh^{-1} \frac{\delta}{2}) \mu y_0 \\ &= (\mu \delta - \frac{1}{6} \mu \delta^3 + \frac{1}{30} \mu \delta^5 \dots) y_0. \end{aligned}$$

Met behulp van bovenstaande formules kunnen we op eenvoudige wijze differentieformules voor de afgeleiden vinden. Centrale differentieformules krijgen we met $hD = 2 \sinh^{-1} \frac{\delta}{2}$

$$\frac{dy_n}{dx} = \frac{1}{h} \mu (\delta - \frac{1}{6} \delta^3 + \frac{1}{30} \delta^5 \dots) y_n$$

$$\frac{d^2 y_n}{dx^2} = \frac{1}{h^2} (\delta^2 - \frac{1}{12} \delta^4 + \frac{1}{90} \delta^6 \dots) y_n$$

$$\frac{d^4 y_n}{dx^4} = \frac{1}{h^4} (\delta^4 - \frac{1}{6} \delta^6 + \frac{7}{240} \delta^8 \dots) y_n$$

en voorwaartse differentieformules met $hD = \log (1 + \Delta)$

$$\frac{dy_n}{dx} = \frac{1}{h} (\Delta - \frac{1}{2} \Delta^2 + \frac{1}{3} \Delta^3 \dots) y_n.$$

In numerieke processen benaderen we iedere afgeleide door een differentieformule, die we na een aantal termen afbreken. De afbreekfout, die we daardoor krijgen, wordt in het algemeen bepaald door de eerste weggelaten term in de differentieformule. Is deze eerste weggelaten term een n -de orde differentie, dan is de afbreekfout van de orde h^n . De differentieformules kunnen we dan weer uitdrukken in de functiewaarden. Bijvoorbeeld

$$\begin{aligned}\frac{dy_n}{dx} &= \frac{1}{h} \mu \delta y_n + O(h^2) = \frac{y_{n+1} - y_{n-1}}{2h} + O(h^2) \\ \frac{dy_n}{dx} &= \frac{1}{h} \Delta y_n + O(h) = \frac{y_{n+1} - y_n}{h} + O(h) \\ \frac{dy_n}{dx} &= \frac{1}{h} \mu \left(\delta - \frac{1}{6} \delta^3 \right) y_n + O(h^4) \\ &= \frac{1}{12h} (-y_{n+2} + 8y_{n+1} - 8y_{n-1} + y_{n-2}) + O(h^4) \\ \frac{dy_n}{dx} &= \frac{1}{2h} (\Delta - \frac{1}{2} \Delta^2) y_n + O(h^2) \\ &= \frac{1}{2h} (-y_{n+2} + 4y_{n+1} - 3y_n) + O(h^2)\end{aligned}\tag{11.9}$$

en voor de tweede afgeleide

$$\begin{aligned}\frac{d^2 y_n}{dx^2} &= \frac{y_{n+1} - 2y_n + y_{n-1}}{h^2} + O(h^2) \\ \frac{d^2 y_n}{dx^2} &= \frac{-y_{n+2} + 16y_{n+1} - 30y_n + 16y_{n-1} - y_{n-2}}{12 h^2} + O(h^4).\end{aligned}\tag{11.10}$$

Voor de partiële afgeleiden kunnen we gelijksoortige formules geven als voor de gewone afgeleiden, bijvoorbeeld

$$\begin{aligned}h \frac{\partial}{\partial x} &= h D_x = \mu_x \delta_x - \frac{1}{6} \mu_x \delta_x^3 + \dots \\ h^2 \frac{\partial^2}{\partial x^2} &= h^2 D_x^2 = \delta_x^2 - \frac{1}{12} \delta_x^4 + \dots\end{aligned}$$

waarbij de index aangeeft welke variabele beschouwd wordt.

Dus

$$\frac{\partial U_{r,s}}{\partial x} = \frac{1}{h} (U_{r+1,s} - U_{r,s}) + O(h),$$

$$\frac{\partial U_{r,s}}{\partial x} = \frac{1}{2h} (U_{r+1,s} - U_{r-1,s}) + O(h^2).$$

Evenzo geldt

$$\frac{\partial U_{r,s}}{\partial y} = \frac{1}{2k} (U_{r,s+1} - U_{r,s-1}) + O(k^2).$$

Voor de gemengde tweede afgeleide geldt

$$\frac{\partial^2 U_{r,s}}{\partial x \partial y} = \frac{\partial}{\partial x} \left(\frac{\partial U_{r,s}}{\partial y} \right) = \frac{1}{h} \mu_x (\delta_x - \frac{1}{6} \delta_x^3 \dots) \cdot \frac{1}{k} \mu_y (\delta_y - \frac{1}{6} \delta_y^3 \dots) U_{r,s}.$$

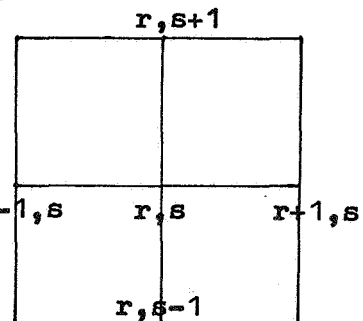
Voor $h = k$ krijgen we dan

$$\frac{\partial^2 U_{r,s}}{\partial x \partial y} = \frac{1}{4h^2} (U_{r+1,s+1} - U_{r-1,s+1} - U_{r+1,s-1} + U_{r-1,s-1}) + O(h^2).$$

Einde intermezzo.

We passen de centrale differentieformules toe op de lineaire partiele differentiaalvergelijking (11.5) waaruit de gemengde term weggelaten is, dus $b = 0$. Door van iedere formule alleen de eerste term te nemen krijgen we

$$\begin{aligned} & \frac{1}{h^2} [(a + \frac{1}{2}eh) u_{r+1,s} + (a - \frac{1}{2}eh) u_{r-1,s}] + \\ & + \frac{1}{k^2} [(c + \frac{1}{2}fk) u_{r,s+1} + (c - \frac{1}{2}fk) u_{r,s-1}] + \quad (11.11) \\ & - (\frac{2a}{h^2} + \frac{2c}{k^2} - g) u_{r,s} + C u_{r,s} = p. \end{aligned}$$



Als we de term $C u_{r,s}$, de afbreekfout, weglaten krijgen we dus een vijfpuntsformule. De coëfficiënten a, c etc. moeten berekend worden in het punt (r, s) . De operator C heeft de gedaante

$$\begin{aligned} C = & \frac{1}{h^2} \{ a(-\frac{1}{12} \delta_x^4 + \frac{1}{90} \delta_x^6 \dots) + eh(-\frac{1}{6} \mu_x \delta_x^3 + \frac{1}{30} \mu_x \delta_x^5 \dots) \} + \\ & + \frac{1}{k^2} \{ c(-\frac{1}{12} \delta_y^4 + \frac{1}{90} \delta_y^6 \dots) + fk(-\frac{1}{6} \mu_y \delta_y^3 + \frac{1}{30} \mu_y \delta_y^5 \dots) \}. \quad (11.12) \end{aligned}$$

De vorm (11. van de lokale afbreekfout in de vijfpuntsformule kunnen we gebruiken als we in tweede instantie een nauwkeuriger resultaat willen krijgen door hogere differenties mee te nemen. Maar in andere gevallen hebben we liever een afbreekfout uitgedrukt in afgeleiden. We kunnen bijvoorbeeld schrijven

$$\delta_x^2 u_{r,s} = h^2 \frac{\partial^2 u_{r,s}}{\partial x^2} + \frac{1}{12} h^4 \frac{\partial^4 u(\xi_1, y_s)}{\partial x^4} \quad (11.13)$$

waarin $x_r - h < \xi_1 < x_r + h$ en de vierde afgeleide continu verondersteld wordt, en

$$\mu_x \delta_x u_{r,s} = h \frac{\partial u_{r,s}}{\partial x} + \frac{1}{6} h^3 \frac{\partial^3 u(\xi_2, y_s)}{\partial x^3} \quad (11.14)$$

waarin ξ_2 tot hetzelfde interval behoort en de derde afgeleide continu moet zijn. Met de gelijksoortige formules voor de y-differenties vinden we voor de afbreekfout in (11.11)

$$\begin{aligned} C u_{r,s} = & -h^2 \left\{ \frac{1}{12} a \frac{\partial^4 u(\xi_1, y_s)}{\partial x^4} + \frac{1}{6} e \frac{\partial^3 u(\xi_2, y_s)}{\partial x^3} \right\} + \\ & -k^2 \left\{ \frac{1}{12} c \frac{\partial^4 u(x_r, \eta_1)}{\partial y^4} + \frac{1}{6} f \frac{\partial^3 u(x_r, \eta_2)}{\partial x^3} \right\} = O(h^2) + O(k^2). \end{aligned} \quad (11.15)$$

Als $b \neq 0$ in (11.5) dan benaderen we de gemengde tweede afgeleide door de differentieformule

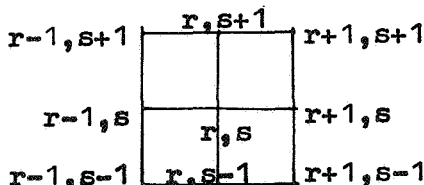
$$\frac{\partial^2 u_{r,s}}{\partial x \partial y} = \frac{u_{r+1,s+1} - u_{r-1,s+1} - u_{r+1,s-1} + u_{r-1,s-1}}{4hk} + T \quad (11.16)$$

met de afbreekfout

$$T = \frac{1}{6} h^2 \frac{\partial^4 u(\xi_3, \eta_3)}{\partial x^3 \partial y} + \frac{1}{6} k^2 \frac{\partial^4 u(\xi_4, \eta_4)}{\partial x \partial y^3} = O(h^2) + O(k^2).$$

Hierin is $x_r - h < \xi_3, \xi_4 < x_r + h$ en $y_s - k < \eta_3, \eta_4 < y_s + k$.

We zien dat als de gemengde term in de differentiaalvergelijking (11.5) voorkomt, de differentieformule een negenpuntsformule is.



Opmerking. Het is wel mogelijk door een coördinatentransformatie de gemengde term te elimineren en daardoor de eenvoudigere vijfpuntsformule te gebruiken, maar dit is niet altijd een voordeel.

Vaak is de randkromme simpel, bijvoorbeeld een rechthoek evenwijdig aan de assen. Door de coördinatentransformatie gaat deze eenvoudige vorm dan verloren.

In verschillende belangrijke partiële differentiaalvergelijkingen met een simpele differentiaal operator, zoals de vergelijking van Laplace of de vergelijking van Poisson

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x,y) \quad (11.17)$$

kunnen we de lokale afbreekfout reduceren, door een relatief klein aantal punten meer op te nemen in de eindige differentieformules. Daartoe is het noodzakelijk dat de stapgrootten in beide richtingen dezelfde zijn, dus $h = k$.

Verder zullen we hierbij gebruik maken van de symbolische operatoren methode.

Omdat zowel de differentiaalvergelijking als het rooster symmetrisch is in x en y beschouwen we de volgende symmetrische sommen

$$\begin{aligned} S_1 &= u_{r+1,s} + u_{r-1,s} + u_{r,s+1} + u_{r,s-1} \\ S_2 &= u_{r+1,s+1} + u_{r+1,s-1} + u_{r-1,s+1} + u_{r-1,s-1} \\ S_3 &= u_{r+2,s} + u_{r-2,s} + u_{r,s+2} + u_{r,s-2} \end{aligned}$$

waarvan de punten respectievelijk op afstand h , $h\sqrt{2}$ en $2h$ van het centrale punt (r,s) gelegen zijn.

