

MC SYLLABUS

9.2

COLLOQUIUM ELLIPTISCHE
DIFFERENTIAALVERGELIJKINGEN

CONSTRUCTIEVE OPLOSSINGSMETHODEN

DEEL 2

DOOR

W.P. VAN DEN BRINK
T.M.T. COOLEN
B. DIJKHUIS
P.P.N. DE GROEN
P.J. VANDER HOUWEN
E.M. DE JAGER
N.M. TEMME
R.J. DE VOGELAERE

MATHEMATISCH CENTRUM AMSTERDAM

1970

VOORWOORD

Het colloquium over elliptische differentiaalvergelijkingen, dat in januari 1969 onder leiding van prof.dr. E.M. de Jager en prof.dr. H.A. Lauwerier aanging, is in het collegejaar 1969/1970 voortgezet. De voordrachten die in deze periode zijn gegeven, zijn in dit tweede deel van de syllabus samengebracht.

Terwijl het eerste deel van het colloquium tot onderwerp had de existentie en uniciteit van oplossingen, wordt in het tweede deel voornamelijk ingegaan op methoden om oplossingen van randwaardeproblemen werkelijk te vinden. De behandelde methoden vallen in drie groepen uiteen: numerieke methoden, singuliere storingsrekening en een Riemann-integraal-methode.

In hoofdstuk 6 worden randwaardeproblemen voor elliptische differentiaalvergelijkingen gediscretiseerd, en enige expliciete iteratieve methoden behandeld om de ontstane differentievergelijkingen op te lossen. In hoofdstuk 10 komen enige aspecten van een impliciete iteratieve methode, de "alternating direction method", ter sprake.

Hoofdstuk 7 is gewijd aan onjuist gestelde ("improperly posed") problemen. Er worden enkele voorbeelden gegeven van dergelijke problemen die fysisch zinvol zijn; verder worden methoden gegeven om oplossingen te construeren.

In hoofdstuk 8 wordt de invloed van numerieke fouten bij variatieberekeningen onderzocht. Deze berekeningen zijn van belang voor de bepaling van de eigenwaarden van elliptische operatoren.

De drie auteurs van hoofdstuk 9 stellen singuliere storingsproblemen voor elliptische differentiaalvergelijkingen aan de orde. Voor verschillende soorten van randen wordt een constructie van benaderde oplossingen gegeven.

In hoofdstuk 11 tenslotte, wordt de recente theorie van I.N. Vekua uiteengezet, die representaties geeft van oplossingen van elliptische differentiaalvergelijkingen met twee onafhankelijke variabelen en analytische coëfficiënten.

INHOUD

Voorwoord	i
6. Numerieke oplossingsmethoden	1
6.1. Discretisering van het Dirichlet-probleem voor de Laplace-vergelijking	1
6.2. Iteratieve oplossingsmethoden	5
6.3. Convergentieversnellingen	15
Literatuur	31
7. Onjuist gestelde problemen	34
7.1. Inleiding en voorbeelden	34
7.2. Regularisatie van onjuist gestelde problemen	39
7.3. Quasi-reversibiliteit	52
Literatuur	66
8. Over de invloed van numerieke fouten bij variatieberekeningen	68
8.1. De Rayleigh-Ritz methode	69
8.2. Afrondingsfouten	71
8.3. Een andere schatting voor de bovengrens	75
Literatuur	78
9. Singuliere storingsproblemen voor elliptische differentiaal- vergelijkingen	79
9.1. Probleemstelling	79
9.2. Toepassingen van het maximumprincipe op singuliere storingsproblemen	85
9.3. De benadering voor de functie ϕ_ϵ	90
9.4. Afschatting van restterm	97
9.5. Plaatselijk niet differentieerbare randvoorwaarden	102
9.6. Asymptotische oplossing van een singulier storings- probleem voor een kwartvlak	112
Literatuur	124

10. Optimal cyclic alternating direction iteration	126
10.1. Introduction	126
10.2. Alternating direction methods	126
10.3. Cyclic alternating direction methods	128
10.4. Optimal choice for the Peaceman-Rachford method in the simplest case	129
10.5. The simple cyclic case for Peaceman-Rachford	131
10.6. A classical min-max problem	131
10.7. The case $r = 2^p$ and heuristics	136
10.8. The solution for arbitrary r	138
10.9. Geometric interpretation	140
10.10. Mechanical interpretation	141
10.11. Rate of convergence	147
10.12. Optimization in the general case	149
10.13. Other related min-max problems	151
10.14. Remarks	152
Literature	155
11. I.N. Vekua's theorie der elliptische differentiaalverge- lijkingen met analytische coëfficiënten	157
11.1. Inleiding	157
11.2. De Riemannfunctie van de vergelijking (E_0)	161
11.3. Analytische oplossingen van (E_0)	168
11.4. Fundamentealoplossingen	172
11.5. Analyticiteit van de oplossingen van (E_0)	175
11.6. Analytische voortzetting van de oplossingen van (E_0)	176
11.7. Algemene representatie van de reguliere oplossingen van (E_0)	178
Literatuur	180

AUTEURS

P.J. van der Houwen	Mathematisch Centrum Amsterdam	hoofdstuk 6
L.M.T. Coolen	Mathematisch Centrum Amsterdam	hoofdstuk 7
B. Dijkhuis	Mathematisch Centrum Amsterdam	hoofdstuk 8
E.M. de Jager	Universiteit van Amsterdam	secties 9.1 t.m. 9.4.
P.P.N. de Groen	Vrije Universiteit Amsterdam	sectie 9.5.
N.M. Temme	Mathematisch Centrum Amsterdam	sectie 9.6.
R.J. de Vogelaere	tijdelijk Universiteit van Amsterdam	hoofdstuk 10 *)
N.P. van den Brink	Universiteit van Amsterdam	hoofdstuk 11

*) Hoofdstuk 10 is door R. Geel (Mathematisch Centrum Amsterdam) uitgewerkt.

6. NUMERIEKE OPLOSSINGSMETHODEN

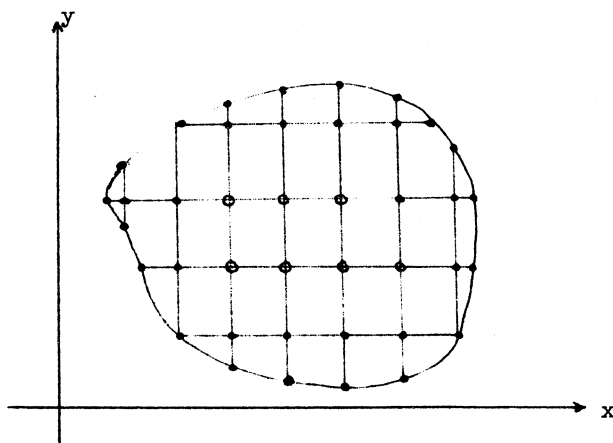
Tot dusver is de constructie van oplossingen van randwaardeproblemen in dit colloquium nog niet ter sprake gekomen. In dit hoofdstuk zullen we ingaan op numerieke methoden om de oplossing te construeren en wel in het bijzonder op differentiemethoden.

6.1. Discretisering van het Dirichlet-probleem voor de Laplace-vergelijking.

Als voorbeeld kiezen we het Dirichlet-probleem voor de vergelijking van Laplace op een begrensde gebied Ω :

$$(6.1) \quad \begin{cases} \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} = 0, (x,y) \in \Omega, \\ U = \phi(x,y), (x,y) \in \partial\Omega. \end{cases}$$

In het algemeen is de oplossing van dit randwaardeprobleem niet op eenvoudige analytische wijze te geven of te benaderen. In de praktijk worden het meest differentiemethoden gebruikt om U te vinden. Men moet zich er dan mee tevreden stellen U niet in alle punten maar in een eindig aantal punten van Ω te berekenen. Deze punten worden meestal gekozen op een rechthoekig rooster zoals aangegeven in figuur 6.1.



Figuur 6.1

Vergelijking (6.1) moet dan ook getransformeerd worden tot een vergelijking voor de waarden van U in deze z.g. roosterpunten. Dit noemt men discretisering van het randwaardeprobleem.

We zullen de roosterpunt-waarden van U met u aangeven en over de roosterfunctie u spreken.

Om het discrete analogen van de Laplace-operator compact te kunnen beschrijven voeren we een verkorte notatie voor differentieoperatoren in, die door de volgende voorbeelden geïllustreerd wordt:

$$(6.2) \quad \begin{aligned} au(x+\xi, y) + bu(x, y) + cu(x-\xi, y) &= \begin{bmatrix} a & b & c \end{bmatrix} u \\ au(x, y+\eta) + bu(x, y) + cu(x, y-\eta) &= \begin{bmatrix} a \\ b \\ c \end{bmatrix} u. \end{aligned}$$

Hierin stellen ξ en η translaties langs de x - en de y -as voor.

Stelling 6.1. In de ruimte van harmonische functies geldt voor $\xi \rightarrow 0$ en $\rho^{-1} = O(1)$

$$(6.3) \quad \begin{aligned} \Delta = D & - \left[\frac{1}{12} \xi^2 \rho^{-2} (\rho^2 - 12L_1 \xi^2 + 1) \frac{\partial^4}{\partial x^4} \right. \\ & + \frac{1}{360} \xi^4 (\rho^2 - 1) (\rho^2 + 1 - 30L_1 \xi^2) \frac{\partial^6}{\partial x^6} \\ & + \frac{1}{20160} \xi^6 \rho^{-6} (\rho^6 + 1 - 28L_1 \xi^2 \rho^4 (2\rho^4 - 5\rho^2 + 2)) \frac{\partial^8}{\partial x^8} \\ & \left. + O(\xi^8) \right] \end{aligned}$$

Hierin is

$$(6.4) \quad D = \begin{bmatrix} L_1 & L_3 & L_1 \\ L_2 & L_4 & L_2 \\ L_1 & L_3 & L_1 \end{bmatrix},$$

$$(6.5) \quad L_1 = \frac{1}{2} \xi^{-2} (1 + \rho^2 - \gamma), \quad L_2 = \xi^{-2} (\gamma - \rho^2), \quad L_3 = \xi^{-2} (\gamma - 1), \quad L_4 = -2\gamma \xi^{-2},$$

$\rho = \xi/\eta$ en γ een reële parameter.

Bewijs Door Du uit te schrijven in termen van $u(x \pm \xi, y \pm \eta)$, deze in Taylorreeksen t.o.v. het punt (x, y) te ontwikkelen en gebruik te maken van het feit dat voor harmonische functies geldt $\partial^2/\partial y^2 = -\partial^2/\partial x^2$ wordt (6.3) verkregen

Deze stelling suggereert om Δ te vervangen door D want DU gaat over in ΔU als $\xi = \eta \rightarrow 0$, waarbij we aannemen dat $\partial^4 U / \partial x^4$ uniform begrensd is op Ω . We zijn er dan van verzekerd dat het differentie-analogon voor $\xi \rightarrow 0$ en $\eta \rightarrow 0$ in de vergelijking van Laplace overgaat, d.w.z. het differentie-analogon is een consistente benadering. Dit betekent nog niet dat de roosterfuncties u naar de analytische oplossing U convergeren. Een analyse van dergelijke discrete benaderingen werd het eerst (1930) gegeven door Gerschgorin [7]. Een algemenere behandeling vindt men in Forsythe en Wasow [4] p. 283. In dit hoofdstuk zal hier niet nader op ingegaan worden; we zullen er van uit gaan dat de differentieoplossingen convergeren. We merken nog op dat in de buurt van de rand $\partial\Omega$ de operator Δ niet door D vervangen kan worden omdat de afstand van de op de rand $\partial\Omega$ gelegen roosterpunten tot de naburige inwendige roosterpunten niet meer ξ of η hoeft te zijn. In figuur 6.1 zijn deze punten met $\bullet$ aangegeven. In deze niet-reguliere punten moet de operator D aangepast worden. Aangezien dit zichzelf wijst zal er niet op ingegaan worden (voor het geval $\gamma = 2$ kan men de formules, afkomstig van Shortley en Weller [16], in Forsythe en Wasow [4] p. 201 vinden). Op de rand $\partial\Omega$ zelf geldt uiteraard $D = 1$.

Het discrete analogon van (6.1) zullen we schrijven als

$$(6.6) \quad Du = f,$$

waarin $f = \Phi$ in de randpunten en $f = 0$ in de niet-randpunten. In het vervolg zullen we u en f voorstellen door vectoren waarin de componenten gerangschikt zijn in de volgorde volgens welke men de roosterpunten in figuur 6.1 zou aftellen. In de praktijk worden meestal alleen de operatoren $D(\gamma)$ gebruikt welke ontstaan voor $\gamma = 1 + \rho^2$ en $\gamma = 5(1 + \rho^2) / 6$.

In het eerste geval wordt $L_1 = 0$ zodat er een vijfpuntsformule ontstaat. We geven deze expliciet voor $\xi = \eta$:

$$(6.7) \quad D(2) = \xi^{-2} \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Ter verduidelijking volgt nu de volledig uitgeschreven vergelijking $D(2) u = f$ voor het discrete Dirichlet-probleem dat correspondeert met figuur 6.2.

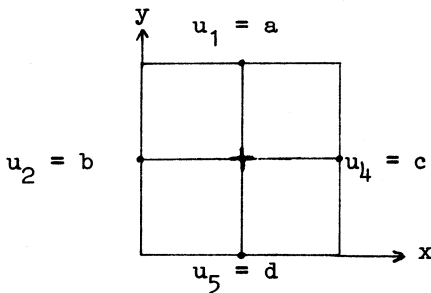


fig. 6.2 Rooster met 1 inwendig roosterpunt voor het geval $\gamma = 2$.