Met behulp van de relatie $E = e^{hD}$ kunnen we deze sommen als volgt schrijven

$$\begin{aligned} S_1 &= (e^{hD_x} + e^{-hD_x} + e^{hD_y} + e^{-hD_y})u_{r,s} = 2(\cosh hD_x + \cosh hD_y)u_{r,s} \\ &= 2\left(1 + \frac{h^2 D_x^2}{2!} + \frac{h^4 D_x^4}{4!} + \dots + 1 + \frac{h^2 D_y^2}{2!} + \frac{h^4 D_y^4}{4!} + \dots\right)u_{r,s} \\ &= \left[4 + h^2(D_x^2 + D_y^2) + \frac{1}{12} h^4 \{(D_x^2 + D_y^2)^2 - 2D_x^2 D_y^2\} + \right. \\ &\quad \left. + \frac{1}{360} h^6 \{(D_x^2 + D_y^2)^3 - 3(D_x^2 + D_y^2)D_x^2 D_y^2\} + \dots\right]u_{r,s} \quad (11.18) \end{aligned}$$

Voor de tweede som vinden we

$$\begin{aligned} S_2 &= (4 \cosh hD_x \cdot \cosh hD_y)u_{r,s} \\ &= \left[4 + 2h^2(D_x^2 + D_y^2) + \frac{1}{6} h^4 \{(D_x^2 + D_y^2)^2 + 4D_x^2 D_y^2\} + \right. \\ &\quad \left. + \frac{1}{180} h^6 \{(D_x^2 + D_y^2)^3 + 12(D_x^2 + D_y^2)D_x^2 D_y^2\} + \dots\right]u_{r,s} \quad (11.19) \end{aligned}$$

Op dezelfde wijze vinden we voor de derde som

$$\begin{aligned} S_3 &= \left[4 + 4h^2(D_x^2 + D_y^2) + \frac{4}{3} h^4 \{(D_x^2 + D_y^2)^2 - 2D_x^2 D_y^2\} + \right. \\ &\quad \left. + \frac{8}{45} h^6 \{(D_x^2 + D_y^2)^3 - 3(D_x^2 + D_y^2)D_x^2 D_y^2\} + \dots\right]u_{r,s} \quad (11.20) \end{aligned}$$

Nemen we uit S_1 alleen de eerste twee termen, dan krijgen we de eerder genoemde vijfpuntsformule, namelijk

$$S_1 = u_{r+1,s} + u_{r-1,s} + u_{r,s+1} + u_{r,s-1} = 4u_{r,s} + h^2(D_x^2 + D_y^2)u_{r,s} + Cu_{r,s}$$

ofwel

$$(D_x^2 + D_y^2)u_{r,s} = \frac{1}{h^2}(u_{r+1,s} + u_{r-1,s} + u_{r,s+1} + u_{r,s-1} - 4u_{r,s}) + T$$

met de afbreekfout

$$T = \frac{1}{12} h^2 \{(D_x^2 + D_y^2)^2 - 2D_x^2 D_y^2\} u_{r,s} + O(h^4). \quad (11.21)$$

Nu geldt voor de vergelijking van Laplace

$$(D_x^2 + D_y^2)u_{r,s} = 0.$$

Dus in dit geval is de afbreekfout T (11.21) van de vorm

$$T = -\frac{1}{6} h^2 D_x^2 D_y^2 u_{r,s} + O(h^4). \quad (11.22)$$

Voor de vergelijking van Poisson (11.17) geldt

$$(D_x^2 + D_y^2)^2 u_{r,s} = (D_x^2 + D_y^2) f_{r,s}$$

en daar $f(x,y)$ gegeven is, is deze uitdrukking te berekenen. De afbreekfout T is dan in dit geval eveneens van de vorm (11.22). Het wegvallen van een gedeelte van de afbreekfout (11.21), geeft geen verbetering in de orde van de fout; deze blijft h^2 .

Door een combinatie van S_1 en S_2 te nemen kunnen we de term $D_x^2 D_y^2 u_{r,s}$ elimineren. Daartoe moeten we nemen

$$4S_1 + S_2 = [20 + 6h^2(D_x^2 + D_y^2) + \frac{1}{2}h^4(D_x^2 + D_y^2)^2 + O(h^6)] u_{r,s}$$

waaruit volgt

$$(D_x^2 + D_y^2)u_{r,s} = \frac{1}{6h^2} (4S_1 + S_2 - 20u_{r,s}) - \frac{1}{12} h^2 (D_x^2 + D_y^2) u_{r,s} + O(h^4). \quad (11.23)$$

Ook hier geldt dat $(D_x^2 + D_y^2) u_{r,s}$ verdwijnt, resp. te berekenen is.

Dus de afbreekfout is in dit geval $O(h^4)$. Het is zelfs nog gunstiger

want men kan gemakkelijk inzien dat ook de term met h^4 verdwijnt, resp.

te berekenen is, dus de afbreekfout is $O(h^6)$.

De gevonden formule (11.23) is de negenpuntsformule. Deze formule is voor de vergelijkingen van Laplace en Poisson dus een zeer nauwkeurige formule.

Een andere formule die ook negenpunten gebruikt krijgen we door een combinatie van S_1 en S_3 te nemen. Voor de combinatie $16S_1 - S_3$ kan men eenvoudig bewijzen dat de term $D_x^2 D_y^2 u_{r,s}$ geelimineerd wordt, en dat de afbreekfout dezelfde nauwkeurigheid heeft, dus $\mathcal{O}(h^6)$. Deze formule heeft echter het grote nadeel dat aan de randen punten buiten het gebied meegenomen moeten worden. Dit in tegenstelling tot de negenpuntsformule (11.23), die daarom ook de voorkeur verdient.

Soms zijn we niet geïnteresseerd in de oplossing van de differentiaalvergelijking maar willen we een van de afgeleiden weten. We leiden daarom ook nog een uitdrukking af voor de gemengde tweede afgeleide. We veronderstellen het rooster weer symmetrisch in x-en y-richting, dus $h = k$, en gebruiken de symbolische operatorenmethode. We gaan uit van de benaderingsformule voor de gemengde tweede afgeleide op blz NA-140 en krijgen dan

$$\begin{aligned} & \frac{1}{4h^2} (u_{r+1,s+1} - u_{r-1,s+1} + u_{r-1,s-1} - u_{r+1,s-1}) = \\ & = \frac{1}{h^2} \sinh hD_x \cdot \sinh hD_y u_{r,s} \\ & = \frac{1}{h^2} \left(hD_x + \frac{h^3 D_x^3}{3!} + \dots \right) \left(hD_y + \frac{h^3 D_y^3}{3!} + \dots \right) u_{r,s} \end{aligned}$$

zodat

$$\begin{aligned} & \frac{1}{4h^2} (u_{r+1,s+1} - u_{r-1,s+1} + u_{r-1,s-1} - u_{r+1,s-1}) = \\ & D_x D_y \left[1 + \frac{1}{6} h^2 (D_x^2 + D_y^2) + \frac{1}{120} h^4 \{ (D_x^2 + D_y^2)^2 + \frac{4}{3} D_x^2 D_y^2 \} + \dots \right] u_{r,s} \end{aligned} \quad (11.24)$$

hetgeen in overeenstemming is met (11.16).

Indien nu $u_{r,s}$ voldoet aan de vergelijking van Laplace of Poisson dan kunnen we $\frac{\partial^2 u}{\partial x \partial y}$ berekenen uit het linkerlid van (11.24) en eventuele bekende grootheden met een afbreekfout van de orde h^4 .

Is $u_{r,s}$ niet de oplossing van de differentiaalvergelijking maar van een differentievergelijking die de vergelijking van Laplace, resp. Poisson, benadert dan geldt niet meer

$$(D_x^2 + D_y^2) u_{r,s} = 0 \quad \text{resp.} = f_{r,s}$$

maar wel

$$(D_x^2 + D_y^2) u_{r,s} = \mathcal{O}(h^2) \quad \text{resp.} f_{r,s} + \mathcal{O}(h^2)$$

zodat ook in dit geval $\frac{\partial^2 u}{\partial x \partial y}$ verkregen wordt met een afbreekfout van de orde h^4 .

Een aantal driedimensionale problemen is terug te brengen tot tweedimensionale doordat er symmetrie aanwezig is. Is het probleem rotatiesymmetrisch dan werken we graag met poolcoördinaten. Daarom beschouwen we de vergelijking van Laplace nu eens voor de poolcoördinaten

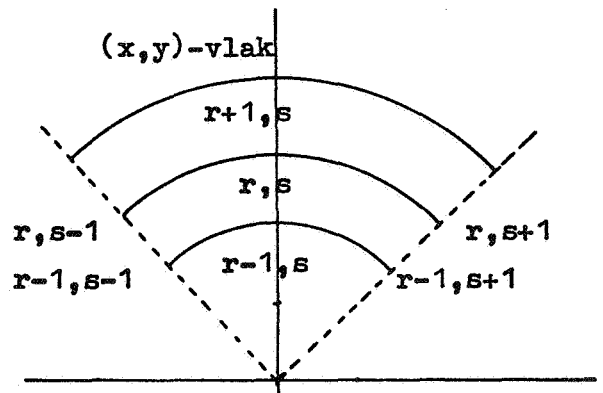
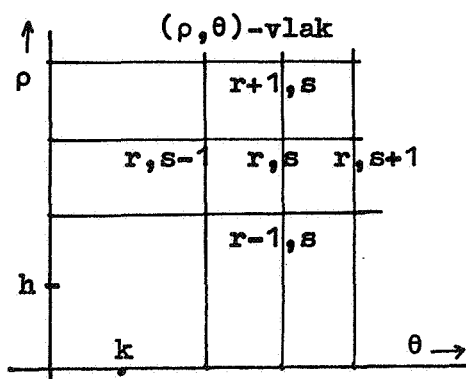
$$x = \rho \cos \theta \quad y = \rho \sin \theta$$

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 u}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial u}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2 u}{\partial \theta^2} = 0. \quad (11.25)$$

Bij rotatiesymmetrie geldt $u_{\theta\theta} = 0$ en dan wordt de vergelijking

$$\frac{\partial^2 u}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial u}{\partial \rho} = 0. \quad (11.26)$$

Voor het bepalen van de eindige differentieformules nemen we in het (ρ, θ) -vlak een rechthoekig rooster. Dit geeft in het (x, y) -vlak het volgende rooster



Er gelden nu de volgende formules

$$u_{\rho\rho} = \frac{1}{h^2} (u_{r+1,s} - 2u_{r,s} + u_{r-1,s}) + \mathcal{O}(h^2)$$

en

$$\begin{aligned} \frac{1}{\rho} u_{\rho} &= \frac{1}{2\rho h} (u_{r+1,s} - u_{r-1,s}) + \mathcal{O}(h^2) \\ &= \frac{1}{2rh^2} (u_{r+1,s} - u_{r-1,s}) + \mathcal{O}(h^2) \end{aligned}$$

$$\frac{1}{\rho^2} u_{\theta\theta} = \frac{1}{(rh)^2} \frac{1}{k^2} (u_{r,s+1} - 2u_{r,s} + u_{r,s-1}) + \mathcal{O}(k^2).$$

Voor de Laplace vergelijking in poolcoördinaten vinden we dan

$$u_{\rho\rho} + \frac{1}{\rho} u_{\rho} + \frac{1}{\rho^2} u_{\theta\theta} = \frac{1}{h^2} \left[\frac{1}{r^2 k^2} u_{r,s+1} + \frac{1}{r^2 k^2} u_{r,s-1} + \left(1 + \frac{1}{2r}\right) u_{r+1,s} + \right. \\ \left. + \left(1 - \frac{1}{2r}\right) u_{r-1,s} - 2\left(1 + \frac{1}{r^2 k^2}\right) u_{r,s} \right] + \mathcal{O}(h^2) + \mathcal{O}(k^2). \quad (11.27)$$

We zullen deze formule bijvoorbeeld ook gebruiken als de rand eenvoudig is in poolcoördinaten. Bij de rotatie-symmetrie is $u_{r,s}$ onafhankelijk van s en daardoor vereenvoudigt (11.27) tot

$$u_{\rho\rho} + \frac{1}{\rho} u_{\rho} = \frac{1}{h^2} \left[\left(1 + \frac{1}{2r}\right) u_{r+1,s} - 2u_{r,s} + \left(1 - \frac{1}{2r}\right) u_{r-1,s} \right] + \mathcal{O}(h^2) \quad (11.28)$$

Het probleem is alleen zinvol in de buurt van de oorsprong als voor $\rho = 0$ of $r = 0$ geldt $\frac{\partial u}{\partial \rho} = 0$. Is dit het geval dan moeten we bij de differentieformule $u_{1,s} = u_{-1,s}$ nemen; dan is

$$(u_{\rho\rho} + \frac{1}{\rho} u_{\rho})_{\rho=0} = \frac{2}{h^2} (u_{1,s} - u_{0,s}) + \mathcal{O}(h^2) \quad (11.29)$$

hetgeen volgt uit

$$\left(\frac{\partial u}{\partial \rho}\right)_{\rho=0} = (u_{\rho\rho})_{\rho=0} = \frac{1}{h^2} (u_{1,s} - 2u_{0,s} + u_{-1,s}) + \mathcal{O}(h^2).$$

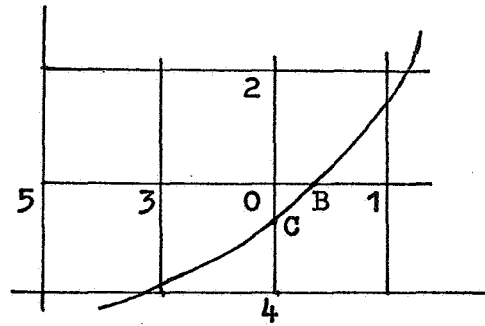
Behandeling van de randen

Tot nu toe hebben we steeds gebieden beschouwd van rechthoekige vorm waarbij de randen evenwijdig liepen aan de coördinaat-assen. Verder veronderstelden we dat h en k zodanig gekozen konden worden dat de randen samenvielen met lijnen van het rooster. Indien h en k verschillend mogen zijn is dit altijd mogelijk, zodat een vijfpuntsformule zoals op blz. NA-140, of een negenpuntsformule voor een vergelijking die de gemengde tweede afgeleide bevat, altijd gebruikt kan worden met een lokale afbreekfout $\mathcal{O}(h^2) + \mathcal{O}(k^2)$ in ieder punt.