Als a, b, c, d de gegeven randwaarden zijn dan is

$$(6.7') \quad \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ L_2 & L_2 & L_4 & L_2 & L_2 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{pmatrix} = \begin{pmatrix} a \\ b \\ 0 \\ c \\ d \end{pmatrix}.$$

We merken op dat (6.7') ook in de gereduceerde vorm

$$(6.7'') \quad L_4 u_3 = -L_2(a+b+c+d)$$

geschreven kan worden.

De operator $D(5 \frac{1+\rho^2}{6})$ staat bekend als de 9-puntsbenadering van de Laplace operator. Deze operator geeft een 4^{de} orde benadering voor Δ aangezien de coëfficiënt van ξ^2 in (6.3) voor $\gamma = 5(1+\rho^2)/6$ nul wordt. Indien $\rho = 1$ ($\xi=\eta$) is er zelfs sprake van een 6^{de} orde benadering. $D(5/3)$ wordt gegeven door

$$(6.8) \quad D(5/3) = \frac{1}{6}\xi^{-2} \begin{bmatrix} 1 & 4 & 1 \\ 4 & -20 & 4 \\ 1 & 4 & 1 \end{bmatrix}.$$

Naast de operatoren (6.7) en (6.8) noemen we nog de operator $D(1)$. Voor $\rho = 1$ wordt deze voorgesteld door

$$(6.9) \quad D(1) = \frac{1}{2}\xi^{-2} \begin{bmatrix} 1 & 0 & 1 \\ 0 & -4 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

Een expliciet voorbeeld van (6.9) corresponderend met figuur 6.3

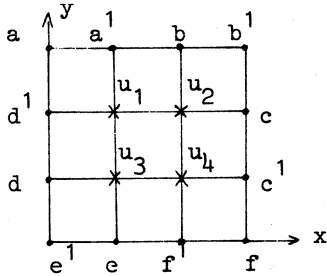


fig. 6.3 Rooster met 4 inwendige roosterpunten voor het geval $\gamma = 1$.

wordt in gereduceerde vorm gegeven door

$$(6.9') \quad \begin{pmatrix} L_4 & 0 & 0 & L_1 \\ 0 & L_4 & L_1 & 0 \\ 0 & L_1 & L_4 & 0 \\ L_1 & 0 & 0 & L_4 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = -L_1 \begin{pmatrix} a + b + d \\ a^1 + b^1 + c^1 \\ d^1 + f^1 + e^1 \\ c + f + e \end{pmatrix}.$$

Evenals (6.7) is (6.9) een vijfpuntsformule. We zullen spreken van de vijf-puntsformule (+) en de vijfpuntsformule (x) in geval van (6.7) en (6.9).

6.2. Iteratieve oplossingsmethoden

In deze sectie zullen een aantal methoden beschreven worden om de matrixvergelijking (6.6) iteratief op te lossen. Dit in tegenstelling tot directe oplossingsmethoden (zoals b.v. eliminatiemethoden); deze worden zelden gebruikt om elliptische differentievergelijkingen op te lossen omdat het feit dat de te inverteren matrix D hoofdzakelijk uit nullen bestaat niet voldoende benut wordt. Behandeld zullen worden:

- (a) methode van Jacobi;
- (b) methode van Gauss-Seidel;
- (c) methode van Frankel.

Deze methoden vormen het uitgangspunt voor het merendeel van de tot nu toe ontwikkelde iteratieve methoden.

Convergentiesnelheid

Om de praktische waarde van een iteratieve methode te meten moet men op de een of andere manier een maat vinden voor de snelheid waarmee het proces convergeert. Vooruitlopend op de exacte definitie van de verschillende iteratieve methoden merken we op dat bovengenoemde basisprocessen geschreven kunnen worden als stationaire lineaire processen, d.w.z. in de vorm

$$(6.10) \quad u_{k+1} = H u_k + M f, \quad k = 0, 1, 2, \dots$$

Hierin stelt u_0 de beginapproximatie van het proces voor.

Als (6.10) convergeert, dan is dit naar de vector

$$(6.11) \quad u = (1-H)^{-1} M f.$$

Definieer nu de vectoren $v_k = u_k - u$. Deze voldoen aan het schema

$$(6.12) \quad v_{k+1} = H v_k, \quad k = 0, 1, 2, \dots$$

Dus na k iteraties wordt de fout v_k gegeven door

$$(6.13) \quad v_k = H^k v_0.$$

De bruikbaarheid van het iteratieproces hangt af van de grootte van $\|H^k\|$, immers

$$(6.14) \quad \|v_k\| \leq \|H^k\| \|v_0\|.$$

Hierin stelt $\| \cdot \|$ een of andere norm voor in de ruimte der roosterfuncties. In 1954 stelde Young [21] voor de gemiddelde convergentiesnelheid na k iteraties $R(k)$ te definiëren door

$$(6.15) \quad R(k) = - \frac{\ln \|H^k\|}{k}.$$

Stelling 6.2. Voor grote waarden van k wordt de gemiddelde convergentiesnelheid (6.15) gegeven door

$$(6.16) \quad R(k) \sim \left(1 - \frac{p-1}{k}\right) \ln [\sigma(H)]^{-1} - \frac{\ln v + (p-1) \ln k}{k},$$

waarin v een constante is, $\sigma(H)$ de spectraalradius van H (d.w.z. het maximum van de absolute waarden van de eigenwaarden van H), en waarin p de grootste orde van alle diagonale ondermatrices J_r van de Jordan-normaalvorm J van H met $\sigma(J_r) = \sigma(H)$ is.

Bewijs Zie [10] p. 31.

We merken nog op dat voor normale matrices H (d.w.z. $HH^* = H^*H$) geldt $p = v = 1$, zodat (6.16) gereduceerd wordt tot

$$(6.17) \quad R(k) = \ln [\sigma(H)]^{-1}.$$

Bepalend voor de bruikbaarheid van een stationair lineair iteratieproces is dus de spectraalradius van de iteratieve matrix. Men moet H zo construeren dat deze zo klein mogelijk is.

Het modelprobleem

Om een indruk te krijgen van de convergentiesnelheden van de verschillende iteratieve methoden kiest men meestal de matrixvergelijking $Du = f$ welke voor $\xi = \eta$ ontstaat uit het Dirichlet-probleem (6.1) waarin Ω een vierkant met zijden π is. Dit wordt het modelprobleem genoemd.

In de constructie van iteratieve processen wordt dit probleem gekozen omdat men hiervoor meestal een gedetailleerde analyse kan geven. Hieruit kan men dan de eigenschappen van het proces leren kennen.

In het volgende zullen de eigenwaarden van de met het modelprobleem corresponderende matrix D een belangrijke rol spelen.

Stelling 6.3 De eigenwaarden $\delta_{n,m}$ en eigenfuncties $e_{n,m}$ van het modelprobleem worden gegeven door

$$(6.18) \quad \delta_{n,m} = 2 \left[-\gamma + (\gamma-1)(\cos n\xi + \cos m\xi) + (2-\gamma) \cos n\xi \cos m\xi \right] \xi^{-2},$$

$$(6.19) \quad e_{n,m} = \sin nj \xi \sin ml \xi,$$

waarin $n, m = 1, 2, \dots, \frac{\pi}{\xi} - 1$ en $(j \xi, l \xi)$ de roosterpunten voorstellen.

Bewijs Zie [12].

Uit deze stelling kan men voor $\xi \rightarrow 0$ de volgende grenzen voor $\delta_{n,m}$ afleiden

$$(6.20) \quad -4 \xi^{-2} + 2(2-\gamma) \leq \delta_{n,m} \leq -2, \quad 1 \leq \gamma \leq \frac{3}{2},$$

$$-8(\gamma-1)\xi^{-2} + 2(3\gamma-4) \leq \delta_{n,m} \leq -2, \quad \gamma \geq \frac{3}{2}.$$

Voor niet-modelproblemen kan men géén expliciete uitdrukkingen geven voor de eigenwaarden en eigenfuncties. Wel kunnen grenzen voor de eigenwaarden gegeven worden. Hiervoor zij verwezen naar [12].

Methode van Jacobi

Het eerste in de literatuur vermelde iteratieproces voor matrixvergelijkingen van het type (6.6) dateert uit 1845. Dit proces wordt de methode van Jacobi genoemd. Om deze methode compact te kunnen weergeven schrijven we D in de vorm (zie Varga [17] p.56)

$$(6.21) \quad D = C - E - F,$$

waarin $C = \text{diag} \{d_{11}, d_{22}, \dots, d_{nn}\}$ (n is de orde van de matrix D), en waarin E en F respectievelijk de onder- en bovendriehoeksmatrices zijn waarvan de elementen de negatieve waarden van de corresponderende matrixelementen van D hebben.

Voor matrix (6.4) geldt dus in de reguliere punten volgens notatie (6.2)

$$(6.22) \quad C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & L_4 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad E = \begin{bmatrix} -L_1 & -L_3 & -L_1 \\ -L_2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad F = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -L_2 \\ -L_1 & -L_3 & -L_1 \end{bmatrix}$$

De methode Jacobi wordt nu gedefinieerd door

$$(6.23) \quad u_{k+1} = C^{-1}(E + F) u_k + C^{-1}f = (1 - C^{-1}D) u_k + C^{-1}f, \quad k = 0, 1, 2, \dots$$

Dus de matrix H uit (6.10) wordt gegeven door

$$(6.24) \quad H = C^{-1}(E + F) = (1 - C^{-1}D) = 1 - \frac{D}{L_4}.$$

Het is duidelijk dat als de methode van Jacobi convergeert, deze naar de oplossing van $Du = f$ convergeert.

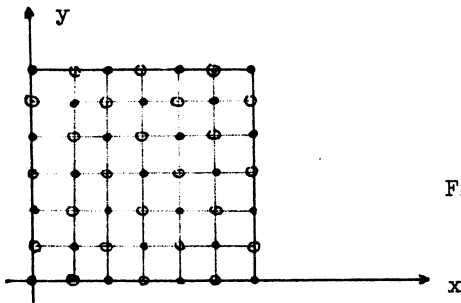
Uit (6.5) en (6.20) volgt dan

$$(6.25) \quad \sigma(H) = \begin{cases} \text{Max} \left\{ 1 - \frac{\xi^2}{\gamma}, \left| 1 - \frac{2}{\gamma} + \frac{2-\gamma}{\gamma} \xi^2 \right| \right\} & = 1 - \frac{\xi^2}{\gamma}, \quad 1 \leq \gamma \leq \frac{3}{2} \\ \text{Max} \left\{ 1 - \frac{\xi^2}{\gamma}, \left| 1 - 4 \frac{\gamma-1}{\gamma} + \frac{3\gamma-4}{\gamma} \xi^2 \right| \right\} & = 1 - \frac{\xi^2}{\gamma}, \quad \gamma \geq \frac{3}{2}. \end{cases}$$

Verder is het duidelijk dat de matrix D en dus ook H symmetrisch is zodat volgens formule (6.17)

$$(6.26) \quad R(k) = -\ln \left(1 - \frac{\xi^2}{\gamma} \right) \sim \frac{\xi^2}{\gamma}.$$

Hieruit volgt dat de vijfpuntsformule ($\times$) tweemaal zo snel convergeert als de vijfpuntsformule (+). Voorts is de ($\times$) formule tweemaal minder bewerkelijk (bij dezelfde maaswijdte) als de (+) formule. Dit volgt uit het feit dat in het iteratieproces met $\gamma = 1$ de vergelijkingen voor u_k in de punten $\circ$ niet gesoppeld zijn met de vergelijkingen in de punten $\bullet$ (Zie figuur 6.4).



Figuur 6.4. Het modelprobleem

Anderzijds mag men, alhoewel beide formules dezelfde orde van nauwkeurigheid hebben, verwachten dat de (+) formule nauwkeuriger is (bij dezelfde ξ). Het zou interessant zijn om te onderzoeken welke formule in de praktijk te prefererener is.

Tenslotte merken we op dat de inverse C^{-1} van C , welke in (6.21) optreedt, direct bepaald kan worden omdat C een diagonaalmatrix is.

Methode van Gauss-Seidel

In 1874 gaf Seidel het volgende iteratieproces aan:

$$(6.27) \quad u_{k+1} = (C-E)^{-1} F u_k + (C-E)^{-1} f = (1 - (C-E)^{-1} D) u_k + (C-E)^{-1} f.$$

Dit proces werd echter al veel eerder (1832) door Gauss gebruikt, maar deze heeft het niet de moeite waard gevonden om het te publiceren (zie Bodewig [1], p.128).

Door de matrixvergelijking (6.27) uit te schrijven kan men eenvoudig inzien dat de vergelijking in wezen expliciet is, d.w.z. er hoeft geen matrix geïnverteerd te worden.

De iteratieve matrix H wordt gegeven door

$$(6.28) \quad H = (C-E)^{-1} F = 1 - (C-E)^{-1} D.$$

Om de convergentiesnelheid te bepalen moet men volgens stelling 6.2 de spectraalradius van de eigenwaarden van H kennen.