Voor de speciale negenpuntsformule (11.23) moeten we veronderstellen dat $h = k$ om de afbreekfout $\mathcal{O}(h^6)$ te krijgen. In dit geval kunnen de randen alleen maar samenvallen met roosterlijnen als de lengten van de randen een simpele verhouding hebben.

Is de verhouding van deze lengten niet geschikt dan zullen we een roosterlijn hebben waarvan de afstand tot de rand kleiner is dan h . We kunnen dan nog wel soortgelijke differentieformules vinden voor de punten dicht bij de rand, maar de lokale afbreekfout wordt van lagere orde in h en k ; bijvoorbeeld van $\mathcal{O}(h^4)$ tot $\mathcal{O}(h)$ voor de speciale negenpuntsformule. In zo'n geval is het niet erg zinvol om zeer nauwkeurige formules te gebruiken in de punten die verder van de rand liggen omdat deze nauwkeurigheid toch niet overal te halen is. Men zou in zo'n geval bijvoorbeeld de eis $h = k$ kunnen laten vallen.

Veel erger wordt het als de randen niet meer recht zijn, maar gekromd zoals bijvoorbeeld in nevenstaande figuur. Als we de vijfpuntsformule gebruiken dan zijn de punten 2 en 5 normale inwendige punten. B en C zijn randpunten. De afstanden OB en OC zijn fracties van de stapgrootten h en k. We stellen daarom



$$OB = \xi_1 h \quad OC = \eta_1 k \quad O3 = \xi_2 h \quad O2 = \eta_2 k \quad (11.30)$$

Hierin zijn $0 < \xi_1 < 1$, $-1 < \eta_1 < 0$, terwijl $\xi_2 = 1$ en $\eta_2 = 1$ ingevoerd zijn om de afleiding van de formule zo algemeen mogelijk te houden.

We beschouwen het geval van een Dirichlet probleem. Dan zijn u_B en u_C gegeven. We bepalen de afgeleiden in het punt 0 met behulp van de Taylor-ontwikkelingen

$$u_B = u_0 + (\xi_1 h) \frac{\partial u_0}{\partial x} + \frac{(\xi_1 h)^2}{2!} \frac{\partial^2 u_0}{\partial x^2} + \frac{(\xi_1 h)^3}{3!} \frac{\partial^3 u_0}{\partial x^3} + \dots$$

$$u_3 = u_0 + (\xi_2 h) \frac{\partial u_0}{\partial x} + \frac{(\xi_2 h)^2}{2!} \frac{\partial^2 u_0}{\partial x^2} + \frac{(\xi_2 h)^3}{3!} \frac{\partial^3 u_0}{\partial x^3} + \dots$$

Om de eerste afgeleide te krijgen elimineren we $\frac{\partial u_0}{\partial x}$, en krijgen dan

$$\frac{\partial u_0}{\partial x} = \frac{1}{h} \left\{ \frac{\xi_2}{\xi_1(\xi_1 + \xi_2)} u_B + \frac{\xi_1 - \xi_2}{\xi_1 \xi_2} u_0 - \frac{\xi_1}{\xi_2(\xi_1 + \xi_2)} u_3 \right\} + \mathcal{O}(h^2). \quad (11.31)$$

Door eliminatie van $\frac{\partial u_0}{\partial x}$ vinden we

$$\frac{\partial^2 u_0}{\partial x^2} = \frac{2}{h^2} \left\{ \frac{1}{\xi_1(\xi_1 + \xi_2)} u_B - \frac{1}{\xi_1 \xi_2} u_0 + \frac{1}{\xi_2(\xi_1 + \xi_2)} u_3 \right\} + \frac{1}{3} h(\xi_2 - \xi_1) \frac{\partial^3 u_0}{\partial x^3} + \mathcal{O}(h^2). \quad (11.32)$$

Soortgelijke formules kunnen we ook afleiden voor de afgeleiden van u naar y.

We merken op dat in de afbreekfout een term $\mathcal{O}(h)$ voorkomt als $\xi_1 \neq \xi_2$.

Het formeren van het systeem differentievergelijkingen wordt een moeizaam karwei. Voor de negenpuntsformule wordt de zaak nog gecompliceerder. Een andere mogelijkheid is de standaard differentieformule alleen te gebruiken in die punten die omgeven zijn door inwendige roosterpunten, en de punten, die dicht bij de rand liggen, te benaderen door interpolatie.

Op deze manier worden de randvoorwaarden verwerkt in de roosterpunten die dicht bij de rand liggen. In bovenstaande figuur, bijvoorbeeld, bevat de vijfpuntsformule in het punt 3 het punt 0 en de waarde in dit punt 0 kunnen we uitdrukken in de waarden in de punten 3 en B door lineaire interpolatie

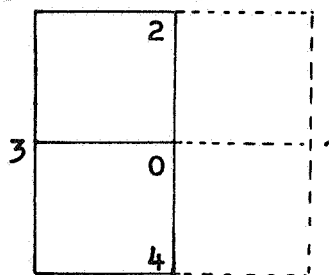
$$u_0 = \frac{\xi_1 u_3 + \xi_2 u_B}{\xi_1 + \xi_2} + \mathcal{O}(h^2) \tag{11.33}$$

of door een kwadratische interpolatie die het punt 5 gebruikt. Een derde mogelijkheid is om ook in het punt 0 aan de differentievergelijking te voldoen. Daarbij maken we gebruik van het uitwendige punt 1. De waarde u_1 kunnen we uitdrukken in u_B en u_0 of in u_B , u_0 en u_3 door lineaire of kwadratische extrapolatie. Deze laatste methode geeft effectief dezelfde formules als die welke we hebben afgeleid uit (11.31) en (11.32)

We beschouwen nu het geval van een Neumann-probleem.

Als op de rand de normale afgeleide (of een combinatie van functie en normale afgeleide) gegeven is, dan is het probleem van de verwerking van de rand nog gecompliceerder. Voor een rechthoekig gebied zijn nog wel betrekkelijk eenvoudige formules af te leiden, doch de afbreekfout zal dan weer groter zijn dan in inwendige van het gebied.

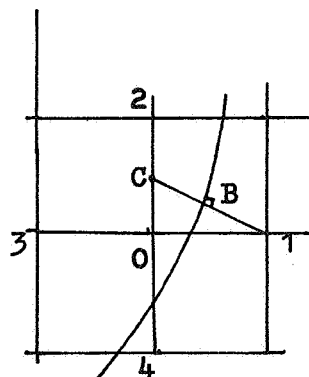
In nevenstaande figuur bijvoorbeeld wordt de rand gevormd door 20^4 , en op deze rand is de normale afgeleide gegeven. In het punt 0 geeft deze normale afgeleide een relatie tussen u_4 en u_3 . De vijfpuntsformule in het punt 0 bevat dus geen uitwendige punten en de afbreekfout is van dezelfde orde, dus $\mathcal{O}(h^2)$. Voor de negenpuntsformule echter reduceert de lokale afbreekfout $\mathcal{O}(h^4)$ in een inwendig punt tot $\mathcal{O}(h^2)$ in een randpunt.



Voor gekromde randen is het probleem veel gecompliceerder, daar de richting van de normaal zelden zal samenvallen met een rechte van het rooster.

Voor dit probleem zijn twee methoden in gebruik. De eerste methode leidt tot eenvoudige, doch onnauwkeurige formules, de tweede methode geeft grotere nauwkeurigheid ten koste van veel gecompliceerdere formules.

De eerste methode introduceert uitwendige punten zoals punt 1 en punt 4 in de vijfpuntsformule voor het punt 0. Deze uitwendige punten worden met behulp van de normale afgeleide op de rand uitgedrukt in inwendige punten en daardoor geelimineerd uit de differentieformule.



Bijvoorbeeld voor punt 1.

We trekken de normaal uit het punt 1 op de randkromme. Het snijpunt noemen we B. Het snijpunt van de normaal met de roosterlijn O2 noemen we C. We benaderen $\frac{\partial u_B}{\partial n}$ door $\frac{u_1 - u_C}{m}$ waarbij m de afstand van 1 tot C is. u_C vinden we door lineaire interpolatie uit u_0 en u_2 . Daarmee hebben we dus u_1 uitgedrukt in $\frac{\partial u_B}{\partial n}$, u_0 en u_2 . De differentieformule in het punt O is natuurlijk zeer onnauwkeurig vergeleken met de formules in punten die verder van de rand liggen.

De tweede methode is meer bewerkelijk. We geven hier alleen de resultaten. We passen in O de vijfpuntsformule toe. In deze formule is alleen het punt 1 op de lijn OA uitwendig. We trekken in A de normaal \underline{n} die een hoek α maakt met de horizontale roosterlijn, en de raakvector \underline{s} . Overal op de randkromme is $\frac{\partial u}{\partial \underline{n}}$ gegeven.

We beschouwen speciaal de afgeleide van $\frac{\partial u}{\partial \underline{n}}$ langs de kromme, dus $\frac{\partial^2 u}{\partial \underline{s} \partial \underline{n}}$.

Nu geldt

$$\frac{\partial}{\partial \underline{n}} = \cos \alpha \frac{\partial}{\partial x} + \sin \alpha \frac{\partial}{\partial y} \quad \text{en} \quad \frac{\partial}{\partial \underline{s}} = -\sin \alpha \frac{\partial}{\partial x} + \cos \alpha \frac{\partial}{\partial y} \quad (11.34)$$

dus

$$\begin{aligned} \frac{\partial^2 u_A}{\partial \underline{s} \partial \underline{n}} &= (\cos^2 \alpha - \sin^2 \alpha) \frac{\partial^2 u_A}{\partial x \partial y} + \cos \alpha \sin \alpha \left(\frac{\partial^2 u_A}{\partial y^2} - \frac{\partial^2 u_A}{\partial x^2} \right) + \\ &+ \left(\sin \alpha \frac{\partial \alpha}{\partial x} - \cos \alpha \frac{\partial \alpha}{\partial y} \right) \left(\sin \alpha \frac{\partial u_A}{\partial x} - \cos \alpha \frac{\partial u_A}{\partial y} \right). \end{aligned} \quad (11.35)$$

In het bijzonder geldt

$$\sin \alpha \frac{\partial \alpha}{\partial x} - \cos \alpha \frac{\partial \alpha}{\partial y} = -\frac{\partial \alpha}{\partial \underline{s}} = -\frac{1}{\rho}$$

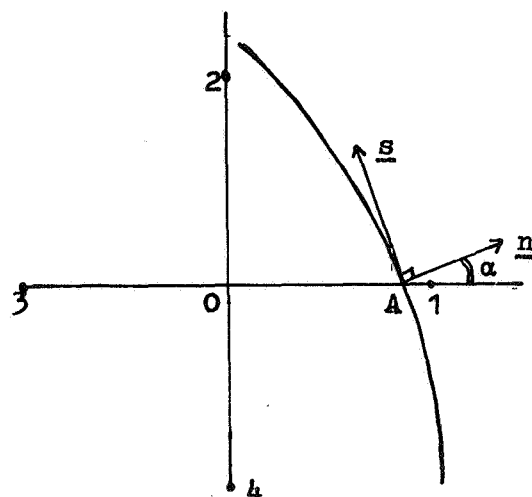
waarbij ρ de kromtestraal is van de rand in het punt A.

De vijf partiele afgeleiden van u naar x en y die in (11.35) voorkomen kunnen we uitdrukken in de afgeleiden in het punt O en de afstand OA = ζh . We doen dit met behulp van Taylor-ontwikkelingen, bijvoorbeeld

$$\frac{\partial u_A}{\partial x} = \frac{\partial u_O}{\partial x} + (\zeta h) \frac{\partial^2 u_O}{\partial x^2} + \frac{(\zeta h)^2}{2!} \frac{\partial^3 u_O}{\partial x^3} + \dots \quad (11.36)$$

Daarmee kunnen we $\frac{\partial u_A}{\partial \underline{n}}$ en $\frac{\partial^2 u_A}{\partial \underline{s} \partial \underline{n}}$ schrijven in termen van de afgeleiden in het punt O.