Stelling 6.4. De eigenwaarden $\mu_{n,m}$ van de matrix (6.28) worden voor het modelprobleem gegeven door de vergelijking:

$$(6.29) \quad R = 2 \sqrt{PQ} \cos n \xi,$$

$$\text{waarin} \quad P = L_2 + 2L_1 \sqrt{\mu_{n,m}} \cos m \xi.$$

$$Q = L_2 \mu_{n,m} + 2L_1 \sqrt{\mu_{n,m}} \cos m \xi,$$

$$R = -L_4 \mu_{n,m} - 2L_2 \sqrt{\mu_{n,m}} \cos m \xi$$

$$n, m = 1, 2, \dots, \frac{\pi}{\xi} - 1.$$

Bewijs. Een bewijs vindt men in [12], p.16 (in deze referentie moet men $p = q = \omega_k = -1/L_4$ en $\lambda = L_4(\mu-1)$ stellen).

In het algemeen leidt (6.29) tot een vergelijking van de vierde graad $\sqrt{\mu_{n,m}}$. In het geval van de beide vijfpuntsformules kunnen we μ echter

expliciet vinden:

$$(6.30) \quad \mu_{n,m}^+ = \frac{1}{4} (\cos m \xi + \cos n \xi)^2,$$

$$(6.31) \quad \mu_{n,m}^x = \cos^2 m \xi \cos^2 n \xi.$$

Hieruit volgt dat voor $\xi \rightarrow 0$

$$(6.32) \quad \sigma(H) = 1 - \frac{2\xi^2}{\gamma}, \quad \gamma = 1, 2.$$

Aangezien de matrix H in de methode van Gauss-Seidel niet symmetrisch is wordt de convergentiesnelheid gegeven door formule (6.16):

$$(6.33) \quad R(k) \sim \frac{2}{\gamma} \left(1 - \frac{p-1}{k}\right) \xi^2 - \frac{\ln v + (p-1) \ln k}{k}, \quad \gamma = 1, 2.$$

Asymptotisch ($k \rightarrow \infty$) geldt dus

$$(6.34) \quad R(\infty) \sim \frac{2\xi^2}{\gamma}.$$

Vergelijken we dit met (6.26) dan zien we dat de vijfpuntsformules van de Gauss-Seidelmethode voor grote waarden van k tweemaal zo snel convergeren als de overeenkomstige vijfpuntsformules van de methode van Jacobi.

Methode van Frankel

In 1950 beschouwde Frankel [6] het volgende tweede orde proces:

$$(6.35) \quad u_{k+1} = (\alpha D + \beta) u_k + (1 - \beta) u_{k-1} - \alpha f, \quad k = 0, 1, \dots$$

Hierin zijn α en β nader te bepalen parameters.

Behalve u_0 is in dit proces ook u_1 nodig om de berekening te starten.

Door een vector

$$(6.36) \quad \vec{w}_k = \begin{pmatrix} u_k \\ u_{k-1} \end{pmatrix}$$

in te voeren kan (6.35) geschreven worden als een stationair lineair proces, namelijk

$$(6.37) \quad \vec{w}_{k+1} = \begin{pmatrix} \beta + \alpha D & 1 - \beta \\ 1 & 0 \end{pmatrix} \vec{w}_k - \begin{pmatrix} \alpha f \\ 0 \end{pmatrix}.$$

De matrix H wordt nu dus gegeven door

$$(3.8) \quad H = \begin{pmatrix} \beta + \alpha D & 1 - \beta \\ 1 & 0 \end{pmatrix}.$$

Stelling 6.5. Wanneer $\delta_{n,m}$ en $d_{n,m}$ respectievelijk de eigenwaarden en eigenfuncties van D voorstellen dan worden de eigenwaarden $\mu_{n,m}$ en eigenfuncties $e_{n,m}$ van H gegeven door

$$(6.39) \quad \mu_{n,m}^2 - (\beta + \alpha \delta_{n,m}) \mu_{n,m} + \beta - 1 = 0,$$

$$(6.40) \quad e_{n,m} = \begin{pmatrix} \mu_{n,m} \\ 1 \end{pmatrix} d_{n,m}.$$

Bewijs. Volgt door uitwerking van de vergelijking $He_{n,m} = \mu_{n,m} e_{n,m}$ (we merken op dat D niet noodzakelijk met het modelprobleem hoeft te corresponderen).

De volgende stap is om α en β zo te kiezen dat $\sigma(H) = \max_{n,m} |\mu_{n,m}|$ minimaal is. Hiertoe gebruiken we het volgende lemma.

Lemma 6.1. De wortels van de vergelijking

$$z^2 - Sz + P = 0, \quad S \text{ en } P \text{ reëel,}$$

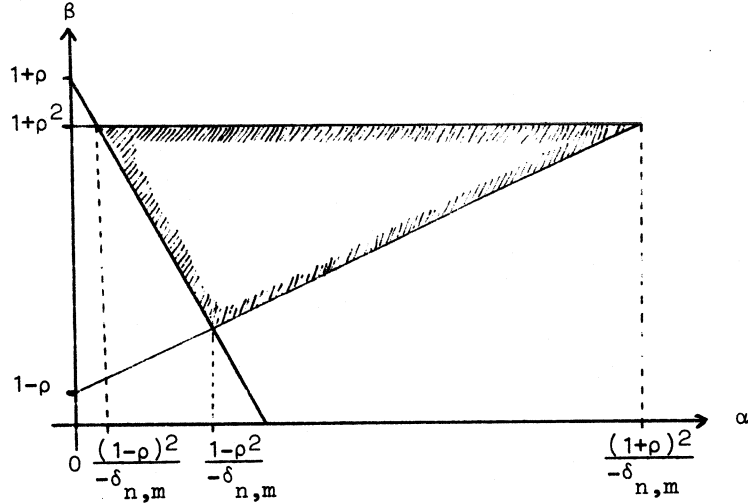
liggen binnen of op de cirkel $|z| = \rho$ wanneer aan de volgende noodzakelijke voorwaarden voldaan is

$$(6.41) \quad P \leq \rho^2, \quad \rho^2 - Sp + P \geq 0, \quad \rho^2 + Sp + P \geq 0.$$

Bewijs. Elementair.

Passen we dit lemma nu toe op vergelijking (6.39) met $S = \beta + \alpha \delta_{n,m}$ en $P = \beta - 1$ dan gelden de volgende ongelijkheden voor α en β ($\delta_{n,m}$ wordt reëel verondersteld):

$$(6.42) \quad \left\{ \begin{array}{l} \beta \leq \rho^2 + 1, \\ \beta \geq \frac{\rho \delta_{n,m}}{1-\rho} \alpha + 1 + \rho, \quad \rho < 1, \\ \beta \geq -\frac{\rho \delta_{n,m}}{1+\rho} \alpha + 1 - \rho, \quad \rho < 1. \end{array} \right.$$



Figuur 6.5

In figuur 6.5 is het gebied gearceerd waarin de punten (α, β) liggen die voldoen aan de voorwaarden (6.42). Voor deze α en β geldt dus $|\mu_{n,m}| \leq \rho$ waarin $\mu_{n,m}$ de met de gekozen $\delta_{n,m}$ corresponderende eigenwaarde is. Voor iedere $\delta_{n,m}$ is er zo'n driehoek in het (α, β) vlak. De doorsnede van al deze driehoeken is alleen dan niet leeg wanneer

$$\frac{(1+\rho)^2}{|\delta_{n,m}|_{\max}} \geq \frac{(1-\rho)^2}{|\delta_{n,m}|_{\min}} \quad \text{ofwel} \quad \rho \geq \frac{\sqrt{|\delta_{n,m}|_{\max}} - \sqrt{|\delta_{n,m}|_{\min}}}{\sqrt{|\delta_{n,m}|_{\max}} + \sqrt{|\delta_{n,m}|_{\min}}}.$$

In het geval van het modelprobleem vinden we voor de minimale waarde van ρ en dus van $\sigma(H)$ (stelling 6.3)

$$(6.43) \quad \rho_{\min} = \sigma(H) \sim \begin{cases} 1 - \sqrt{2} \xi, & 1 \leq \gamma \leq \frac{3}{2} \\ 1 - \frac{1}{\sqrt{\gamma-1}} \xi, & \gamma \geq \frac{3}{2}. \end{cases}$$

De waarden van α en β waarvoor dit minimum bereikt wordt zijn (zie figuur 6.5)

$$(6.44) \quad \left\{ \begin{array}{l} \alpha = \frac{(1-\rho_{\min})^2}{|\delta_{n,m}|_{\min}} \sim \frac{(1-\rho_{\min})^2}{2}, \\ \beta = 1 + \rho_{\min}^2. \end{array} \right.$$

Uit (6.40) volgt dat, alhoewel H niet symmetrisch is, zijn eigenfuncties in het geval van het modelprobleem toch orthogonaal zijn ($\mu_{n,m}^+ \mu_{n,m}^- = 1$, $d_{n,m} \sin n\xi \sin m\xi$) zodat H een normale matrix is. Hieruit volgt dat de convergentiesnelheid gegeven wordt door formule (6.17), dus

$$(6.45) \quad R(k) \sim \begin{cases} \sqrt{2} \xi & , \quad 1 \leq \gamma \leq \frac{3}{2}, \\ \frac{1}{\sqrt{\gamma-1}} & , \quad \gamma \geq \frac{3}{2}. \end{cases}$$

In tegenstelling tot de methoden van Jacobi en Gauss-Seidel die een convergentiesnelheid $O(\xi^2)$ hebben, is de convergentiesnelheid van het Frankel-proces $O(\xi)$. Voor kleine waarden van ξ betekent dit een aanzienlijke verbetering. Daar tegenover staat echter dat de methode van Frankel tweemaal meer geheugenruimte vergt als de eerste methoden.

Tenslotte merken we nog op dat de formules (6.43) - (6.45) zonder meer gegeneraliseerd kunnen worden voor matrixproblemen $Du = f$ waarin D negatieve eigenwaarden heeft. Stellen we

$$(6.46) \quad a = |\delta_{n,m}|_{\min}, \quad b = |\delta_{n,m}|_{\max},$$

dan vinden we

$$(6.43') \quad \sigma(H) = \frac{\sqrt{b} - \sqrt{a}}{\sqrt{b} + \sqrt{a}},$$

$$(6.44') \quad \alpha = \frac{4}{(\sqrt{b} + \sqrt{a})^2}, \quad \beta = \frac{2(b+a)}{(\sqrt{b} + \sqrt{a})^2},$$

$$(6.45') \quad R(k) \sim 2\left(1 - \frac{p-1}{k}\right) \sqrt{\frac{a}{b}} - \frac{\ln v + (p-1) \ln k}{k}, \quad a \ll b.$$

Is D bovendien symmetrisch dan geldt

$$(6.45'') \quad R(k) \sim 2\sqrt{\frac{a}{b}}, \quad a \ll b.$$

6.3. Convergentieversnellingen

In deze sectie zal een methode beschreven worden waarmee een gegeven iteratieproces van de vorm

$$(6.10) \quad u_{k+1} = H u_k + M f$$

versneld kan worden. Deze versnellingsmethode bestaat uit de introductie van een aantal parameters welke zo gekozen worden dat de fout v_k sneller naar 0 convergeert. Als maat hiervoor neemt men meestal, zoals al eerder opgemerkt is, de convergentiesnelheid, dus in feite de spectraalradius van de iteratieve operator. In (6.10) is dit dus $\sigma(H)$. In sommige gevallen is het echter eenvoudiger om niet $\sigma(H) = \max_j |\mu_j|$ te nemen maar een andere norm voor de eigenwaarden μ_j van de iteratieve operator, b.v.

$$(6.47) \quad \|\vec{\mu}\|_p = \left[\sum_j |\mu_j|^p \right]^{1/p}.$$

Hierin stelt $\vec{\mu}$ de vector $(\mu_1, \mu_2, \dots)$ voor.

Dit wordt gerechtvaardigd door het volgende: stel dat H een volledig stelsel eigenfuncties $\{e_j\}$ heeft dan kan de fout v_k geschreven worden als

$$(6.48) \quad v_k = \sum_j c_j e_j$$

en de fout v_{k+1} als

$$(6.49) \quad v_{k+1} = \sum_j c_j \mu_j e_j.$$

Hieruit volgt

$$\|v_{k+1}\| < \sum_j |c_j| |\mu_j|,$$

waarin verondersteld is dat $\|e_j\| = 1$.

Het beste zou zijn dat $\sum_j |c_j| |\mu_j|$ geminimaliseerd kon worden, maar dat eist te veel kennis van de oplossing. Schrijven we echter (Hölder's ongelijkheid)

$$(6.50) \quad \|v_{k+1}\| \leq \sum_j |c_j| |\mu_j| \leq \left[\sum_j |\mu_j|^p \right]^{1/p} \left[\sum_j |c_j|^q \right]^{1/q} \\ = \|\vec{\mu}\|_p \|\vec{c}\|_q, \quad \frac{1}{p} + \frac{1}{q} = 1,$$

dan kunnen we door minimalisering van $\|\vec{\mu}\|_p$ bereiken dat v_k in de zin van de norm $\|\cdot\|_p$ maximaal gedempt wordt.

In dit hoofdstuk zal ingegaan worden op de z.g. methode van Richardson. Voor andere versnellingsmethoden, zoals de SOR-methode en de SSOR-methode wordt verwezen naar de literatuur, b.v. [4], [5], [8], [9], [10], [13], [15], [17], [18], [19] en [21].

Methode van Richardson voor het geval van reële eigenwaarden

Richardson [14] beschouwde in 1910 het volgende niet-stationaire iteratieproces

$$(6.51) \quad u_{k+1} = (1-\omega_k) u_k + \omega_k (Hu_k + Mf), \quad k = 0, 1, 2, \dots$$

Hierin stellen de ω_k 's de zogenaamde relaxatieparameters voor. De fout v_k wordt gegeven door

$$(6.52) \quad v_k = P_k(H) v_0,$$

Waarin $P_k(H)$ een polynoom van de graad k in H is met de eigenschap dat $P_k(1) = 1$; immers

$$(6.53) \quad P_k(H) = \prod_{j=0}^{k-1} (1 + \omega_j(H-1)).$$

Richardson gebruikte dergelijke schema's om iteratieprocessen te versnellen waarin H reële eigenwaarden μ heeft. Daartoe trachtte hij polynomen $P_k(\mu)$ te construeren die voldoen aan $P_k(1) = 1$ en die op het interval $[\mu_{\min}, \mu_{\max}]$ een zo klein mogelijke maximum-norm hebben. Hij koos voor de nulpunten van $P_k(\mu)$ de punten die het interval $[\mu_{\min}, \mu_{\max}]$ in gelijke stukken verdelen en wist daarmee (aannemende dat $\delta f \mu < 1$ of $\mu > 1$) een aanzienlijke convergentieversnelling te verkrijgen. Toch waren de door Richardson gebruikte polynomen niet optimaal (in de maximum-norm).

Stelling 6.6. Het polynoom

$$(6.54) \quad P_k(\mu) = \frac{T_k\left(\frac{2\mu - (\mu_{\max} + \mu_{\min})}{\mu_{\max} - \mu_{\min}}\right)}{T_k\left(\frac{2 - (\mu_{\max} + \mu_{\min})}{\mu_{\max} - \mu_{\min}}\right)}$$

heeft van alle k^{de} graads polynomen in μ die voldoen aan $P_k(1) = 1$ de kleinste maximum-norm over het interval $[\mu_{\min}, \mu_{\max}]$.

Bewijs Volgt direct uit de minimax-eigenschap van de Chebyshev-polynomen $T_k(y)$. Deze minimax-eigenschap werd door Markov in 1886 bewezen.

Uit (6.54) kan eenvoudig de convergentiesnelheid van het optimale Richardson-proces afgeleid worden. Definieer

$$(6.55) \quad a = 1 - \mu_{\max}, \quad b = 1 - \mu_{\min},$$

en veronderstel dat $a > 0$ (van nu af aan zal dit steeds aangenomen worden) dan is

$$(6.56) \quad \sigma(P_k(\mu)) = \|P_k(\mu)\|_{\infty} = \max_{\mu_{\min} \leq \mu \leq \mu_{\max}} |P_k(\mu)| = \left\{ \cosh \left[\ln \frac{\sqrt{b} + \sqrt{a}}{\sqrt{b} - \sqrt{a}} \right] \right\}^k$$

voor $a < b$ en H normaal levert dit

$$(6.57) \quad R(k) \sim \frac{1}{k} \ln \left[\cosh 2\sqrt{\frac{a}{b}} \right] \sim 2\sqrt{\frac{a}{b}} - \frac{\ln 2}{k}.$$

Laten we de methode van Richardson toepassen op het iteratieproces van Jacobi. Eenvoudigheidshalve kiezen we in de matrix D de diagonaalmatrix $C = -I$; dan is $H = I + D$ (zie (6.24)), zodat $\mu = 1 + \delta$. Hieruit volgt dat $a = |\delta|_{\min}$ en $b = |\delta|_{\max}$. Voor symmetrische matrices D (of nog algemener voor normale matrices D) geldt (6.57) waarin a en b op dezelfde manier gedefinieerd zijn als in formule (6.45"). We zien hieruit dat het Frankel proces en de methode van Richardson toegepast op Jacobi's schema asymptotisch dezelfde convergenti snelheid hebben.

Een belangrijk verschil tussen de methoden van Frankel en Richardson is echter dat de eerste stabiel is en de tweede voor grote waarden van k en b/a instabiel. Deze instabiliteit is zeer gevoelig voor de volgorde van de relaxatieparameters ω_j en men kan veel bereiken door deze in een zeer specifieke volgorde te kiezen (zie [11]). Heeft men echter voldoende geheugenruimte ter beschikking dan kan men m.b.v. een drie-niveau-schema een stabiele versie van het Richardson-proces verkrijgen, het z.g. tweede orde Richardson-proces. Dit is gebaseerd op de recurrente betrekkingen die tussen orthogonale polynomen bestaan. Zo geldt voor de polynomen (6.54):

$$(6.58) \quad \begin{cases} P_0(\mu) = 1, \\ P_1(\mu) = \frac{2}{b-a} (H-1) + 1, \\ P_{k+1}(\mu) = (2y_0 + \frac{4}{b-a}(\mu-1)) \frac{T_k(y_0)}{T_{k+1}(y_0)} P_k(\mu) - \frac{T_{k-1}(y_0)}{T_{k+1}(y_0)} P_{k-1}(\mu), \quad k > 0, \end{cases}$$

waarin $y_0 = (b+a)/(b-a)$ (vergelijk [10], p. 91).

Schrijf nu het drie-niveauschema in de vorm (vergelijk methode van Frankel)

$$(6.59) \quad \begin{cases} u_1 = (\alpha_0(H-1)+1) u_0 + \alpha_0 f, \\ u_{k+1} = (\alpha_k(H-1)+\beta_k) u_k + (1-\beta_k) u_{k-1} + \alpha_k f, \quad k > 0, \end{cases}$$

dan is

$$v_k = Q_k(H) v_0,$$

waarin

$$(6.60) \quad \begin{cases} Q_0(\mu) = 1, \\ Q_1(\mu) = \alpha_0(\mu-1) + 1, \\ Q_{k+1}(\mu) = (\alpha_k(\mu-1) + \beta_k) Q_k(\mu) + (1-\beta_k) Q_{k-1}(\mu). \end{cases}$$

Door $P_k(\mu)$ en $Q_k(\mu)$ te identificeren bereiken we dat het Richardson proces gesimuleerd kan worden door een niet-stationair Frankel proces. Er geldt dan in (6.59)

$$(6.61) \quad \begin{cases} \alpha_0 = \frac{2}{b-a}, \quad \beta_0 = 1, \\ \alpha_k = \frac{4}{b-a} \frac{T_k(y_0)}{T_{k+1}(y_0)}, \quad \beta_k = 2 y_0 \frac{T_k(y_0)}{T_{k+1}(y_0)}. \end{cases}$$

Deze formulering van het Richardson proces werd door Stiefel in 1958 gegeven.

Stelling 6.7. Zij $H = I + D$ en laten de eigenwaarden van δ_j van D voldoen aan $-b \leq \delta_j \leq -a$, dan is schema (6.59) stabiel als

$$(6.62) \quad \alpha_k > 0, \quad \frac{1}{2} \alpha_k b < \beta_k < 2.$$

Bewijs Zie [10], p. 53.

Uit deze stelling volgt direct dat de methode van Frankel gedefinieerd door (6.44') en het tweede orde Richardson proces gedefinieerd door (6.61) stabiel zijn.

Naast de polynomen (6.54) welke optimaal in de maximum-norm zijn beschouwen we polynomen $P_k(\mu)$ welke optimaal zijn in de kwadraat-norm, d.w.z.

$$(6.63) \quad \|P_k(\mu)\|_2 = \left[\int_{\mu_{\min}}^{\mu_{\max}} P_k^2(\mu) d\mu \right]^{\frac{1}{2}}$$

is minimaal onder de nevenvoorwaarde $P_k(1) = 1$.

Deze integraal wordt niet geminimaliseerd door één van de bekende orthogonal polynomen. Er geldt echter de volgende stelling:

Stelling 6.8. Het polynoom

$$(6.64) \quad P_k(\mu) = 1 - a_1 \frac{1-\mu}{b} - a_2 \left(\frac{1-\mu}{b}\right)^2 - \dots - a_k \left(\frac{1-\mu}{b}\right)^k,$$

waarin

$$(6.65) \quad \sum_{l=1}^k \frac{j+1}{j+1+l} \left(1 - \left(\frac{a}{b}\right)^{j+1+l}\right) a_l = 1 - \left(\frac{a}{b}\right)^{j+1}, \quad j = 1, 2, \dots, k$$

en waarin $a = 1 - \mu_{\max}$, $b = 1 - \mu_{\min}$, heeft van alle k^{de} graads polynomen in μ die voldoen aan $P_k(1) = 1$ de kleinste kwadraat -norm over het interval $[\mu_{\min}, \mu_{\max}]$.

Bewijs Vergelijking (6.65) voor de coëfficiënten a_j volgt direct uit differentiatie van $\|P_k(\mu)\|_2$ naar a_j .

In de praktijk is $a \ll b$ zodat (6.65) benaderd kan worden door

$$(6.65') \quad \sum_{l=1}^k \frac{j+1}{j+1+l} a_l = 1.$$

In tabel 6.1 zijn de oplossingen van deze vergelijking voor $k = 1, 2, \dots, 8$ gegeven:

Tabel 6.1

k	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
1	1.500							
2	4.000	-3.333						
3	7.500	-15.00	8.750					
4	12.00	-42.00	56.00	-25.20				
5	17.50	-93.33	210.0	-210.0	77.00			
6	24.00	-180.0	600.0	-990.0	792.0	-245.1		
7	31.50	-315.0	1444	-3464	4503	-3002	804.1	
8	39.71	-506.8	3027	-9802	18240	-19492	11113	-2619

Naar aanleiding van deze stelling werd door Prof. Lauwerier opgemerkt dat het polynoom $P_k(\mu)$ wel expliciet uitgedrukt kan worden in een stelsel Legendre-polynomen, of, zo men wil, in een willekeurig stelsel orthogonale polynomen, met men de bijbehorende gewichtsfunctie in de integraal (6.63) opneemt

Stelling 6.8.' Zij $\{p_j(y)\}$ een stelsel polynomen dat op het interval $-1 \leq y \leq 1$ orthogonaal is t.o.v. de gewichtsfunctie $w(y)$, dan minimaliseert het polynoom

$$(6.64') \quad P_k(\mu) = \frac{\sum_{j=0}^k p_j\left(\frac{b+a}{b-a}\right) h_j^{-1} p_j\left(\frac{b+a}{b-a} - 2 \frac{1-\mu}{b-a}\right)}{\sum_{j=0}^k p_j^2\left(\frac{b+a}{b-a}\right) h_j^{-1}},$$

onder alle polynomen van de orde k , waarin $a = 1 - \mu_{\max}$, $b = 1 - \mu_{\min}$ en

$$(6.68) \quad h_j = \int_{-1}^1 w(y) p_j^2(y) dy,$$

de integraal (gewogen kwadraat-norm)

$$(6.63') \quad \int_{\mu_{\min}}^{\mu_{\max}} w\left(\frac{b+a}{b-a} - 2 \frac{1-\mu}{b-a}\right) P_k^2(\mu) d\mu, \quad P_k(1) = 1.$$

Bewijs. We definiëren de functie

$$y(\mu) = \frac{b+a}{b-a} - 2 \frac{1-\mu}{b-a}$$

en we schrijven $y(1) = y_0$ (vgl. (6.68)).

stel dat

$$P_k(\mu) = \sum_{j=0}^k a_j p_j(y(\mu)),$$

dan volgt uit de orthogonaliteit dat

$$\int_{\mu_{\min}}^{\mu_{\max}} w(y(\mu)) P_k^2(\mu) d\mu = \frac{1}{2}(b-a) \sum_{j=0}^k \int_{-1}^1 a_j^2 w(y) P_j^2(y) dy = \frac{1}{2}(b-a) \sum_{j=0}^k a_j^2 h.$$

Minimalisering van deze uitdrukking onder de nevenconditie

$$P_k(1) = \sum_{j=0}^k a_j p_j(y_0) = 1 \text{ levert (6.64')}. \quad \text{levert (6.64')}$$

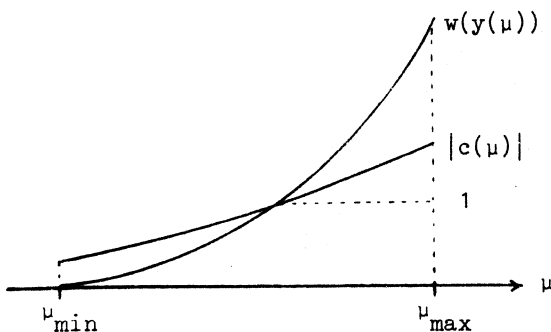
Men kan van de vrijheid dat de gewichtsfunctie $w(y)$ nog te kiezen is gebruik maken als men enige globale informatie over de coëfficiënten $c_j = c(\mu_j)$ in de ontwikkeling van de beginfout v_0 heeft. Stel dat H een orthonormaal stelsel eigenfuncties e_j heeft, dan geldt:

$$||v_k|| = \sum_{j=1}^m c^2(\mu_j) P_k^2(\mu_j) \sim \frac{m}{b-a} \int_{\mu_{\min}}^{\mu_{\max}} c^2(\mu) P_k^2(\mu) d\mu.$$

Het iteratie-proces is dus optimaal voor de gekozen beginapproximatie u_0 wanneer de polynomen $p_j(y)$ orthogonaal zijn t.o.v. de gewichtsfunctie $w(y)$ gedefinieerd door

$$w(y(\mu)) = c^2(\mu).$$

Dit eist te veel kennis van de oplossing. Maar we kunnen in sommige gevallen $w(y(\mu))$ wel het karakter van het gedrag van $c^2(\mu)$ geven. Bijvoorbeeld, wanneer in de beginfout v_0 de eigenfuncties behorende bij grote μ -waarden sterk vertegenwoordigd zijn, dus $|c(\mu)|$ heeft het in figuur 6.6 geschetste gedrag,



Figuur 6.6
Gedrag van $w(y(\mu))$

dan heeft $w(y(\mu))$ dit gedrag nog veel uitgesprokener; men zou kunnen kiezen

$$w(y(\mu)) = (1-y(\mu))^\alpha, \alpha < 0$$

of

$$w(y(\mu)) = (1+y(\mu))^\beta, \beta > 0.$$

Dit leidt tot de Jacobi-polynomen $p_j(y) = P_j^{(\alpha,0)}(y)$ respectievelijk $p_j(y) = P_j^{(0,\beta)}(y)$.