Drie andere relaties vinden we door u_2 , u_3 , en u_4 uit te drukken in u_0 en de afgeleiden van u in het punt O. De gegeven differentiaalvergelijking



tenslotte geeft ook nog een vergelijking. Aldus verkrijgen we zes vergelijkingen waaruit we de vijf partiele afgeleiden in het punt O kunnen elimineren. Daarmee hebben we een eindige differentieformule voor de differentiaalvergelijking gevonden in de grootheden u_0, u_2, u_3, u_4 en de gegeven randvoorwaarden.

Voor het speciale geval van de vergelijking van Poisson (11.17) wordt deze differentieformule

$$\frac{1}{h^2} (au_2 + bu_3 + cu_4 - 4eu_0) - gf_0 + \frac{1}{h} p \frac{\partial u_A}{\partial \underline{n}} + q \frac{\partial^2 u_A}{\partial \underline{s} \partial \underline{n}} = \mathcal{O}(h) \quad (11.37)$$

waarbij f_0 de waarde van $f(x,y)$ in het punt O is en

$$\begin{aligned} a &= 1 - m + 2\xi \frac{1+m^2}{1-m^2} + \frac{\xi h m \sqrt{1+m^2}}{\rho_A 1-m} \\ b &= 2\left(1 + \frac{\xi h m^2 \sqrt{1+m^2}}{\rho_A 1-m^2}\right) \\ c &= 1 + m + 2\xi \frac{1+m^2}{1-m^2} - \frac{\xi h m \sqrt{1+m^2}}{\rho_A 1+m} \\ e &= 1 + \xi \frac{1+m^2}{1-m^2} + \frac{\xi h m^2 \sqrt{1+m^2}}{\rho_A 1-m^2} \\ g &= 1 + \frac{2\xi}{1-m^2} \quad p = 2\sqrt{1+m^2} \quad q = -2\xi m \frac{1+m^2}{1-m^2} \end{aligned} \quad (11.38)$$

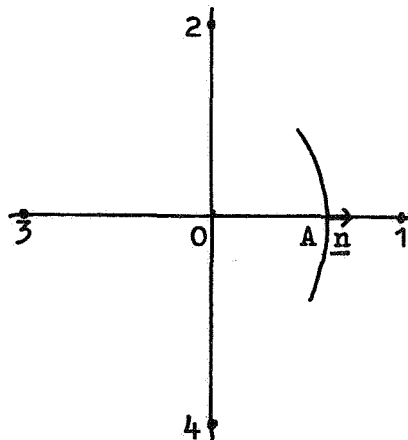
$$m = \tan \alpha .$$

We zien dat de totale afbreekfout van de orde h is, dus van dezelfde orde als bij het Dirichlet-probleem, waarbij we de formules (11.31) en (11.32) gebruiken voor punten dicht bij de rand.

We beschouwen nu het speciale geval $\alpha = 0$, d.w.z. de normaal in A valt samen met een roosterlijn. Dan is $m = 0$ en de formules (11.38) worden eenvoudig

$$\begin{aligned} a &= 1 + 2\xi \\ b &= 2 \\ c &= 1 + 2\xi \\ e &= 1 + \xi \\ g &= 1 + 2\xi \quad p = 2 \quad q = 0. \end{aligned}$$

De differentieformule (11.37) wordt dan



$$\begin{aligned} \frac{1}{h^2} \{ (1+2\xi)u_2 + 2u_3 + (1+2\xi)u_4 - 4(1+\xi)u_0 \} = \\ = (1+2\xi)f_0 - \frac{2}{h} \frac{\partial u_A}{\partial \underline{n}} + \mathcal{O}(h). \end{aligned} \quad (11.39)$$

Veronderstellen we bovendien nog $\xi = 0$, m.a.w. het punt A valt samen met het punt O, dan krijgen we

$$\frac{1}{h^2} (u_2 + 2u_3 + u_4 - 4u_0) = f_0 - \frac{2}{h} \frac{\partial u_0}{\partial \underline{n}} + \mathcal{O}(h). \quad (11.40)$$

Deze formule kunnen we ook krijgen door in de vijfpuntsformule voor het punt O

$$\frac{1}{h^2} (u_2 + u_3 + u_4 + u_1 - 4u_0) = f_0 + \mathcal{O}(h^2) \quad (11.41)$$

te substitueren

$$\frac{u_1 - u_3}{2h} = \frac{\partial u_0}{\partial \underline{n}} + \mathcal{O}(h^2) \quad \text{ofwel} \quad u_1 = 2h \frac{\partial u_0}{\partial \underline{n}} + u_3 + \mathcal{O}(h^3).$$

Ook voor $\xi \neq 0$ kunnen we (11.39) uit de vijfpuntsformule (11.41) afleiden. In dit geval geldt

$$\begin{aligned} \frac{\partial u_A}{\partial \underline{n}} &= \frac{\partial u_A}{\partial x} = \frac{\partial u_0}{\partial x} + \xi h \frac{\partial^2 u_0}{\partial x^2} + \mathcal{O}(h^2) \\ &= \frac{u_1 - u_3}{2h} + \xi h \frac{u_1 - 2u_0 + u_3}{h^2} + \mathcal{O}(h^2) \\ &= \frac{(1+2\xi)u_1 - 4\xi u_0 + (2\xi - 1)u_3}{2h} + \mathcal{O}(h^2) \end{aligned}$$

dus

$$u_1 = \frac{2h}{1+2\xi} \frac{\partial u_A}{\partial \underline{n}} + \frac{4\xi}{1+2\xi} u_0 + \frac{1-2\xi}{1+2\xi} u_3 + \mathcal{O}(h^3). \quad (11.42)$$

Vullen we deze uitdrukking in (11.41) in dan krijgen we (11.39).

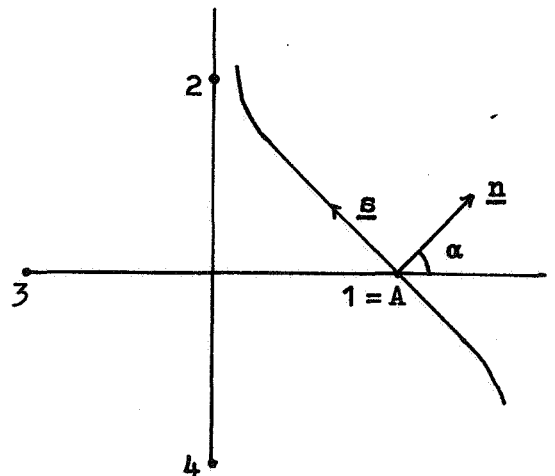
Opmerking. Als $m = 1$, dus als $\alpha = \frac{\pi}{4}$, dan hebben de formules (11.38) geen betekenis. Door in dit geval de gehele vergelijking (11.37), dus alle coëfficiënten (11.38), met $m-1$ te vermenigvuldigen en vervolgens $m = 1$ te stellen krijgen we de differentievergelijking

$$\begin{aligned} \frac{\xi}{h^2} \left[\left(2 + \frac{h}{\rho_A} \sqrt{2} \right) u_2 + \frac{h}{\rho_A} \sqrt{2} u_3 + 2u_4 - 4 \left(1 + \frac{h}{\rho_A} \frac{\sqrt{2}}{2} \right) u_0 \right] = \\ = \xi f_0 + 2\xi \frac{\partial^2 u_A}{\partial \underline{s} \partial \underline{n}} + \mathcal{O}(h). \end{aligned} \quad (11.43)$$

We zullen deze formule aan een speciaal geval controleren. Daartoe nemen we in de buurt van A een rechte rand en $\zeta = 1$, dan gaat de randkromme door het punt 1 en $\frac{1}{\rho_A} = 0$.

Formule (11.43) wordt in dit geval

$$\frac{1}{h^2} (2u_2 + 2u_4 - 4u_0) = f_0 + 2 \frac{\partial^2 u_1}{\partial \underline{s} \partial \underline{n}} + \mathcal{O}(h). \quad (11.44)$$



Ook deze formule kunnen we rechtstreeks uit de vijfpuntsformule (11.41) afleiden, door gebruik te maken van de voorwaartse differentie benadering

$$\frac{\partial^2 u_1}{\partial \underline{s} \partial \underline{n}} = \frac{u_2 - u_3 - u_1 + u_4}{2h^2} + \mathcal{O}(h)$$

dus

$$u_1 = u_2 + u_4 - u_3 - 2h^2 \frac{\partial^2 u_1}{\partial \underline{s} \partial \underline{n}} + \mathcal{O}(h^3).$$

Vullen we deze uitdrukking in (11.41) in dan krijgen we (11.44).

Om de convergentie van een differentiemethode te onderzoeken beschouwen we het verschil van de eindige differentie-oplossing en de exacte oplossing van de differentiaalvergelijking

$$e_{r,s} = u_{r,s} - u(x_r, y_s). \quad (11.45)$$

$e_{r,s}$ noemen we de discretiseringsfout of afbreekfout van de oplossing.

De differentiemethode is convergent als $e_{r,s} \rightarrow 0$ voor $h, k \rightarrow 0$.

We nemen het geval van de algemene lineaire vergelijking (11.5) met als differentieformule de negenpuntsformule

$$\begin{aligned} & \left(\frac{a}{h^2} + \frac{e}{2h} \right) u_{r+1,s} + \left(\frac{a}{h^2} - \frac{e}{2h} \right) u_{r-1,s} + \left(\frac{c}{k^2} + \frac{f}{2k} \right) u_{r,s+1} + \left(\frac{c}{k^2} - \frac{f}{2k} \right) u_{r,s-1} + \\ & + \frac{b}{4hk} (u_{r+1,s+1} - u_{r-1,s+1} + u_{r-1,s-1} - u_{r+1,s-1}) + \\ & - \left(\frac{2a}{h^2} + \frac{2c}{k^2} - g \right) u_{r,s} = L u_{r,s} \end{aligned} \quad (11.46)$$

L heeft de eindige differentie-operator. De lokale afbreekfout van de differentieformule noemen we \mathbb{T} .

Dan luidt de differentie-vergelijking

$$L u_{r,s} = p \quad (11.47)$$

en de differentiaalvergelijking

$$L u(x_r, y_s) = p + T. \quad (11.48)$$

Dus de fout in de oplossing $e_{r,s}$ voldoet aan de vergelijking

$$L e_{r,s} = T. \quad (11.49)$$

Voor een elementaire behandeling van de convergentie moet geëist worden dat de differentie-operator L van positief type of tenminste niet-negatief type is, voor voldoende kleine waarden van h en k . Hiermee bedoelen we dat het teken van de coefficient van $u_{r,s}$ verschillend is van de tekens van alle overige coefficienten in (11.46).

Deze voorwaarde is vervuld als $b = 0$, want a en c hebben hetzelfde teken omdat de vergelijking elliptisch is, en de termen waarin a of c voorkomen zijn voor kleine waarden van h en k overheersend.

Verder moet geëist worden dat de matrix behorende bij de operator L diagonaal-dominant is, d.w.z. de coefficient van $u_{r,s}$ is in absolute

waarde niet kleiner dan de som van de absolute waarden van de overige coefficienten en in minstens één geval zelfs groter. Dit is bijvoorbeeld vervuld als $b = 0$ en het teken van g verschillend is van het teken van a en c .

Als aan deze twee eisen voldaan is dan heeft de oplossing van $Lu = 0$ met u gegeven op de rand de eigenschap dat het maximum en het minimum niet in het inwendige maar op de rand van het gebied liggen; want iedere waarde is een gewogen gemiddelde, met positieve gewichten, van omliggende waarden. Hieruit volgt direct dat als de randwaarden nul zijn, zodat de vergelijkingen werkelijk homogeen zijn, de enige oplossing de nuloplossing is. Dan is de matrix van L niet singulier, dus bestaat de oplossing van het inhomogene stelsel vergelijkingen en deze oplossing is éénduidig.

Deze feiten maken het mogelijk in sommige gevallen convergentiebewijzen te geven. Bijvoorbeeld kan bewezen worden dat de vijfpuntsformule voor de vergelijking van Laplace in een rechthoekig gebied als op de rand de functiewaarden gegeven zijn, een fout geeft die niet groter is dan $O(h^2) + O(k^2)$, mits de vierde partiële afgeleiden van de oplossing begrensd zijn.

Een bewijs hiervan is gegeven door Kantorowitz en Krylow. Voor gebieden met gekromde randen heeft Collatz een bewijs gegeven waarbij voor punten dicht bij de rand benaderingen zoals (11.33) gebruikt worden.

Om een eerste benadering van de oplossing te verbeteren hebben we, evenals bij de gewone differentiaalvergelijkingen, in het algemeen drie methoden.

De eerste methode bestaat eenvoudig uit het verfijnen van het rooster. Deze methode is echter zeer tijdrovend. Halvering van de stapgrootten heeft tot gevolg dat het aantal vergelijkingen vier keer zo groot wordt. Op deze manier wordt het aantal vergelijkingen na enige verfijning onttellend groot.

De tweede methode staat bekend als de deferred-correction methode. We zullen deze methode bespreken aan de hand van de gewone differentiaalvergelijking

$$y'' + f(x)y' + g(x)y = k(x). \quad (11.50)$$

Met behulp van centrale differentieformules, zie blz NA-138, kunnen we (11.50) als volgt schrijven

$$y_{n+1} - 2y_n + y_{n-1} + \frac{1}{2}hf_n(y_{n+1} - y_n - y_{n-1}) + h^2g_n y_n + Cy_n = h^2k_n$$

ofwel

$$(1 + \frac{1}{2}hf_n)y_{n+1} - (2 - h^2g_n)y_n + (1 - \frac{1}{2}hf_n)y_{n-1} + Cy_n = h^2k_n \quad (11.51)$$

waarbij

$$C = (-\frac{1}{12}\delta^4 + \frac{1}{90}\delta^6 \dots) + hf_n(-\frac{1}{6}\mu\delta^3 + \frac{1}{30}\mu\delta^5 \dots). \quad (11.52)$$

Als nu twee beginwaarden y_0 en y_1 gegeven zijn en de term Cy_n verwaarloosd wordt, dan kunnen we y_n in opeenvolgende punten berekenen met bovenstaande recurrente relatie (11.51). Op deze wijze vinden we de eerste benadering $y_n^{(1)}$. Met de differenties van $y_n^{(1)}$ berekenen we de term $Cy_n^{(1)}$, als benadering voor de waarde Cy_n . We vullen $Cy_n^{(1)}$ in (11.51) in, en berekenen dan uitgaande van dezelfde y_0 en y_1 een nieuwe benadering $y_n^{(2)}$. Dit proces kan op dezelfde wijze worden voortgezet tot het verschil in opvolgende benaderingen voor y_n voldoende klein is.

De correctie formule voor de nieuwe benadering kan geschreven worden in de vorm

$$(1 + \frac{1}{2}hf_n)\eta_{n+1} - (2 - h^2g_n)\eta_n + (1 - \frac{1}{2}hf_n)\eta_{n-1} + Cy_n^{(v)} = 0$$

met $\eta_0 = \eta_1 = 0$.

Dan is de nieuwe benadering $y_n^{(v+1)} = y_n^{(v)} + \eta_n$.

De berekening van Cy_n voor kleine waarden van n vereist dat we y_{-1} , y_{-2} , etc kennen. Deze berekenen we door de recurrente betrekking (11.51) in omgekeerde richting toe te passen.

Ook bij de partiële differentiaalvergelijkingen kunnen we de afbreekfout uitdrukken in differenties van de oplossing. Evenals bij de gewone differentiaalvergelijkingen kunnen we met de oplossing van de differentievergelijking benaderingen voor de differenties, en dus voor de afbreekfout berekenen en deze in de vergelijking invullen. Op deze wijze verbeteren we de oplossing op hetzelfde net. Een nadeel van deze methode is dat in ieder punt differenties in twee richtingen bepaald moeten worden, hetgeen tot veel werk aanleiding kan geven.

Een derde methode welke soms met succes toegepast kan worden is de methode van de deferred approach to the limit. (zie ook blz NA-38). Ook deze methode demonstreren we eerst aan de hand van een gewone differentiaalvergelijking

$$y'' = f(x,y). \quad (11.53)$$

Een benaderende oplossing krijgen we uit de eenvoudige eindige differentievergelijking

$$\delta^2 y = h^2 f(x,y). \quad (11.54)$$

We noemen $y(x)$ de oplossing van de differentiaalvergelijking (11.53), en $y(x,h)$ de oplossing van de differentievergelijking (11.54). We veronderstellen dat de afrondingsfouten verwaarloosbaar zijn, dat de oplossing zich fatsoenlijk gedraagt en dat de stapgrootte h voldoende klein is, zodat voor de afbreekfout in de oplossing geldt

$$y(x) - y(x,h) = Ah + Bh^2 + Ch^3 + \dots \quad (11.55)$$

waarbij A, B, C, \dots functies van x zijn.

Indien nu een dergelijke relatie bestaat, dan zijn we in staat de functies A, B, C, \dots te bepalen. Van belang is de eerste functie uit deze rij te bepalen die niet identiek nul is. Daarmee vinden we immers de orde van de fout.

We vullen (11.55) in de vergelijking (11.54) in, en krijgen dan

$$\delta^2(y(x) - Ah - Bh^2 - Ch^3 - \dots) = h^2 f(x, y(x) - Ah - Bh^2 - \dots).$$

Met behulp van de formule

$$\delta^2 = (2 \sinh \frac{1}{2}hD)^2 = h^2 D^2 + \frac{1}{12} h^4 D^4 + \frac{1}{360} h^6 D^6 + \dots$$

vinden we

$$\begin{aligned} (D^2 + \frac{1}{12} h^2 D^4 + \frac{1}{360} h^4 D^6 + \dots)(y(x) - Ah - Bh^2 - Ch^3 - \dots) = \\ = f(x, y(x) - Ah - Bh^2 - Ch^3 - \dots). \end{aligned} \quad (11.56)$$

Deze formule is een identiteit in h en bijgevolg moeten de coëfficiënten van overeenkomstige machten van h gelijk zijn. Dit geeft achtereenvolgens

$$\begin{aligned} h^0 : D^2 y &= f(x, y) \\ h^1 : (D^2 - \frac{\partial f}{\partial y})A &= 0 \\ h^2 : (D^2 - \frac{\partial f}{\partial y})B &= \frac{1}{12} D^4 y - \frac{1}{2} A^2 \frac{\partial^2 f}{\partial y^2} \\ h^3 : (D^2 - \frac{\partial f}{\partial y})C &= -\frac{1}{12} D^4 A - AB \frac{\partial^2 f}{\partial y^2} + \frac{1}{6} A^3 \frac{\partial^3 f}{\partial y^3} \end{aligned} \quad (11.57)$$

waarbij we gebruik maken van de ontwikkeling

$$f(x, y-k) = f(x, y) - k \frac{\partial f}{\partial y} + \frac{1}{2} k^2 \frac{\partial^2 f}{\partial y^2} - \dots$$

De eerste betrekking van (11.57) levert juist de oorspronkelijke differentiaalvergelijking op, terwijl de volgende successievelijk vergelijkingen zijn voor A , B , C , ... etc.

We veronderstellen nu dat we een randwaarde-probleem hebben, met voorgeschreven waarden $y(x)$ aan de randen. Dan is de fout (11.55) aan de randen nul voor iedere waarde van h . Dus alle functies A , B , C , ... zijn nul aan de randen.

Hieruit kunnen we direct concluderen dat A identiek nul is, als éénduidige oplossing van een homogene differentiaalvergelijking (11.57). Dus ook C is identiek nul. De vergelijking voor B is inhomogeen door de term $-\frac{1}{12} D^4 y$, dus B is niet identiek nul. Dit alles geldt natuurlijk onder de voorwaarde dat de vergelijking

$$z'' - \frac{\partial f}{\partial y} z = 0 \quad z(x_0) = z(x_n) = 0$$

alleen de triviale oplossing bezit.

De fout (11.55) kan dus geschreven worden als

$$y(x) - y(x, h) = B(x) \cdot h^2 + E(x) \cdot h^4 + \dots$$

Uit twee benaderende oplossingen voor verschillende stapgrootten h_1 en h_2 bepalen we een nauwkeuriger resultaat door eliminatie van $B(x)$. Er geldt

$$y(x) = y(x, h_1) + B h_1^2 + E h_1^4 + \dots$$

$$y(x) = y(x, h_2) + B h_2^2 + E h_2^4 + \dots$$

dus

$$y(x) = \frac{h_2^2 y(x, h_1) - h_1^2 y(x, h_2)}{h_2^2 - h_1^2} + O(h^4).$$

Hebben we benaderende oplossingen dan kunnen we zowel B als E elimineren. Dit proces staat bekend onder de naam deferred approach to the limit of ook wel h^p -extrapolatie, waarbij p de macht van h aangeeft behorende bij de eerste coëfficiënt, die ongelijk nul is.

We passen dit proces nu toe op het geval van de partiële differentiaalvergelijkingen. Hierbij is het noodzakelijk dat in de uitdrukking voor de fout slechts één parameter voorkomt. We veronderstellen dus dat $k = \lambda h$ en nemen h als parameter.

We veronderstellen dat voor de oplossing u van de differentiaalvergelijking en de oplossing u(h) van de differentievergelijking geldt

$$u = u(h) + Ah + Bh^2 + \dots \quad (11.58)$$

Evenals in het voorgaande willen we de coëfficiënten A, B, ... bepalen.

We veronderstellen dat de elliptische differentiaalvergelijking de vorm heeft

$$f(D) u = g(x, y) \quad (11.59)$$

waarin L de eindige differentie-operator voorstelt.

De eindige differentievergelijking schrijven we in de vorm

$$L u(h) = g(x, y) \quad (11.60)$$

waarbij f(D) een gegeven differentiaal-operator voorstelt.

We veronderstellen dat lokale afbreekfout bekend is, en geschreven kan worden als een machtreeks in h, dus

$$L = f(D) + bh + ch^2 + \dots \quad (11.61)$$

Substitueren we (11.61) en vervolgens u(h) uit (11.58) in (11.60) en stellen we vervolgens de coëfficiënten van overeenkomstige machten van h aan elkaar gelijk, dan krijgen we

$$\begin{aligned} f(D) u &= g(x, y) \\ f(D) A - bu &= 0 \\ f(D) B - cu + bA &= 0 \\ &\text{etc.} \end{aligned} \quad (11.62)$$

Is nu de rand recht, en zijn de functiewaarden gegeven op de rand dan zijn alle coëfficiënten A, B, C, ... nul op de rand. We kunnen in dit geval overal centrale differenties gebruiken, zodat de afbreekfout tenminste $\mathcal{O}(h^2)$ is, m.a.w. $b = 0$. Hieruit kunnen we afleiden dat $A \equiv 0$, maar B niet overal nul.

Op deze wijze kunnen we dus weer komen tot h^2 -extrapolatie.

Opmerking. Voor speciale vergelijkingen en differentie-formules zoals de vergelijking van Laplace en de speciale negenpuntsformule op een rechthoekig gebied, is het mogelijk om tot een h^6 -extrapolatie te komen.

12. Integraalvergelijkingen

Een integraalvergelijking is een vergelijking waarin de onbekende functie onder het integraalteken voorkomt.

We verdelen de integraalvergelijkingen in twee hoofdtypen

- a. Fredholm integraalvergelijkingen. Bij deze vergelijkingen zijn de integratie-grenzen vaste getallen.
- b. Volterra integraalvergelijkingen. In deze vergelijkingen heeft de integraal een vaste grens, de andere integratiegrens is de onafhankelijk variabele.

We beschouwen alleen lineaire integraalvergelijkingen die we onderverdelen in drie soorten

$$\int_a^b k(x,y) f(y) dy = g(x) \quad (12.1)$$

$$\int_a^b k(x,y) f(y) dy = g(x) + f(x) \quad (12.2)$$

$$\int_a^b k(x,y) f(y) dx = \lambda f(x). \quad (12.3)$$

Deze vergelijkingen noemen we respectievelijk de Fredholm vergelijking van de eerste, tweede, derde soort. In ieder van deze vergelijkingen is de kern $k(x,y)$ een bekende functie, $g(x)$ is eveneens bekend. De functie $f(x)$ is de onbekende functie, die we moeten bepalen. De vergelijking van de derde soort (12.3) is een eigenwaarde vergelijking. Hierbij wordt gevraagd naar de waarden van λ (eigenwaarden) waarvoor de vergelijking een niet-triviale oplossing bezit.

De bij een eigenwaarde λ behorende niet-triviale oplossing $f(x)$ noemen we een eigenfunctie.