Het gebruik van de polynoomoperator $P_k(H)$ in de praktijk vereist enige toelichting.

Er zijn drie mogelijkheden om de operatoren $P_k(H)$ in het iteratie-proces (6.51) toe te passen:

- (a) Men schrijft $P_k(H)$ in de gefactoriseerde vorm (6.53) en bepaalt de relaxatieparameters ω_j uit de wortels μ_j van $P_k(\mu)$,

$$(6.66) \quad \omega_j = \frac{1}{1-\mu_j}.$$

- (b) Men schrijft $P_k(H)$ in de vorm

$$(6.67) \quad P_k(H) = 1 + \sum_{j=1}^k a_j (1-H)^j$$

en het iteratie-proces (6.51) in de vorm

$$(6.68) \quad u_k = P_k(H)u_0 + (1-P_k(H))u.$$

Met behulp van de relatie $(1-H)u = Mf$ kan de onbekende u geelimineerd worden.

- (c) Men tracht door middel van recurrente betrekkingen voor de polynomen $P_k(\mu)$, u_{k+1} uit te drukken in voorgaande iteranden.

Methode (a) is geschikt voor polynomen $P_k(\mu)$ waarvan de wortels expliciet bekend zijn zoals b.v. de polynomen gedefinieerd in stelling 6.6. Zoals echter al opgemerkt is bestaat er gevaar voor numerieke instabiliteit voor grote waarden van k . Bovendien moet men van te voren besluiten

hoe groot k is want de relaxatieparameters ω_j hangen van k af. Dit is bezwaarlijk omdat men tijdens het proces pas schattingen kan maken hoe groot de fout is en dus hoe groot k zou moeten zijn.

Methode (b) is direct van toepassing op het polynoom (6.64). Overigens heeft deze methode dezelfde bezwaren als methode (a), namelijk: numerieke instabiliteit (coëfficiënten a_j zijn zeer groot en alternerend als k groot is, zie tabel 6.1) en a priori bepaling van k .

Methode (c) werd door Stiefel toegepast op de polynomen (6.54) en ondervangt de hierboven genoemde bezwaren. Daar tegenover staat echter dat er meer geheugenruimte in de rekenmachine gevraagd wordt. We zullen nu laten zien hoe ook op de polynomen $P_k(\mu)$, welke optimaal in de kwadraat-norm zijn, methode (c) toegepast kan worden.

Daartoe definiëren we

$$(6.69) \quad \begin{cases} \alpha_k = p_k^{-1}(y_0) h_k \sum_{j=0}^k p_j^2(y_0) h_j^{-1}, \\ \beta_k = \alpha_k - p_k(y_0). \end{cases}$$

Stelling 6.9. De polynomen (6.64') voldoen aan de recurrente betrekking

$$(6.70) \quad P_{k+1}(\mu) = [A_k + B_k y(\mu)] P_k(\mu) - [C_k + D_k y(\mu)] P_{k-1}(\mu) + E_k P_{k-2}(\mu).$$

Hierin is

$$\begin{aligned} A_k &= \alpha_{k+1}^{-1} (\beta_{k+1} + \alpha_k a_k), \\ B_k &= \alpha_{k+1}^{-1} \alpha_k b_k, \\ C_k &= \alpha_{k+1}^{-1} (\beta_k a_k + \alpha_{k-1} c_k), \\ D_k &= \alpha_{k+1}^{-1} \beta_k b_k, \\ E_k &= \alpha_{k+1}^{-1} \beta_{k-1} c_k, \end{aligned}$$

en zijn a_k , b_k , c_k de coëfficiënten van de voor $p_k(y)$ geldende recurrente betrekking

$$(6.71) \quad p_{k+1}(y) = (a_k + b_k y) p_k(y) - c_k p_{k-1}(y).$$

ewijs. Het is eenvoudig te verifiëren dat

$$p_k(y(\mu)) = \alpha_k P_k(\mu) - \beta_k P_{k-1}(\mu).$$

substitutie in (6.71) geeft (6.70).

Met behulp van (6.70) kan een vier-niveauschema voor u_{k+1} afgeleid worden:

it

$$v_k = P_k(H) v_0 \text{ en } v_k = u_k - u$$

olgt dat

$$\begin{aligned} u_{k+1} &= (A_k + B_k y(H)) u_k - (C_k + D_k y(H)) u_{k-1} + E_k u_{k-2} \\ &\quad + (1 - A_k - B_k y(H) + C_k + D_k y(H) - E_k) u. \end{aligned}$$

erder volgt uit (6.70) dat

$$1 - A_k - B_k y_0 + C_k + D_k y_0 - E_k = 0, \quad .$$

odat

$$\begin{aligned} u_{k+1} &= (A_k + B_k y(H)) u_k - (C_k + D_k y(H)) u_{k-1} + E_k u_{k-2} \\ &\quad - 2 \frac{B_k - D_k}{b-a} (1-H) u. \end{aligned}$$

it de definitie van B_k en D_k , en uit $(1-H) u = Mf$ volgt tenslotte

$$6.72) \quad u_{k+1} = (A_k + B_k y(H)) u_k - (C_k + D_k y(H)) u_{k-1} + E_k u_{k-2} - 2 \frac{b_k p_k(y_0)}{\alpha_{k+1}(b-a)} Mf.$$

f de equivalente schrijfwijze

$$6.72') \quad \begin{pmatrix} u_{k+1} \\ u_k \\ u_{k-1} \end{pmatrix} = \begin{pmatrix} A_k + B_k y(H) - (C_k + D_k y(H)) - E_k \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} u_k \\ u_{k-1} \\ u_{k-2} \end{pmatrix} - 2 \frac{b_k p_k(y_0)}{\alpha_{k+1}(b-a)} \begin{pmatrix} Mf \\ 0 \\ 0 \end{pmatrix}$$

Als beginapproximatie is nodig de vector $(u_2, u_1, u_0)^T$. Hiervan is alleen u_0 vrij te kiezen; u_1 en u_2 volgen uit (6.68), dus

$$(6.68') \quad u_k = P_k(H) u_0 + [1 - P_k(H)] u, \quad k = 1, 2.$$

Men zal dus $P_1(u)$ en $P_2(u)$ expliciet moeten kennen.

Verder moet men beschikken over de coëfficiënten a_k, b_k, c_k, h_k (zie b.v. Handbook of Mathematical functions, National Bureau of Standards, Applied Mathematics Series 55 voor expliciete formules) en de coëfficiënten $p_k(y_0)$ en α_k . Voor deze laatste gelden de recurrente betrekkingen

$$\begin{cases} p_{k+1}(y_0) = (a_k + b_k y_0) p_k(y_0) - c_k p_{k-1}(y_0), \\ \alpha_{k+1} = \frac{p_k(y_0) h_{k+1}}{p_{k+1}(y_0) h_k} \left(\alpha_k + \frac{p_{k+1}^2(y_0)}{h_{k+1}} \right). \end{cases}$$

Rest nog de numerieke stabiliteit van het schema (6.70) of van het equivalente schema (6.72') te onderzoeken. Voor stabiliteit is het noodzakelijk dat de eigenwaarden λ van de in (6.72') optredende matrix binnen of op de eenheidscirkel liggen. Stel

$$F_k = A_k + B_k y \text{ en } G_k = C_k + D_k y,$$

dan voldoen de eigenwaarden λ aan de vergelijking

$$(6.73) \quad \det \begin{bmatrix} F_k - \lambda & -G_k & E_k \\ 1 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{bmatrix} = 0.$$

Hierin neemt y de waarden $y(\mu_j) = \frac{b+a}{b-a} - 2 \frac{1-\mu_j}{b-a}$, μ_j eigenwaarde van H , aan. Vergelijking (6.73) is te schrijven als

$$(6.73') \quad \lambda^3 - F_k \lambda^2 + G_k \lambda - E_k = 0.$$

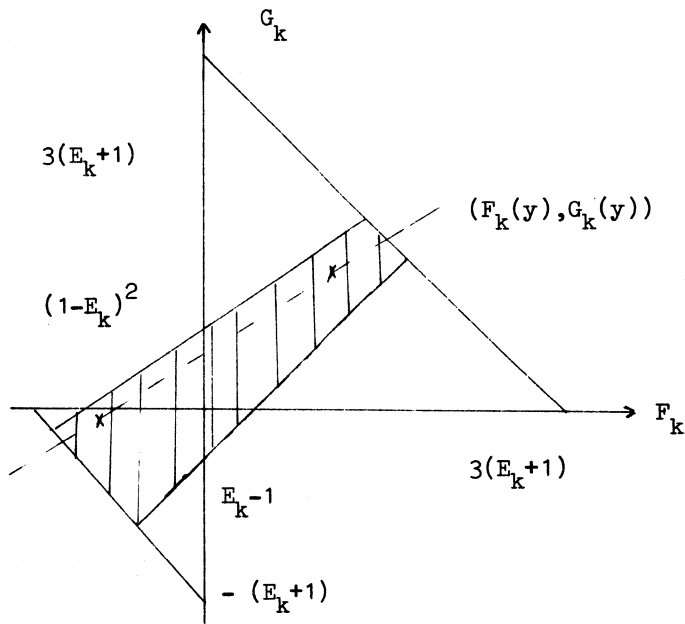
We onderzoeken het geval waarin $|E_k| \leq 1$. Volgens het Hurwitz-criterium

(vergelijk [10], p. 78) is $|\lambda| \leq 1$ als

$$(6.74) \quad \begin{cases} G_k \geq F_k - (1 - E_k), \\ G_k \geq -F_k - (1 + E_k), \\ G_k \leq -F_k + 3(1 + E_k), \\ G_k \leq E_k F_k + (1 - E_k^2). \end{cases}$$

In figuur 6.7 is het stabiliteitsgebied gearceerd aangegeven. We zien hieruit dat het voldoende voor stabiliteit is wanneer alle punten $\{(F_k(y), G_k(y))\}_{y=-1}^1$ binnen dit gebied liggen.

Omdat het stabiliteitsgebied convex is wordt hieraan voldaan wanneer de "uiteinden" van deze "puntenlijn", dus $(A_k \pm B_k, C_k \pm D_k)$ er binnen liggen, met andere woorden schema (6.70) is stabiel wanneer het voor $y = \pm 1$ stabiel is.



6.7. Stabiliteitsgebied van (6.70)

Dit geeft in elk geval al een eenvoudig criterium om te verifiëren tijdens de berekeningen. Een analyse om a priori iets over de stabiliteit te zeggen is nog niet afgerond.

Voor experimenten met het hier beschreven iteratieproces zij verwezen naar Jansen [23].

Methode van Richardson voor het geval van complexe eigenwaarden

Het polynoom $P_k(\mu)$ gedefinieerd in stelling 6.6 is optimaal in de maximum-norm voor reële eigenwaarden μ . Men kan zich afvragen of er een dergelijke optimale polynoom te vinden is voor complexe eigenwaarden μ . In zijn algemeenheid is deze vraag moeilijk te beantwoorden; daarom vereenvoudigen we het probleem door niet $P_k(\mu)$ te beschouwen, maar $P_k(\rho e^{i\phi})$, waarin ρ een constante is met de eigenschap dat de eigenwaarden μ_j van H binnen of op de cirkel $|\mu| = \rho$ liggen. Het polynoom $P_k(\rho e^{i\phi})$ moet dus in de een of andere norm geminimaliseerd worden. We herinneren er in dit verband aan dat de eigenwaarden μ_j van het Frankel-proces op een cirkel met straal

$$(6.43') \quad \rho = \frac{\sqrt{b} - \sqrt{a}}{\sqrt{b} + \sqrt{a}}$$

liggen.

Stelling 6.9. Het polynoom

$$(6.75) \quad P_k(\mu) = \mu^k$$

heeft van alle k^{de} graads polynomen in μ die voldoen aan $P_k(1) = 1$ de kleinste maximum-norm over de cirkel $|\mu| = \rho \leq 1$.

Bewijs Het bewijs van deze stelling werd gegeven door E.H. Zarantonello en is te vinden in een artikel van Varga. A comparison of the successive overrelaxation method and semi-iterative methods using Chebyshev polynomials. J. Soc. Indust. Appl. Math., 5(1957), p. 39.

Uit deze stelling volgt dat in het geval van complexe eigenwaarden het stationaire proces in de zin van de maximum-norm optimaal is.