Als de constante b in de vergelijkingen (12.1), (12.2) en (12.3) vervangen wordt door x dan zijn de vergelijkingen van het Volterra type.

We noemen een Fredholm vergelijking niet-singulier als de integratiegrenzen a en b eindig zijn, de kern $k(x,y)$ in het gebied (a,b) voor x en y geen singulariteit heeft en $g(x)$ voldoende glad is.

We beschouwen nu de numerieke oplossing van een lineaire niet-singuliere Fredholm vergelijking van de tweede soort (12.2).

Volgens de definitie van de Riemann-integraal kunnen we deze vergelijking ook schrijven als

$$\lim_{N \rightarrow \infty} \sum_{i=1}^N (x_i - x_{i-1}) k(x, y_i) f(y_i) = f(x) + g(x) \quad (12.4)$$

waarin $a \leq x_0 < x_1 \dots < x_N \leq b$, y_i een punt van het interval (x_{i-1}, x_i) is, terwijl bij het limietproces $N \rightarrow \infty$ en $\max(x_{i-1}, x_i) \rightarrow 0$. We geven x in (12.4) achtereenvolgens de waarden x_1, x_2, \dots, x_N en krijgen dan als benadering voor de integraalvergelijking het stelsel lineaire vergelijkingen

$$\sum_{i=1}^N (x_i - x_{i-1}) k(x_j, x_i) f(x_i) = f(x_j) + g(x_j) \quad (12.5)$$

$$j = 1, 2, \dots, N$$

voor de bepaling van $f(x_j)$.

Dit stelsel vergelijkingen heeft in matrix-notatie de vorm

$$(K - I) \underline{f} = \underline{g} \quad (12.6)$$

waarbij de matrix K afkomstig is van de kern $k(x, y)$. Als de matrix $K - I$ van (12.6) singulier is, dan heeft het algebraïsche probleem in het algemeen geen oplossing, en in ieder geval geen eenduidig bepaalde oplossing. Overeenkomstig zal de integraalvergelijking (12.2) geen oplossing, of tenminste geen eenduidig bepaalde oplossing hebben, als de homogene vergelijking (12.2) met $g(x) = 0$ een niet-triviale oplossing heeft; of anders gezegd, als $\lambda = 1$ een eigenwaarde is van (12.3).

Verder, als (12.3) een eigenwaarde dicht bij één heeft, dan is het inhomogene probleem ill-conditioned, en we zullen numerieke moeilijkheden hebben bij het bepalen van de oplossing van (12.6) met enige nauwkeurigheid.

In het geval van een Volterra integraalvergelijking wordt de matrix K uit (12.6) een driehoeksmatrix. Het is daarom voor de hand liggend te veronderstellen dat een Volterra probleem gemakkelijker op te lossen is dan een Fredholm probleem.

We leiden nu allereerst voor de trapeziumregel

$$\int_a^b y(x) dx = h \left[\left(\frac{1}{2} y_0 + y_1 + \dots + y_{n-1} + \frac{1}{2} y_n \right) + C \right] \quad (12.7)$$

een uitdrukking af voor de afbreekfout C in voorwaartse differenties van y_0 en achterwaartse differenties van y_n . We doen dit met behulp van de symbolische operatorenmethode. Daartoe definiëren we eerst de operator

$$D^{-1} \text{ integraaloperator } D^{-1} y(x) = \int_a^x y(t) dt.$$

Men kan eenvoudig nagaan dat D^{-1} inderdaad de inverse is van de differentiaaloperator D . Dan geldt

$$\int_{x_r}^{x_{r+1}} y(x) dx = D^{-1} (y_{r+1} - y_r) = D^{-1} \Delta y_r = D^{-1} \nabla y_{r+1}. \quad (12.8)$$

Met behulp van de relaties $\Delta = E - 1$ en $hD = \log E$ vinden we

$$D^{-1} \Delta = h \frac{\Delta}{\log(1+\Delta)} = h \left(1 + \frac{\Delta}{2} - \frac{\Delta^2}{12} + \frac{\Delta^3}{24} - \dots \right).$$

Op analoge wijze vinden we

$$D^{-1} \nabla = h \left(1 - \frac{\nabla}{2} - \frac{\nabla^2}{12} - \frac{\nabla^3}{24} - \dots \right).$$

Dus in termen van voorwaartse differenties geldt

$$\int_{x_r}^{x_{r+1}} y(x) dx = \frac{h}{2} (y_{r+1} + y_r) + h \left(-\frac{\Delta^2}{12} + \frac{\Delta^3}{24} \dots \right) y_r$$

en met achterwaartse differenties

$$\int_{x_r}^{x_{r+1}} y(x) dx = \frac{h}{2} (y_{r+1} + y_r) + h \left(-\frac{\nabla^2}{12} - \frac{\nabla^3}{24} \dots \right) y_{r+1}.$$

Noemen we C_r de afbreekfout in het r -de interval bij de trapeziumregel dan is volgens bovenstaande formules

$$C_r = \left(-\frac{\Delta^2}{12} + \frac{\Delta^3}{24} - \dots \right) y_r \quad (12.9)$$

en

$$C_r = \left(-\frac{\nabla^2}{12} - \frac{\nabla^3}{24} - \dots \right) y_{r+1}. \quad (12.10)$$

Nu geldt

$$\begin{aligned} \int_{x_0}^{x_n} y(x) dx &= \sum_{r=0}^{n-1} \int_{x_r}^{x_{r+1}} y(x) dx = \sum_{r=0}^{n-1} \left[\frac{h}{2} (y_{r+1} + y_r) + h C_r \right] \\ &= h \left(\frac{1}{2} y_0 + y_1 + \dots + y_{n-1} + \frac{1}{2} y_n \right) + h \sum_{r=0}^{n-1} C_r \end{aligned}$$

en daar $C_r = E^r C_0$ geldt dus voor de afbreekfout C uit (12.7)

$$C = \sum_{r=0}^{n-1} C_r = \sum_{r=0}^{n-1} E^r C_0 = \frac{1-E^n}{1-E} \cdot C_0 = \frac{1}{1-E} C_0 - \frac{E^n}{1-E} C_0. \quad (12.11)$$

In de eerste term van (12.11) vullen we de uitdrukking (12.9) in, dus

$$\frac{1}{1-E} C_0 = \frac{1}{1-E} \left(-\frac{\Delta^2}{12} + \frac{\Delta^3}{24} - \dots \right) y_0 = \left(\frac{\Delta}{12} - \frac{\Delta^2}{24} - \dots \right) y_0.$$

In de tweede term van (12.11) gebruiken we (12.10) en krijgen dan

$$\frac{E^n}{1-E} C_0 = \frac{E^{n-1}}{E^{-1}-1} \left(-\frac{\nabla}{12} - \frac{\nabla^2}{24} - \dots \right) y_1 = \left(\frac{\nabla}{12} + \frac{\nabla^2}{24} - \dots \right) y_n$$

zodat we tenslotte voor C de formule vinden

$$C = \left(\frac{\Delta}{12} - \frac{\Delta^2}{24} + \dots \right) y_0 - \left(\frac{\nabla}{12} + \frac{\nabla^2}{24} + \dots \right) y_n \quad (12.12)$$

(Gregory)

We beschouwen nu de numerieke oplossingsmethode van het tweede Fredholm probleem, die we krijgen als we de integraal in (12.2) met behulp van de trapeziumregel berekenen. De Fredholm vergelijking (12.2) wordt dan

$$h\left[\frac{1}{2}k(x, x_0)f_0 + k(x, x_1)f_1 + \dots + k(x, x_{n-1})f_{n-1} + \frac{1}{2}k(x, x_n)f_n + C(x)\right] = g(x) + f(x) \quad (12.13)$$

waarin de correctieterm C natuurlijk een functie van x is. Deze vergelijking geldt voor iedere x waarvoor $k(x, y)$ gedefinieerd is. Nemen we nu speciaal de waarden $x = x_0, x_1, \dots, x_n$ in (12.13) dan krijgen we het stelsel lineaire vergelijkingen

$$h\left(\frac{1}{2}k_{i,0}f_0 + k_{i,1}f_1 + \dots + k_{i,n-1}f_{n-1} + \frac{1}{2}k_{i,n}f_n + C_i\right) = g_i + f_i \quad i=0,1,\dots,n \quad (12.14)$$

waarin $k_{i,j} = k(x_i, x_j)$ en $g_i = g(x_i)$ bekend zijn, en $f_i = f(x_i)$ bepaald moeten worden. Of in matrix notatie

$$\underline{A} \underline{f} = -\underline{g} + h \underline{C}(f)$$

waarbij de matrix A van de vorm is

$$\underline{A} = \begin{pmatrix} 1 - \frac{1}{2}h k_{0,0} & -h k_{0,1} & \dots & -\frac{1}{2}h k_{0,n} \\ -\frac{1}{2}h k_{1,0} & 1 - \frac{1}{2}h k_{1,1} & & \cdot \\ & & & \cdot \\ -\frac{1}{2}h k_{n,0} & \dots & \dots & 1 - \frac{1}{2}h k_{n,n} \end{pmatrix} \quad (12.15)$$

De correctieterm C_i kunnen we op twee verschillende manieren in rekening brengen.

Als we weten hoeveel differenties van (12.12) een significante bijdrage geven kunnen we deze differenties uitdrukken in functiewaarden. Daarbij blijft het stelsel vergelijkingen (12.14) lineair. De moeilijkheid hierbij is echter dat de differenties mede bepaald worden door de onbekende functie $f(x)$. In de meeste gevallen zal het daarom niet mogelijk zijn te bepalen welke differenties nog significant zijn. Het is daarom aantrekkelijker om een iteratieve methode zoals de "deferred correction" methode te gebruiken. Voor de eerste benadering van de oplossing $f(x)$ negeren we de correctieterm en krijgen dan het stelsel vergelijkingen

$$\underline{A} \underline{f}^{(0)} = -\underline{g}$$

Met de oplossing $f^{(0)}$ bepalen we de correctieterm $\underline{C}(f^{(0)})$ en vinden vervolgens als tweede benadering $f^{(1)}$ uit

$$A \underline{f}^{(1)} = -g + h \underline{C}(f^{(0)}).$$

Dit proces herhalen we op de gebruikelijke manier. Daarbij kunnen we voor het bepalen van de correctieterm C gebruik maken van de formule (12.12), of een centrale differentieformule in de eindpunten

$$C = \left(-\frac{1}{12} \mu \delta + \frac{11}{720} \mu \delta^3 - \dots\right) (y_n - y_0). \quad (12.16)$$

Deze heeft het voordeel dat de coëfficiënten sneller afnemen. Om (12.16) te gebruiken moeten we y_r buiten het interval (a,b) kennen, en hiervoor is nodig dat we $k(x,y)$ en $g(x)$ buiten (a,b) definiëren. Want dan kunnen we met behulp van de vergelijking (12.13) $f(x)$ in het buitengebied berekenen.

De deferred correction methode is nogal bewerkelijk in dit geval omdat voor iedere vergelijking (12.14) een differentietabel gemaakt moet worden van $y_j = k_{i,j} f_j$.

Ook kunnen we de methode van de "deferred approach to the limit" gebruiken, met h^2 -extrapolatie in het geval van de trapeziumregel en h^4 -extrapolatie bij het gebruik van de integratieformule van Simpson.

De eindige differentieformules voor integratie geven bij $n+1$ aequidistante basispunten alleen exacte resultaten bij polynomen van de graad $\leq n$ zelfs als we de differentiecorrectie meenemen.

Als we de Gauss integratieformule gebruiken krijgen we exacte resultaten voor polynomen tot en met de graad $2n+1$, alleen de basispunten zijn niet meer aequidistant gelegen. De integraalvergelijking wordt in dit geval

$$\sum_{r=1}^p w_r k(x, x_r) f(x_r) + C(x) = g(x) + f(x). \quad (12.17)$$

Deze formule geldt voor alle x . Kiezen we $x = x_1, x_2, \dots, x_p$ dan krijgen we een stelsel van p lineaire vergelijkingen, waaruit we de waarden $f(x_r)$ kunnen bepalen.

Een nadeel van deze methode is de moeilijkheid om de nauwkeurigheid te schatten zonder de gehele berekening te herhalen met een formule van hogere orde.

Opmerking. Als we de oplossing $f(x_r)$ gevonden hebben in de basispunten x_r , dan kunnen we $f(x)$ in andere punten eenvoudig vinden uit de vergelijking (12.13) of (12.17). Dit gaat in het algemeen sneller en nauwkeuriger dan interpolatie met de gevonden basispunten.

We beschouwen nu enkele bijzondere gevallen.
We noemen de kern $k(x,y)$ symmetrisch als geldt

$$k(x,y) = k(y,x). \quad (12.18)$$

Dan is de matrix A uit (12.15) eveneens symmetrisch als we de eerste en de laatste rij delen door 2. Het oplossen van het stelsel vergelijkingen (12.14) is eenvoudiger. Bij gebruik van een eliminatiemethode is de hoeveelheid werk half zo groot als bij een willekeurig stelsel vergelijkingen van dezelfde afmetingen.

Een verdere vereenvoudiging krijgen we als de symmetrische kern van het convolutie-type is, d.w.z. de kern is een functie van $x - y$ alleen

$$k(x,y) = k(x - y) \quad (12.19)$$

In dit geval wordt de matrix A behalve symmetrisch, ook centro-symmetrisch; dit betekent

$$a_{ij} = a_{n-i,n-j} = a_{ji} = a_{n-j,n-i}$$

We kunnen dan het stelsel van $n+1$ vergelijkingen splitsen in twee stelsels beide van de orde $\frac{1}{2}(n+1)$ als n oneven is, of van de orde $\frac{1}{2}(n+2)$ en $\frac{1}{2}n$ als n even is.

Voorbeeld

$$\begin{aligned} af_0 + bf_1 + cf_2 + df_3 &= g_0 \\ bf_0 + ef_1 + pf_2 + cf_3 &= g_1 \\ cf_0 + pf_1 + ef_2 + bf_3 &= g_2 \\ df_0 + cf_1 + bf_2 + af_3 &= g_3 \end{aligned}$$

Door vergelijking vier bij een, en vergelijking drie bij twee op te tellen krijgen we de vergelijkingen

$$\begin{aligned} (a+d)(f_0 + f_3) + (b+c)(f_1 + f_2) &= g_0 + g_3 \\ (b+c)(f_0 + f_3) + (e+p)(f_1 + f_2) &= g_1 + g_2 \end{aligned}$$

Door aftrekking van dezelfde vergelijkingen krijgen we

$$\begin{aligned} (a-d)(f_0 - f_3) + (b-c)(f_1 - f_2) &= g_0 - g_3 \\ (b-c)(f_0 - f_3) + (e-p)(f_1 - f_2) &= g_1 - g_2 \end{aligned}$$

In plaats van één stelsel van vier vergelijkingen hebben we dus twee stelsels van elk twee vergelijkingen gekregen met als onbekenden $f_0 + f_3$, $f_1 + f_2$ resp. $f_0 - f_3$, $f_1 - f_2$. We bepalen dus eerst de oplossing van de twee kleine stelsels vergelijkingen, en met behulp hiervan vinden we de oplossing van het grote stelsel.

In sommige speciale gevallen zijn we in staat de integraalvergelijking exact op te lossen. Zulke oplossingen kunnen we ook gebruiken als we de integraalvergelijking kunnen benaderen door een vergelijking van speciale gedaante.

Als de kern het product is van een functie van x , en een functie van y

$$k(x,y) = X(x) Y(y) \quad (12.20)$$

dan spreken we van een ontaarde kern. De integraalvergelijking wordt dan

$$\int_a^b X(x)Y(y) f(y) dy = f(x) + g(x). \quad (12.21)$$

We schrijven

$$A = \int_a^b Y(y)f(y) dy \quad (12.22)$$

en vinden dan een expliciete uitdrukking voor de oplossing

$$f(x) = A.X(x) - g(x) \quad (12.23)$$

De constante A vinden we door substitutie van (12.23) in (12.22)

$$A = \int_a^b (AX(y) - g(y)) Y(y) dy$$

zodat

$$A = \frac{-\int_a^b g(y)Y(y) dy}{1 - \int_a^b X(y)Y(y) dy} \quad (12.24)$$

We zullen in het algemeen een kern $k(x,y)$ ontaard noemen (en wel p-voudig ontaard) als

$$k(x,y) = \sum_{r=1}^p X_r(x) Y_r(y). \quad (12.25)$$

We voeren nu in de grootheden

$$A_r = \int_a^b Y_r(y) f(y) dy \quad (12.26)$$

dan is

$$f(x) = \sum_{r=1}^p A_r X_r(x) - g(x) \quad (12.27)$$

Substitutie van (12.27) in (12.26) geeft een stelsel lineaire vergelijkingen voor A_r

$$A_r = \sum_{s=1}^p A_s \int_a^b X_s(y) Y_r(y) dy - \int_a^b g(y) Y_r(y) dy \quad (12.28)$$

$$r = 1, 2, \dots, p$$

Uit dit stelsel vergelijkingen bepalen we A_r $r = 1, 2, \dots, p$, substitueren deze waarden in (12.27) en krijgen dan een oplossing van de integraalvergelijking.

Is het stelsel (12.28) niet oplosbaar dan heeft ook de integraalvergelijking geen oplossing.

In praktische gevallen zal de kern vaak niet de simpele vorm (12.25) hebben, maar kan wel benaderd worden door een ontaarde kern. Bijvoorbeeld als de kern van het convolutie type is kunnen we als benadering voor $k(x-y)$ de eerste $n+1$ termen van de machtreeks nemen

$$k(x-y) \approx \sum_{r=0}^n a_r (x-y)^r = \sum b_{ij} x^i y^j.$$

In andere gevallen kunnen we soms gebruik maken van een dubbele Fourierreeks.

Een iteratieve methode om de Fredholm vergelijking van de tweede soort (12.2) op te lossen is de methode van de successieve approximaties.
Stel

$$\begin{aligned} f_1(x) &= -g(x) \\ f_{n+1}(x) &= -g(x) + \int_a^b k(x,y) f_n(y) dy \end{aligned} \quad (12.29)$$

Dan is de oplossing van de Fredholm vergelijking

$$f(x) = \lim_{n \rightarrow \infty} f_n(x)$$

indien deze limiet bestaat.

We kunnen ook gebruik maken van de formules

$$\begin{aligned} \varphi_0(x) &= -g(x) \\ \varphi_{n+1}(x) &= \int_a^b k(x,y) \varphi_n(y) dy \end{aligned} \quad (12.30)$$

en de oplossing wordt gegeven door de Neumann reeks

$$f(x) = \varphi_0(x) + \varphi_1(x) + \varphi_2(x) + \dots \quad (12.31)$$

Om de convergentie van deze methode te kunnen nagaan beschouwen we

$$\begin{aligned} \epsilon_n(x) &= f(x) - f_n(x) = \int_a^b k(x,y) [f(y) - f_{n-1}(y)] dy \\ &= \int_a^b k(x,y) \epsilon_{n-1}(y) dy. \end{aligned}$$

Noemen we

$$N_{n+1} = \max_{a \leq x \leq b} |\epsilon_{n+1}(x)|$$

dan is

$$\begin{aligned} N_{n+1} &= |\epsilon_{n+1}(\tilde{x})| = \left| \int_a^b K(\tilde{x},y) \epsilon_n(y) dy \right| \\ &\leq M \cdot (b-a) N_n \end{aligned}$$

als we veronderstellen dat $|k(x,y)| \leq M$.

Dus als $M(b-a) < 1$, dan is de methode convergent. Deze voorwaarde is wel voldoende, doch niet noodzakelijk, en zal niet altijd vervuld zijn.

Men kan ook bewijzen dat de methode van de successieve approximaties convergent is als alle eigenwaarden van de kern in absolute kleiner zijn dan één.

We beschouwen nog eens de formules (12.30)

$$\begin{aligned}\varphi_0(x) &= -g(x) \\ \varphi_1(x) &= -\int_a^b k(x,y) g(y) dy \\ \varphi_2(x) &= \int_a^b k(x,y) \varphi_1(y) dy = -\int_a^b k_2(x,y) g(y) dy \\ \varphi_3(x) &= \int_a^b k(x,y) \varphi_2(y) dy = -\int_a^b k_3(x,y) g(y) dy \\ \text{etc.}\end{aligned}$$

waarin

$$\begin{aligned}k_2(x,y) &= \int_a^b k(x,t) k(t,y) dt \\ k_3(x,y) &= \int_a^b k(x,t) k_2(t,y) dt \\ \text{etc.}\end{aligned}$$

meer algemeen

$$\begin{aligned}k_r(x,y) &= \int_a^b k(x,t) k_{r-1}(t,y) dt \\ k_1(x,y) &= k(x,y)\end{aligned}\tag{12.32}$$

De functie $k_r(x,y)$ noemen we de r-de geïtereerde kern.

Als de methode van de successieve approximaties convergent is dan convergeert de reeks (12.31). Invullen van $\varphi_n(x)$ in de reeks geeft dan

$$f(x) = -g(x) + \int_a^b k^*(x,y) g(y) dy\tag{12.33}$$

waarin

$$k^*(x,y) = -\sum_{r=1}^{\infty} k_r(x,y)$$

de reciproke of resolvente kern genoemd wordt.

We beschouwen nu de Fredholm vergelijking van de eerste soort

$$\int_a^b k(x,y) f(y) dy = g(x)\tag{12.1}$$

Dit probleem is moeilijker dan het probleem van de tweede soort. De moeilijkheden komen tevoorschijn als we eindige differentiemethoden, zoals in het voorgaande behandeld, willen gebruiken.

In veel gevallen heeft het probleem van de eerste soort bij gegeven kern $k(x,y)$ alleen een oplossing bij zeer speciale keuze van de functie $g(x)$. Dit is bijvoorbeeld het geval als de kern ontaard is.

Voorbeeld

Neem

$$k(x,y) = X(x)Y(y)$$

dan wordt (12.1)

$$\int_a^b X(x) Y(y) f(y) dy = g(x)$$

en deze vergelijking heeft alleen een oplossing als

$$g(x) = \alpha \cdot X(x) \quad (12.34)$$

In dit geval is de oplossing niet eenduidig want de enige eis voor de oplossing $f(x)$ is

$$\int_a^b Y(y) f(y) dy = \alpha \quad (12.35)$$

Bij een gevonden oplossing $f(x)$ kunnen we dus iedere functie $f_0(x)$ optellen waarvoor

$$\int_a^b Y(y) f_0(y) dy = 0. \quad (12.36)$$

Voor het algemene geval van de ontaarde kern

$$k(x,y) = \sum_{r=1}^p X_r(x) Y_r(y)$$

heeft het probleem alleen een oplossing als

$$g(x) = \sum_{r=1}^p \alpha_r X_r(x)$$

en ook in dit geval is er van eenduidigheid van de oplossing geen sprake. Iedere functie $f_0(x)$ waarvoor geldt

$$\int_a^b Y_r(y) f_0(y) dy = 0, \quad r = 1, 2, \dots, p \quad (12.36)$$

mag bij een oplossing opgeteld worden.

Om in het algemeen de mogelijkheid van oplossen te onderzoeken maken we gebruik van de eigenfuncties van de kern $k(x,y)$. In sommige gevallen is het namelijk mogelijk de oplossing van (12.1) uit te drukken met behulp van de eigenfuncties.We veronderstellen dat de kern $k(x,y)$ de eigenwaarden λ_r met bijbehorende eigenfuncties $f_r(x)$ heeft, zodat dus

$$\int_a^b k(x,y) f_r(y) dy = \lambda_r f_r(x) \quad (12.37)$$

Als het mogelijk is om de functie $g(x)$ te schrijven als

$$g(x) = \sum_{r=1}^{\infty} b_r f_r(x) \quad (12.38)$$

en we nemen aan dat

$$f(x) = \sum_{r=1}^{\infty} a_r f_r(x) \quad (12.39)$$

dan geldt

$$a_r = \frac{b_r}{\lambda_r}$$

hetgeen direct volgt als we (12.38) en (12.39) in (12.1) invullen en gebruik maken van (12.37). Daarmee hebben we de oplossing dan gevonden. Het vinden van een oplossing op deze manier is alleen mogelijk als aan verschillende voorwaarden voldaan is. Bijvoorbeeld moet $g(x)$ te schrijven zijn als een reeks van eigenfuncties (12.38). Dit is alleen mogelijk voor willekeurige $g(x)$ als de eigenfuncties een volledig stelsel vormen, en hiervoor is noodzakelijk dat er oneindig veel onderling verschillende eigenwaarden zijn.

(Een voorbeeld van een volledig stelsel functies is het stelsel van de goniometrische functies $\sin kx$, $\cos kx$, $k = 0, 1, \dots$, . Iedere continue functie is te schrijven als een Fourierreeks). Verder mag $\lambda_0 = 0$ niet voorkomen want dan bestaat $a_0 = \frac{b_0}{\lambda_0}$ niet, tenzij $b_0 = 0$. Anders gezegd, als $\lambda_0 = 0$ dan mag $g(x)$ niet een willekeurige functie zijn want b_0 moet nul zijn. In dat geval is de oplossing niet eenduidig want met $f(x)$ voldoet ook $f(x) + \mu f_0(x)$ aan de vergelijking.