Stelling 6.10. Het polynoom

$$(6.76) \quad P_k(\mu) = \rho^{2k} \frac{1-\rho^2}{1-\rho^{2(k+1)}} \sum_{j=0}^k (\mu/\rho^2)^j$$

heeft van alle k^{de} graads polynomen in μ die voldoen aan $P_k(1) = 1$ de kleinste kwadraat-norm over de cirkel $|\mu| = \rho$.

Bewijs Stel

$$(6.77) \quad P_k(\mu) = \sum_{j=0}^k a_j \mu^j, \quad \sum_{j=0}^k a_j = 1.$$

Het is eenvoudig te verifiëren dat

$$(6.78) \quad \|P_k(\mu)\|_2^2 = \int_0^{2\pi} |P_k(\rho e^{i\phi})|^2 d\phi = 2\pi \sum_{j=0}^k |a_j|^2 \rho^{2j}.$$

Deze som moet geminimaliseerd worden onder de nevenvoorwaarde dat $\sum_{j=0}^k a_j = 1$.

Dit geeft de volgende vergelijkingen voor de coëfficiënten a_j en de Lagrange-parameter λ :

$$\frac{\partial}{\partial a_j} \left[\sum_{l=0}^k (|a_l|^2 \rho^{2l} + \lambda a_j) - \lambda \right] = 0, \quad j = 0, \dots, k,$$

$$\sum_{j=0}^k a_j = 1.$$

Hieraan voldoen de coëfficiënten

$$a_j = \rho^{2k} \frac{1-\rho^2}{1-\rho^{2(k+1)}} \rho^{-2j}.$$

Substitutie in (6.77) geeft uitdrukking (6.76).

Uit (6.78) volgt direct de waarde van $\|P_k(\mu)\|_2$, namelijk

$$(6.79) \quad \|P_k(\mu)\|_2 = \sqrt{2\pi} \rho^{2k} \frac{1-\rho^2}{1-\rho^{2(k+1)}} \left[\sum_{j=0}^k \rho^{-2j} \right]^{\frac{1}{2}} = \sqrt{2\pi} \rho^k \left[\frac{1-\rho^2}{1-\rho^{2(k+1)}} \right]^{\frac{1}{2}}.$$

Hieruit kan men het gedrag voor grote waarden van k en ρ -waarden in de buurt van 1 afleiden. Er geldt:

$$(6.80) \quad \begin{aligned} & \sim \sqrt{2\pi} \rho^k (k+1)^{-\frac{1}{2}} \quad \text{als } k(1-\rho) \rightarrow 0, \\ \|P_k(\mu)\|_2 & \sim \sqrt{2\pi} \rho^k (1-e^{-c})^{-\frac{1}{2}} (1-\rho^2)^{\frac{1}{2}} \quad \text{als } (k+1)(1-\rho^2) = c, \quad 0 < c < \infty, \\ & \sim \sqrt{2\pi} \rho^k (1-\rho^2)^{\frac{1}{2}} \quad \text{als } k(1-\rho) \rightarrow \infty. \end{aligned}$$

Vergelijken we dit met de kwadraat-norm van het polynoom $P_k(\mu) = \mu^k$,

$$\|\mu^k\|_2 = \sqrt{2\pi} \rho^k,$$

dan zien we dat de polynomen (6.76) in de kwadraat-norm voor grote waarden van k en/of ρ -waarden in de buurt van 1 een veel grotere demping geven dan de polynomen $P_k(\mu) = \mu^k$.

Uiteraard is de maximum-norm van de polynomen (6.76) groter dan $\|\mu^k\|_\infty = \rho^k$. Er geldt volgens (6.76)

$$|P_k(\mu)| = [P_k(\mu) P_k(\bar{\mu})]^{1/2} = \rho^{2k} \frac{1-\rho^2}{1-\rho^{2(k+1)}} \left[\frac{1-(\mu/\rho^2)^{k+1}}{1-\mu/\rho^2} \frac{1-(\bar{\mu}/\rho^2)^{k+1}}{1-\bar{\mu}/\rho^2} \right]^{1/2}$$

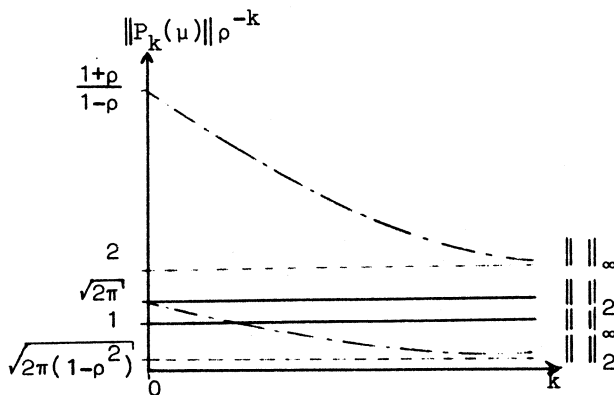
Substitutie van $\mu = \rho e^{i\phi}$ geeft na uitwerking

$$\begin{aligned} |P_k(\mu)| &= \rho^k \frac{1-\rho^2}{1-\rho^{2(k+1)}} \left[\frac{1-2\rho^{k+1} \cos(k+1)\phi + \rho^{2(k+1)}}{1-2\rho \cos \phi + \rho^2} \right]^{1/2} \\ &< \rho^k \frac{1-\rho^2}{1-\rho^{2(k+1)}} \sqrt{\frac{(\rho^{k+1}+1)^2}{(\rho-1)^2}} = \rho^k \frac{1+\rho}{1-\rho^{k+1}}, \end{aligned}$$

dus

$$(6.81) \quad \|P_k(\mu)\|_\infty < \rho^k \frac{1+\rho}{1-\rho^{k+1}}.$$

In figuur 6.8 is het gedrag van $\|P_k(\mu)\|$ als functie van k voor een vaste $\rho \sim 1$ gegeven.



Figuur 6.8

— polynomen (6.75)
 - - - polynomen (6.76)

Literatuur

- [1] Bodewig, E. Matrix calculus.
North-Holland Publishing Company, Amsterdam,
1956.
- [2] Coolen, T.M.T., On the acceleration of Richardson's method IV.
Houwen, P.J. van der A non-symmetrical case.
Report TW 109, Math. Centre, Amsterdam, 1968.
- [3] Flanders, D.A., Numerical determination of fundamental modes.
Shortley, G. J. Appl. Phys. 21, 1326-1332, 1950.
- [4] Forsythe, G.E. Finite difference methods for partial dif-
Wasow, W.R. ferential equations.
John Wiley & Sons, Inc., New York, 1960.
- [5] Frank, W. Solution of linear systems by Richardson's
method.
J. Assoc. Comput. Mach. 7, 274-286, 1960.
- [6] Frankel, S.P. Convergence rates of iterative treatments
of partial differential equations.
Math. Tables Aids Comput., 4, 65-75.
- [7] Gerschgorin, S. Fehlerabschätzung für das Differenzenver-
fahren zur Lösung partieller Differential-
gleichungen.
Z. Angew. Math. Mech. 10, 373-382, 1930.
- [8] Garabedian, P.R. Partial differential equations.
Wiley, New York, 1964.
- [9] Garabedian, P.R. Estimates of the relaxation factor for small
mesh size.
Math. Tables Aids Comput., 10, 183-185, 1956.
- [10] Houwen, P.J. van der Finite difference methods.
MC tracts 20, Math. Centre, Amsterdam, 1968.
- [11] Houwen, P.J. van der On the acceleration of Richardson's method II.
Numerical aspects.
Report TW 107, Math. Centre, Amsterdam, 1967.

- [12] Houwen, P.J. van der On the acceleration of Richardson's method III
Applications.
Report TW 108, Math. Centre, Amsterdam, 1967.
- [13] Houwen, P.J. van der A method for solving elliptic difference
equations.
Report MR 104, Math. Centre, Amsterdam, 1969
- [14] Richardson, L.F. The approximate arithmetical solution by
finite differences of physical problems in-
volving differential equations.
Philos. Trans. Roy. Soc., London, A210,
307-357, 1910.
- [15] Sheldon, J.W. On the numerical solution of elliptic
difference equations.
Math. Tables. Aids Comput., 9, 101-112, 1955.
- [16] Shortley, G.H., The numerical solution of Laplace equation.
Weller, R. J. Appl. Phys., 9, 334-348, 1938.
- [17] Varga, R.S. Matrix iterative analysis.
Prentice-Hall, Inc., Englewood Cliffs,
New Jersey, 1962.
- [18] Vooren, A.J. van de On the 9-point difference formula for
Vliegenthart, A.C. Laplace's equation.
Report TW 42, Math. Inst., University
Groningen, 1967.
- [19] Wachspress, E.L. Iterative solution of Elliptic systems.
Prentice-Hall, Inc., Englewood Cliffs,
New Jersey, 1966.
- [20] Young, D.M. On Richardson's method for solving linear
systems with positive definite matrices.
J. Math. Phys. 32, 243-255, 1953.
- [21] Young, D.M. Iterative methods for solving partial dif-
ferential difference equations of elliptic
type.
Trans. Amer. Soc. 76, 92-111, 1954.

[22] Young, D.M.,
Warlick, C.H.

On the use of Richardson's method for the
numerical solution of Laplace's equation on
Ordvac.

S.R.L. Memo Report 707, Aberdeen Proving
Ground, 1953.

[23] Jansen, M.

Semi-iteratieve methoden voor de oplossing
van de vergelijking van Laplace.

Technische notie. Math. Centrum.

Amsterdam (verschijnt in 1970).

7. ONJUIST GESTELDE PROBLEMEN

7.1 Inleiding en voorbeelden

Het is bekend dat veel begin- en randwaardeproblemen voor partiële differentiaalvergelijkingen voortkomen uit problemen van de fysica, mechanica enz.

Randwaardeproblemen blijken typisch bij elliptische differentiaalvergelijkingen voor te komen, terwijl beginwaardeproblemen en begin-randwaardeproblemen optreden in verband met hyperbolische en parabolische vergelijkingen.

Men heeft lange tijd gesteld, dat een mathematisch probleem dat met een fysische werkelijkheid correspondeert, moet voldoen aan de volgende drie voorwaarden:

- (i) de oplossing moet bestaan;
- (ii) de oplossing moet eenduidig bepaald zijn;
- (iii) de oplossing moet in een of andere zin continu afhangen van begin- en/of randvoorwaarden, of algemener: continu afhangen van alle gegevens ("data").

De eerste voorwaarde drukt uit dat er niet te veel voorwaarden aan de oplossing mogen worden opgelegd, de tweede dat het er niet te weinig mogen zijn. De derde eis lijkt noodzakelijk als van de mathematische formulering verwacht wordt dat deze meetbare verschijnselen in de natuur beschrijft. De gegevens door meting verkregen zijn nooit helemaal exact, en men wil nu, dat een kleine fout in de gegevens slechts een kleine fout in de oplossing van het begin-randwaardeprobleem teweegbrengt. Deze "stabiliteits"-eis lijkt niet alleen essentieel voor zinvolle problemen in de mathematische fysica, maar is, zodra we met numerieke procedures de problemen proberen op te lossen, onontbeerlijk.

Een probleem, dat voldoet aan de drie eisen, zullen we een goed (of juist) gesteld probleem noemen (Engels: properly posed, well posed, correctly set).

Het hyperbolische beginwaardeprobleem

$$\begin{cases} u_{xx} - u_{tt} = 0 & -\infty < x < \infty, \quad t > 0, \\ u(x,0) = \phi(x) & -\infty < x < \infty, \\ u_t(x,0) = \psi(x) & -\infty < x < \infty, \end{cases}$$

heeft, zoals bekend, als oplossing

$$u(x,t) = \frac{1}{2} [\phi(x+t) + \phi(x-t)] + \int_{x-t}^{x+t} \psi(\tau) d\tau$$

en is kennelijk goed gesteld. Bekijken we echter het minder bekende hyperbolische probleem

$$\begin{cases} u_{xy} = 0 & 0 < x < a, \quad 0 < y < b, \\ u(x,0) = f(x) & 0 \leq x \leq a, \\ u(0,y) = g(y) & 0 \leq y \leq b, \quad f(0) = g(0), \end{cases}$$

dan is gemakkelijk na te gaan (zie bijv. [6] p. 133 e.v.) dat u hierdoor reeds in de gehele rechthoek eenduidig vastligt, zodat het in het algemeen niet zal lukken een oplosbaar Dirichlet-probleem voor de hyperbolische vergelijking op te stellen. Het probleem is overbepaald.

Het Dirichlet-probleem voor een elliptische vergelijking in een begrensde gebied Ω is daarentegen goed gesteld. Existentie en uniciteit zijn in dit colloquium uitvoerig aan de orde gekomen (zie hoofdstukken 3 en 5 van het eerste deel van de syllabus [1]) en de continue afhankelijkheid van de randvoorwaarden volgt onmiddellijk uit het maximumprincipe: als twee oplossingen op de rand slechts ϵ verschillen, dan ook in het inwendige.

Het Cauchyprobleem voor elliptische differentiaalvergelijkingen is in het algemeen onjuist gesteld. We geven hiervan twee bekende voorbeelden.

Voorbeeld 7.1. Het probleem

$$\begin{cases} u_{xx} + u_{yy} = 0 & x > 0, \\ u(0,y) = 0 & -a < y < a, \\ u_x(0,y) = f(y) & -a < y < a, \end{cases}$$

heeft geen oplossing als $f(y)$ niet analytisch is. Dit is dus reeds het geval als $f(y) = |y|$. Merk op dat de stelling van Cauchy-Kowalewski een oplossing garandeert, mits $f(y)$ analytisch is.