In het geval van een p -voudig ontaarde kern (12.25) wordt de eigenwaardevergelijking

$$\sum_{r=1}^p X_r(x) \int_a^b Y_r(y) f(y) dy = \lambda f(x) \quad (12.40)$$

ofwel

$$\sum_{r=1}^p A_r X_r(x) = \lambda f(x) \quad (12.41)$$

Vullen we (12.41) in (12.40) in dan krijgen we

$$\sum_{r=1}^p X_r(x) \sum_{s=1}^p A_s \int_a^b Y_r(y) X_s(y) dy = \lambda \sum_{r=1}^p X_r(x) \cdot A_r$$

dus

$$\sum_{s=1}^p C_{rs} A_s = \lambda A_r \quad r = 1, 2, \dots, p \quad (12.42)$$

met

$$C_{rs} = \int_a^b Y_r(y) X_s(y) dy.$$

Hieruit zien we dat bij een p-voudig ontaarde kern hoogstens p verschillende eigenwaarden ongelijk nul behoren, nl. de eigenwaarden van (12.42). De bijbehorende eigenfuncties zijn een lineaire combinatie van $X_1(x)$...
... $X_p(x)$ volgens (12.41). Bij een eigenwaarde $\lambda = 0$ behoort als eigenfunctie iedere functie waarvoor

$$A_r = \int_a^b Y_r(y) f(y) dy = 0.$$

We kunnen nu gemakkelijk nagaan dat het voorbeeld op blz. NA-167 een speciaal geval is van de oplossingsmethode met behulp van eigenfuncties.

We zullen nu de numerieke oplossing van het probleem van de eerste soort behandelen. Daarbij moeten we verschijnselen tegenkomen welke corresponderen met de gevonden theoretische resultaten. We nemen het geval van de eindige differentiemethode waarbij de integraal van (12.1) benaderd wordt door de trapeziumformule. Dan krijgen we een stelsel lineaire vergelijkingen, dat in matrix-notatie van de gedaante is

$$A\underline{f} = h^{-1}\underline{g} \quad (12.43)$$

waarin \underline{f} en \underline{g} de vectoren zijn van functiewaarden van $f(x)$ en $g(x)$ op de basispunten en

$$A = KD \quad (12.44)$$

met

$$K = \begin{pmatrix} k_{0,0} & \dots & k_{0,n} \\ k_{1,0} & \dots & k_{1,n} \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ k_{n,0} & \dots & k_{n,n} \end{pmatrix} \quad D = \begin{pmatrix} \frac{1}{2} & & & \\ & 1 & \ominus & \\ & & \cdot & \\ & & & \cdot & \\ \ominus & & & & 1 & \\ & & & & & \frac{1}{2} \end{pmatrix}$$

Als A^{-1} bestaat dan is de oplossing van het numerieke probleem

$$\underline{f} = A^{-1}(h^{-1}\underline{g}) \quad (12.45)$$

en de vorm van de oplossing van (12.1) hangt af van het gedrag van (12.45) als $h \rightarrow 0$.

Als de kern een eigenwaarde $\lambda = 0$ heeft dan zal de matrix K ook een eigenwaarde nul (ev. bijna nul) hebben. De matrix K is dan singulier (ev. bijna singulier). Hetzelfde geldt voor de matrix A , en dus heeft in dit geval (12.43) slechts een oplossing voor zeer speciale keuze van \underline{g} .

Als de kern p-voudig ontaard is dan is er een eigenwaarde $\lambda = 0$. De matrix K is singulier als de orde groter is dan p . Want als we de basispunten x_1, x_2, \dots, x_n nemen en met \underline{X}_s en \underline{Y}_s , $s = 1, \dots, n$ bedoelen de

vectoren

$$\underline{X}_S = \begin{pmatrix} X_1(x_S) \\ \vdots \\ X_p(x_S) \end{pmatrix} \quad \underline{Y}_S = \begin{pmatrix} Y_1(y_S) \\ \vdots \\ Y_p(y_S) \end{pmatrix} \quad (12.46)$$

dan is

$$K = \begin{pmatrix} X_1^T Y_1 & \dots & X_1^T Y_n \\ X_2^T Y_1 & \dots & X_2^T Y_n \\ \vdots & & \vdots \\ X_n^T Y_1 & \dots & X_n^T Y_n \end{pmatrix}$$

ofwel

$$K = \begin{pmatrix} X_1^T \\ X_2^T \\ \vdots \\ X_n^T \end{pmatrix} (Y_1, Y_2, \dots, Y_n) \quad (12.47)$$

De X matrix van (12.47) is een $n \times p$ matrix, de Y matrix is een $p \times n$ matrix. De matrix K heeft de orde n en de rang p als $n > p$.

Voor een ontaarde kern is het daarom niet mogelijk het stelsel vergelijkingen (12.43) eenduidig op te lossen als we meer dan p basispunten kiezen. Dit is niet zo verrassend want we weten dat de integraalvergelijking geen eenduidige oplossing heeft.

De vraag is of het in dit geval toch mogelijk is een oplossing te vinden op meer dan p basispunten. We hebben in het voorgaande gezien dat een noodzakelijke voorwaarde is dat $g(x)$ van de vorm is

$$g(x) = \sum_{r=1}^p \alpha_r X_r(x).$$

De vector g heeft dan de componenten $g_i = (X_i^T, \underline{\alpha})$ waarbij $\underline{\alpha}$ de vector van de coëfficiënten α_r is. De vergelijking (12.43) wordt dan

$$X^T Y D \underline{f} = h^{-1} X^T \underline{\alpha} \quad (12.48)$$

waarbij X^T en Y de matrices van K uit (12.47) zijn. Oplossingen van deze vergelijking zijn de vectoren \underline{f} waarvoor

$$Y D \underline{f} = h^{-1} \underline{\alpha} \quad (12.49)$$

en daar Y een $p \times n$ matrix is kunnen we $n - p$ componenten van \underline{f} willekeurig nemen.

Speciaal kunnen we deze componenten zo kiezen dat we de additionele functies in de oplossing, d.w.z. de eigenfuncties behorende bij $\lambda = 0$, onderdrukken.

We nemen als voorbeeld de integraalvergelijking

$$\int_0^1 (x+y) f(y) dy = g(x). \quad (12.50)$$

Hierin is

$$k(x,y) = \sum_{r=1}^2 X_r(x) Y_r(y)$$

met $X_1(x) = x$, $X_2(x) = Y_1(y) = 1$, $Y_2(y) = y$.

De eigenwaarden met bijbehorende eigenfuncties zijn

$$\lambda_{1,2} = \frac{1}{2} \pm \frac{1}{\sqrt{3}} \quad f_{1,2}(x) = x\sqrt{3} \pm 1 \quad (12.51)$$

en bij de eigenwaarde $\lambda = 0$ behoren als eigenfuncties alle functies $\varphi(x)$ waarvoor geldt

$$\int_0^1 \varphi(y) dy = 0 \quad \int_0^1 y \varphi(y) dy = 0. \quad (12.52)$$

De integraalvergelijking (12.50) is slechts oplosbaar als $g(x)$ een lineaire functie is. De oplossingen zijn dan van de vorm

$$f(x) = A_1 f_1(x) + A_2 f_2(x) + \mu \varphi(x) \quad (12.53)$$

De constanten A_1 en A_2 zijn te berekenen maar μ en ook $\varphi(x)$ is niet bepaald.

Willen we nu de oplossing hebben welke de term $\varphi(x)$ uitsluit dan moeten we meer dan twee basispunten nemen. Bijvoorbeeld is Simpson over het hele interval al voldoende. De trapeziumregel met drie basispunten geeft het stelsel vergelijkingen

$$\begin{aligned} \frac{1}{2} f\left(\frac{1}{2}\right) + \frac{1}{2} f(1) &= 2 g(0) \\ \frac{1}{4} f(0) + f\left(\frac{1}{2}\right) + \frac{3}{4} f(1) &= 2 g\left(\frac{1}{2}\right) \\ \frac{1}{2} f(0) + \frac{3}{2} f\left(\frac{1}{2}\right) + f(1) &= 2 g(1) \end{aligned} \quad (12.54)$$

De matrix is singulier maar als $g(x)$ lineair is dan heeft het stelsel (12.54) de algemene oplossing

$$\begin{aligned} f(0) &= z \\ f\left(\frac{1}{2}\right) &= -12 g(0) + 8 g\left(\frac{1}{2}\right) - z \\ f(1) &= 16 g(0) - 8 g\left(\frac{1}{2}\right) + z \end{aligned}$$

met z als parameter. Bij iedere gegeven functie $g(x)$ kunnen we dan z zo kiezen dat de oplossing een lineaire functie is. Daarmee elimineren we dan de term $\varphi(x)$ uit (12.53).

Dit proces wordt zeer bewerkelijk en moeilijk in het algemene geval, met veel punten, willekeurige constanten en afgeronde coëfficiënten in de vergelijkingen. Verder heeft deze manier voor het geval van een ontaarde kern

geen praktische waarde, want zoals we gezien hebben is dit geval analytisch oplosbaar. We hebben hiermee evenwel aangetoond dat het mogelijk is met behulp van een eindige differentiemethode een oplossing te vinden.

Het geval van een niet ontaarde kern is veel moeilijker. We verwachten hierbij oneindig veel eigenwaarden die zich naar de oorsprong toe verdichten. Bij toename van het aantal basispunten worden meer eigenfuncties toegevoegd behorende bij kleine eigenwaarden en de matrix wordt meer ill-conditioned naarmate het aantal basispunten toeneemt. Het stelsel vergelijkingen (12.43) dat we oplossen is slechts een benadering van het werkelijke stelsel

$$(A + C)\underline{f} = h^{-1}\underline{g} \quad (12.55)$$

waarbij $C\underline{f}$ de differentiecorrectie is. Weliswaar wordt $C\underline{f}$ kleiner als het aantal basispunten toeneemt maar de matrix A wordt tenslotte zo ill-conditioned, dat de kleine residuen aanleiding geven tot grote veranderingen in de componenten van \underline{f} .

We kunnen niet verwachten dat de methode met de eigenfuncties van blz. NA-168 in dit geval significante resultaten oplevert. Want we kunnen slechts eindig veel eigenfuncties bepalen en kleine fouten bij het bepalen van f_r behorende bij een kleine λ_r hebben een groot effect op

$$a_r = \frac{b_r}{\lambda_r} \text{ en dus op de oplossing.}$$

Het probleem is essentieel ill-conditioned, inherent instabiel in de zin van paragraaf 5 en het is de vraag of het probleem in fysische zin wel goed gesteld (well posed) is. Het zoeken naar methoden om het probleem zeer nauwkeurig op te lossen is dus onvruchtbaar. Men zal echter dikwijls al tevreden zijn als men een redelijk gladde functie $f(x)$ kan vinden waarvoor de residue functie

$$e(x) = \int_a^b k(x,y) f(y) dy - g(x)$$

in een of andere zin klein is in het gehele interval. In dat geval kunnen we bovenstaande methoden wel gebruiken, als tenminste de matrix niet te groot is en de integratieformule voldoende nauwkeurig. Men denkt hierbij direct aan de Gauss integratieformules. Het gebruik van deze formules zal dikwijls resultaten geven, welke, zeker voor fysische doeleinden, aanvaardbaar zijn.

Litteratuur

- G.E.Forsythe en W.R.Wasow; Finite Difference Methods for Partial Differential Equations. John Wiley and Sons, Inc. [1960]
- L.Fox; Numerical Solution of Ordinary and Partial Differential Equations. Pergamon Press [1962].
- L.Fox and E.T.Goodwin; Some New Methods of the Numerical Integration of Ordinary Differential Equations. Proc.Cambr.Phil.Soc.45 [1949]. pag.373-388.
- G.Hamel; Integralgleichungen. Springer-Verlag [1949]
- S.G.Mikhlin; Linear Integral Equations. Hindustan Publ. Corp. (India) [1960].
- L.F.Richardson and J.A.Gaunt; The Deferred Approach to the Limit. Phil. Trans. A 226 [1926] pag. 299-361.
- R.D.Richtmyer; Difference Methods for Initial-Value Problems. Interscience Publishers Inc., New York [1957].

Einde eerste deel Numerieke Analyse

(De syllabus van dit gedeelte van het college werd verzorgd door drs.A.J.Geurts, Technische Hogeschool Eindhoven).