Bewijs: Volgens het reflectieprincipe is elke functie u die harmonisch is in een halve bol D^+ om 0 in het rechter halfvlak, en die op de diameter 0 is, analytisch voort te zetten in het spiegelbeeld D^- van D^+ , m.b.v. de regel $u(-x, y) = -u(x, y)$. Maar dit zou betekenen dat u analytisch is op de diameter, en daarmee u_x eveneens.

Voorbeeld 7.2. (Hadamard [5]). Beschouw de vergelijking

$$u_{xx} + u_{yy} = 0 \quad x > 0,$$

tezamen met de Cauchydata

$$(7.1) \quad \begin{aligned} u(0, y) &= 0, \\ u_x(0, y) &= \frac{1}{n} \sin ny \quad n > 0. \end{aligned}$$

De methode van scheiding van variabelen suggereert de oplossing

$$(7.2) \quad u(x, y) = \frac{1}{n} \sin nx \sin ny.$$

Voor $n \rightarrow \infty$ gaan de Cauchydata (7.1) uniform naar 0, maar de oplossing (7.2) oscilleert voor $x \neq 0$ tussen grenzen die met n onbepaald toenemen. Aan de derde eis is dus niet voldaan.

Hadamard kon in zijn tijd zeggen dat elk begin- en randwaardeprobleem dat correspondeert met een fysische werkelijkheid, juist gesteld moet zijn. Dat is nu niet meer vol te houden. Laten we hier twee voorbeelden geven van fysisch zinvolle problemen die onjuist gesteld zijn.

Voorbeeld 7.3. Beschouw het probleem van de warmtegeleiding in een staaf van lengte π . Laat u de temperatuurverdeling voorstellen, t de tijd en x de plaatscoördinaat. We zijn nu niet geïnteresseerd in het bekende, juist gestelde, probleem de eindtemperatuur bij gegeven randvoorwaarden uit de beginsituatie te bepalen, t.w.

$$(7.3) \quad \begin{cases} u_t - u_{xx} = 0 & 0 < x < \pi, \quad t > 0, \\ u(x, 0) = \xi(x) & 0 < x < \pi \text{ (beginvoorwaarde)}, \\ u(\pi, t) = u(0, t) = 0 & t > 0 \text{ (randvoorwaarden)}. \end{cases}$$

In plaats daarvan willen we de temperatuurverdeling bepalen op een (vroeger) tijdstip $t < T$ als deze bekend is op een (later) tijdstip $t = T > 0$; mathematisch weergegeven:

$$(7.4) \quad \begin{cases} u_t - u_{xx} = 0 & 0 < x < \pi, \quad 0 < t < T, \\ u(x, T) = \chi(x) & 0 < x < \pi \text{ ("eind"-voorwaarde)}, \\ u(\pi, t) = u(0, t) = 0 & t > 0 \text{ (randvoorwaarden)}. \end{cases}$$

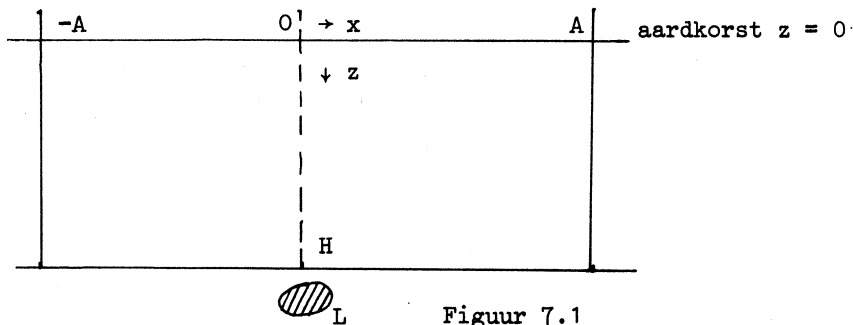
Het proces van warmtegeleiding, waarvan (7.3) het mathematische model is, is een irreversibel fysisch proces. Het verbaast ons dan ook niet dat (7.4) onjuist gesteld is. En toch is de vraag die gesteld is fysisch zinvol.

Voorbeeld 7.4. Het grootste deel van het zwaartekrachtsveld van de aarde wordt veroorzaakt door het inwendige van de aarde en heeft weinig te maken met de bovenste 5 km van de aardkorst. Afwijkingen van de gemiddelde structuur van de aardkorst, zoals scheuren, verschuivingen en ertsaders, hebben echter kleine fluctuaties in de versnelling van de zwaartekracht g ten gevolge, die meetbaar zijn. Op grond van deze fluctuaties wil men zich een beeld vormen waar deze afwijkingen zich bevinden.

Stellen we de gravitatie potentiaal u , dan is u te schrijven als $u_{\text{gemiddeld}} + \bar{u}$, waarin $\bar{u}$ de excessopotential ten gevolge van de afwijking (bijv. een lichaam van grote dichtheid) in de aardkorst voorstelt. Evenzo $g = g_{\text{gemiddeld}} + \bar{g}$, waarin $\bar{g}$ het effect van de afwijking in de aardkorst is op g . We hebben hier te maken met een "invers" potentiaalprobleem. Een van de vragen die zich hierbij voordoen is: hoe zet ik een harmonische functie voort als zijn Cauchydata bekend zijn?

Laten we aannemen dat tot op diepte H zich geen extra lichamen bevinden die een effect op de gravitatie potentiaal hebben, en laten we aannemen dat de verstoring van het gravitatieveld slechts door één stervormig lichaam dieper dan H wordt veroorzaakt, hetgeen tot uitdrukking komt

in de Cauchydata, dan kunnen we het hierboven genoemde probleem mathematisch in gesimplificeerde vorm als volgt weergeven:



Zoek de functie $\bar{u}(x,y,z)$, die voldoet aan

$$(7.5) \quad \left\{ \begin{array}{l} \bar{u}_{xx} + \bar{u}_{yy} + \bar{u}_{zz} = 0 \quad 0 < z < H, \\ \bar{u}(x,y,0) = \phi(x,y) \\ \frac{\partial \bar{u}}{\partial z}(x,y,0) = \bar{g}(x,y,0) = \psi(x,y) \end{array} \right\} x^2 + y^2 \leq A, \\ \bar{u}(x,y,z) = 0 \text{ voor } x^2 + y^2 = A^2.$$

De laatste voorwaarde drukt uit dat bij benadering verweg het effect van het lichaam L te verwaarlozen is. Dit is een typies onjuist gesteld probleem. Door $\bar{u}(x,y,z)$ naar beneden voort te zetten (boven de aardkorst is hij in principe bekend) kunnen we ons een beeld vormen waar (voor welke x,y) ongeveer het lichaam ligt, als we additionele aannames hebben over de diepte en dichtheid van het lichaam L. Meer over dit soort zaken is bijv. te vinden in [4], p. 189 e.v.

Randwaardeproblemen voor elliptische vergelijkingen, waarbij op een deel van de rand "te veel" voorwaarden gegeven zijn, en op het resterende deel "te weinig" of geen, zijn typische voorbeelden van onjuist gestelde elliptische randwaardeproblemen. Men heeft met een in zekere zin "invers" probleem te maken, als men de randvoorwaarden op een deel van de rand wil bepalen, terwijl de differentiaalvergelijking in het gebied en de Cauchydata op het resterende deel van de rand bekend zijn.

Het hierboven geschetste voorbeeld is zo'n probleem.

We zullen ons in dit hoofdstuk bezig houden met oplossingsmethoden voor onjuist gestelde elliptische problemen, eerst met analytische methoden, en daarna met numerieke.

7.2 Regularisatie van onjuist gestelde problemen

We formuleren eerst de begrippen onjuist gesteld en regularisatie abstract in termen van metrische ruimten, en gaan daarna na hoe onjuist gestelde elliptische randwaardeproblemen geregulariseerd kunnen worden.

Laten U en F volledige metrische functieruimten zijn, en laat A een (niet noodzakelijk lineaire) operator op U zijn met waarden in F . Voor functies $u \in U$ en $f \in F$ bekijken we de vergelijking

$$(7.6) \quad Au = f.$$

Het probleem (7.6) is goed gesteld als

- (i) met elke $f \in F$ een oplossing $u \in U$ correspondeert;
- (ii) de oplossing u binnen U uniek is bij gegeven $f \in F$;
- (iii) u continu afhangt van f in de metrieken van U en F .

Anders geformuleerd: (7.6) is een goed gesteld probleem als er een continue operator B op F bestaat, die de inverse is van A .

Laten we nu aannemen dat de uniciteit wel gewaarborgd is, maar dat aan eis (i) of (iii) niet voldaan is. Voor dergelijke onjuist gestelde problemen voerde Tychonov ([10], [11], [12]) het begrip regularisator als volgt in. Een regularisator voor de vergelijking (7.6) is een familie operatoren $B_\tau: F \rightarrow U$, die voldoen aan

- (i) voor iedere $\tau > 0$ is B_τ gedefinieerd en continu op de gehele ruimte F ;
- (ii) voor iedere $u \in U$ geldt (in de norm van U)

$$(7.7) \quad \lim_{\tau \rightarrow 0} B_\tau Au = u.$$

We zullen het probleem (7.6) regulariseerbaar noemen, als er tenminste één regularisator bestaat. Een goed gesteld probleem is regulariseerbaar: kies $B_\tau = A^{-1}$ voor alle τ .

Stel nu dat we reeds weten dat voor een zekere rechterlid f een oplossing u van (7.6) bestaat. De regularisator geeft ons de mogelijkheid een oplossing met een zekere gegarandeerde nauwkeurigheid te construeren met behulp van een rechterlid f_δ dat f benadert in de zin

$$(7.8) \quad \|f_\delta - f\| \leq \delta.$$

Hier en later in deze sectie zal $\| \cdot \|$ de supremum-norm voorstellen. We schrijven

$$u_{\tau\delta} = B_\tau f_\delta$$

en

$$u_\tau = B_\tau A u = B_\tau f.$$

Merk op dat u als oplossing van $Au = f$ bestaat, maar dat de oplossing van $Au = f_\delta$ niet hoeft te bestaan als functie van τ genomen. Er volgt

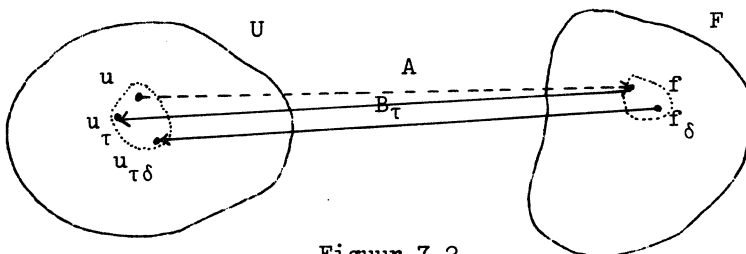
$$(7.9) \quad \|u - u_{\tau\delta}\| \leq \|u - u_\tau\| + \|u_\tau - u_{\tau\delta}\|$$

We stellen

$$\|u - u_\tau\| \leq \varepsilon_1(\tau),$$

$$\|u_\tau - u_{\tau\delta}\| \leq \varepsilon_2(\tau, \delta).$$

Wegens (7.7) kan τ zo klein worden gekozen dat $\varepsilon_1(\tau)$ willekeurig klein is, en wegens continuïteit van B_τ is daarna (voor vaste τ) $\varepsilon_2(\tau, \delta)$ willekeurig klein te krijgen voor voldoende kleine δ . We hebben dus δ als functie van τ genomen.



Figuur 7.2

Het kan ook anders om : $\varepsilon_1(\tau)$ is op te vatten als de mate van convergentie van $B_\tau Au$ tot u , $\varepsilon_2(\tau, \delta)$ als de mate van continuïteit van B_τ in het punt Au . Wanneer nu, zoals meestal, δ gegeven is (dit correspondeert met een niet precies bekende randvoorwaarde), dan gaat het er om $\tau = \tau(\delta)$ zo te kiezen dat het linkerlid van (7.9) zo klein mogelijk wordt. Daar vaak $\varepsilon_2(\tau, \delta)$ bij vaste δ groot wordt voor $\tau \rightarrow 0$, en $\varepsilon_1(\tau)$ juist klein, ligt het voor de hand $\tau = \tau(\delta)$ zo te kiezen dat

$$(7.10) \quad c_1 \varepsilon_1(\tau(\delta)) \leq \varepsilon_2(\tau(\delta), \delta) \leq c_2 \varepsilon_1(\tau(\delta)) \quad , \quad c_2 \geq c_1 > 0,$$

waaruit dan volgt

$$(7.9') \quad \|u - u_{\tau\delta}\| \leq (1+c_2) \varepsilon_1(\tau(\delta)).$$

We willen nu bewijzen dat $u \rightarrow u_{\tau\delta}$ (in de maximum-norm) als $\delta \rightarrow 0$.

Lemma 7.1. Laat gegeven zijn dat

- (a) $\varepsilon_1(\tau)$ een niet-negatieve continue sterk monotoon stijgende functie is voor $0 \leq \tau \leq a$;
- (b) $\varepsilon_2(\tau, \delta)$ een niet-negatieve functie op de rechthoek $[0, a; 0, b]$ is, die continu en voor vaste $\delta \rightarrow 0$ (zwak) monotoon stijgend is in δ ;
- (c) $\varepsilon_1(0) = 0$ en $\varepsilon_2(\tau, 0) = 0$ voor alle $\tau > 0$;
- (d) de functie $\varepsilon_2(\tau, \delta)$ begrensd is op iedere rechthoek $0 \leq \delta \leq b$, $\eta \leq \tau \leq a$, waarin $\eta > 0$;
- (e) dat er een constante $c_1 > 0$ bestaat, zodanig dat

$$\sup_{0 < \tau \leq a} \frac{\varepsilon_2(\tau, \delta)}{\varepsilon_1(\tau)} > c_1 \quad \text{voor } 0 \leq \delta \leq b.$$

Dan geldt: er bestaat een constante $c_2 \geq c_1$, zodanig dat bij iedere δ een $\tau = \tau(\delta)$ bestaat die voldoet aan

$$(7.10) \quad c_1 \varepsilon_1(\tau(\delta)) \leq \varepsilon_2(\tau(\delta), \delta) \leq c_2 \varepsilon_1(\tau(\delta)).$$

Voor deze $\tau = \tau(\delta)$ geldt dan bovendien $\lim_{\delta \rightarrow 0} \tau(\delta) = 0$

Bewijs: De laatste eis houdt in dat $\lim_{\delta \rightarrow 0} \varepsilon(\tau, \delta)$ niet uniform geldt; en

dat is bij een onjuist gesteld probleem dan ook niet te verwachten: immers " $\varepsilon(0, \delta)$ " is zoiets als de norm van " A^{-1} ". Een constante c_2 met bovenstaande eigenschappen is gemakkelijk te vinden. Neem bijv. $c_2 = \inf_{0 < \tau \leq a} \varepsilon_2(\tau, b)$ als dit groter is dan c_1 .

Laat η willekeurig, > 0 zijn. De functie

$$\sup_{\eta \leq \tau \leq a} \varepsilon_2(\tau, \delta) = f(\delta)$$

is monotoon niet-dalend voor $0 \leq \delta \leq b$, zodat $\lim_{\delta \rightarrow 0} f(\delta)$ bestaat. De punts-
gewijze limiet

$$\lim_{\delta \rightarrow 0} \varepsilon_2(\tau, \delta) = \varepsilon_2(\tau, 0) = 0$$

is dan op $\eta \leq \tau \leq a$ uniform, want $\varepsilon_2(\tau, \delta)$ wordt door een van τ onafhankelijke functie gemajoreerd. Het gevolg is dat $\lim_{\delta \rightarrow 0} f(\delta) = 0$. Er bestaat dus een δ_0 zo dat voor $\delta < \delta_0$ geldt $f(\delta) < c_1 \varepsilon_1(\eta)$. Bepaal voor zo'n δ de $\tau = \tau(\delta)$ die voldoet aan (7.10). Stel dat $\tau(\delta) > \eta$, dan zou gelden

$$c_1 \varepsilon_1(\eta) < c_1 \varepsilon_2(\tau(\delta)) \leq \varepsilon_2(\tau(\delta), \delta) \leq \sup_{\eta \leq \tau \leq a} \varepsilon_2(\tau, \delta) < c_1 \varepsilon_1(\eta)$$

hetgeen tot een tegenspraak leidt. Dus $\tau(\delta) < \eta$, hetgeen te bewijzen was.

Gevolg. De functies $\varepsilon_1(\tau)$, $\varepsilon_2(\tau, \delta)$ en $\tau = \tau(\delta)$ voorkomende in (7.9) en (7.10) voldoen aan de voorwaarden van het lemma, als $\|u - u_\tau\| \neq 0$ voor alle $\tau > 0$. Is $\|u - u_\tau\| = 0$ voor alle τ , dan kunnen we een monotone $\varepsilon_1(\tau)$ kiezen met $\|u - u_\tau\| \leq \varepsilon_1(\tau)$. Het lemma kunnen we dus toepassen, en we vinden

$$\lim_{\delta \rightarrow 0} \tau(\delta) = 0$$

waaruit volgt

$$\lim_{\delta \rightarrow 0} \varepsilon_1(\tau(\delta)) = 0,$$

zodat we uit (7.9') kunnen concluderen dat

$$(7.11) \quad \lim_{\delta \rightarrow 0} \|u_{\tau(\delta)} - u\| = 0 \quad (\tau = \tau(\delta)).$$

Of anders geformuleerd,

$$(7.11') \quad \lim_{\delta \rightarrow 0} B_{\tau(\delta)} f_\delta = u.$$

Natuurlijk geldt (7.11') voor elke keuze van een weg in het τ - δ -vlak waarlangs τ en δ naar 0 gaan, waarvoor geldt $\lim_{\delta \rightarrow 0} \tau(\delta) = \lim_{\delta \rightarrow 0} \varepsilon_1(\tau(\delta)) = 0$.
 $\lim_{\delta \rightarrow 0} \varepsilon_2(\tau(\delta), \delta) = 0$

Het komt erop neer dat als we $\tau = \tau(\delta)$ op een slimme manier naar nul laten gaan, we erin slagen de benaderende oplossing te laten convergeren naar de echte oplossing. Om dit nu te kunnen toepassen op Cauchy-problemen voor tweede orde elliptische differentiaalvergelijkingen in een gebied Ω , modificeren we eerst de definitie van regularisator als volgt:

- (i') Voor iedere $\tau > 0$ is de operator $B_\tau : \Phi \times \Psi \rightarrow U$ op het cartesische produkt $\Phi \times \Psi$ gedefinieerd, en continu. Hierin zijn Φ en Ψ klassen van data voor u en $\partial u / \partial \nu$ op de rand $\partial\Omega$ of een deel $\partial_1\Omega$ ervan, en U de oplossingsruimte. We nemen op Φ en Ψ de supremumnorm, bijv.

$$\|\phi\| = \sup_{x \in \partial_1\Omega} \phi(x).$$

- (ii') Beschouw alle ϕ en ψ , waarvoor een oplossing u van het Cauchy-probleem bestaat (we weten dan reeds dat het probleem dan eenduidig oplosbaar is); voor deze ϕ en ψ moet dan gelden

$$(7.7') \quad \lim_{\tau \rightarrow 0} B_\tau(\phi, \psi) = u \quad (\text{in de norm van } U).$$

Het Cauchyprobleem voor de vergelijking van Laplace is onjuist gesteld, omdat de oplossing instabiel is met betrekking tot kleine veranderingen in de randgegevens. We laten aan een voorbeeld zien hoe zulke problemen aan te pakken zijn.

Voorbeeld 7.5. Het Cauchyprobleem voor de potentiaalvergelijking in een oneindige strook (zie ook [7]):

$$(7.12) \quad \begin{cases} \Omega = \{(x, y) \mid -\infty < x < \infty, \quad 0 < y < b\} \\ u_{xx} + u_{yy} = 0 \\ u \in C^0(\Omega) \\ u(x, 0) = f(x) \\ u_y(x, 0) = g(x), \end{cases}$$

waarin $f(x)$ en $g(x)$ gegeven functies zijn van de reële variabele x . We zullen aannemen dat we reeds weten dat het probleem een eenduidig bepaalde oplossing heeft (wat bijv. het geval is als f en g begrensde analytische functies zijn).

Fouriertransformatie naar x (eventueel in distributionele zin)

$$\int_{-\infty}^{\infty} e^{-ixs} f(x) dx = \mathcal{F}[f] = \hat{f}(s), \text{ enz.,}$$

voert het probleem (7.12) over in

$$(7.13) \quad \hat{u}(s,y) = \hat{f}(s) \cosh sy + \hat{g}(s) \frac{\sinh sy}{s}$$

We vervangen nu $\hat{u}(s,y)$ door

$$(7.14) \quad U_{\tau}(s,y) = e^{-\tau^2 s^2} \hat{u}(s,y)$$

De invers-Fouriergetransformeerde van $e^{-\tau^2 s^2}$ is

$$(7.15) \quad w_{\tau}(x) = \mathcal{F}^{-1} \left[e^{-\tau^2 s^2} \right] = \frac{1}{2\pi^{\frac{1}{2}}\tau} \exp -\frac{x^2}{4\tau^2}$$

We bepalen nu

$$(7.16) \quad u_{\tau}(x,y) = \mathcal{F}^{-1} [U_{\tau}(s,y)] = f(x) * \mathcal{F}^{-1} [\cosh sy e^{-\tau^2 s^2}] + g(x) * \mathcal{F}^{-1} \left[\frac{\sinh sy}{s} e^{-\tau^2 s^2} \right].$$

We vinden

$$(7.17) \quad R_{\tau}(x,y) = \mathcal{F}^{-1} [\cosh sy e^{-\tau^2 s^2}] = \frac{1}{2\tau\pi^{\frac{1}{2}}} \exp\left(\frac{y^2 - x^2}{4\tau^2}\right) \cos \frac{xy}{2\tau^2},$$

en

$$(7.18) \quad T_{\tau}(x,y) = \mathcal{F}^{-1} \left[\frac{\sinh sy}{s} e^{-\tau^2 s^2} \right] = \frac{1}{2\tau\pi^{\frac{1}{2}}} \exp\left(-\frac{x^2}{4\tau^2}\right) \int_0^{y/2\tau} e^{t^2} \cos \frac{tx}{\tau} dt$$

zodat (7.16) in de vorm

$$(7.16') \quad u_{\tau}(x,y) = R_{\tau}(x,y) * f(x) + T_{\tau}(x,y) * g(x)$$

geschreven kan worden, waarin $*$ het convolutieproduct voorstelt.

We vragen naar de oplossing op de lijn $y = b$. We definiëren nu voor alle $\tau > 0$

$$B_{\tau} : F \times G \rightarrow U,$$

waarin F en G resp. ruimten van data f en g zijn, bestaande uit de begrensde analytische functies op $-\infty < x < \infty$ met maximumnorm, en U de ruimte van functies $u(x,b)$ eveneens met maximumnorm, als volgt:

$$(7.19) \quad u_{\tau} = B_{\tau}(f,g) = R_{\tau}(x,b) * f(x) + T_{\tau}(x,b) * g(x).$$

We bewijzen dat de familie van operatoren $(B_\tau)_{\tau>0}$ een regularisator voor (7.12) vormt. De lineariteit en continuïteit volgen onmiddellijk uit de definitie. We moeten eis (ii') aantonen. Uit (7.14) volgt

$$u_\tau(x,b) = w_\tau(x) * u(x,b),$$

waaruit, in verband met

$$\int_{-\infty}^{\infty} w_\tau(t) dt = 1$$

volgt

$$\begin{aligned} (7.20) \quad |u_\tau(x,b) - u(x,b)| &= \left| \int_{-\infty}^{\infty} w_\tau(t) [u(x-t,b) - u(x,b)] dt \right| \\ &\leq \left| \left\{ \int_{-\infty}^{-N} + \int_N^{\infty} \right\} w_\tau(t) [u(x-t,b) - u(x,b)] dt \right| + \\ &\quad + \left| \int_{-N}^N w_\tau(t) [u(x-t,b) - u(x,b)] dt \right|. \end{aligned}$$

Uit het karakter van de functies w_τ volgt nu dat de eerste term van het rechterlid bij juiste keuze van N klein gemaakt kan worden, als we het bereik van τ beperken tot bijv. $0 < \tau \leq 1$.

Definiëren we de continuïteitsmodulus over $[a,b]$ van de functie $u(x,b)$ als ([8])

$$\omega(\delta) = \max_{\substack{|x-y| \leq \delta \\ a \leq x, y \leq b}} |u(x,b) - u(y,b)|,$$

dan geldt blijkbaar

$$|u(x-t,b) - u(x,b)| \leq \omega(|t|) \text{ voor } -N < x < N, \text{ als}$$

ω de continuïteitsmodulus over $-N - |t| < x < N + |t|$ is. We maken gebruik van de eigenschap ([8], p. 63)

$$\omega(\lambda\delta) \leq (\lambda+1) \omega(\delta) \quad (\lambda > 0),$$

en vinden voor $\mu > 0$ willekeurig

$$\omega(|t|) = \omega\left(\frac{1}{\mu} |t| \mu\right) \leq \omega\left(\frac{1}{\mu}\right) (|t| \mu + 1),$$

waaruit volgt dat de tweede term kleiner is dan

$$\left| \int_{-N}^N w_{\tau}(t) \left[\omega\left(\frac{1}{\mu}\right) (|t| \mu + 1) \right] dt \right| \leq \omega\left(\frac{1}{\mu}\right) \left(\frac{2\tau\mu}{\pi^{\frac{1}{2}}} + 1 \right).$$

Kiezen we nu $\mu = 1/\tau$, dan vinden we

$$\left| \int_{-N}^N \right| \leq \omega(\tau) \left(\frac{2}{\pi^{\frac{1}{2}}} + 1 \right),$$

zodat de tweede term van (7.20) klein is voor voldoende kleine τ .

Hiermee is bewezen dat

$$\lim_{\tau \rightarrow 0} u_{\tau}(x, b) = \lim_{\tau \rightarrow 0} w_{\tau}(x) = u(x, b) = u(x, b),$$

en dat deze uniform geldt op $-\infty < x < \infty$, zodat (B_{τ}) een regularisator is.

Opmerking 7.1. Het is duidelijk dat voor alle $y \in (0, b)$ eveneens uniform voor $-\infty < x < \infty$ geldt

$$\lim_{\tau \rightarrow 0} u_{\tau}(x, y) = u(x, y).$$

Als nu f en g niet precies bekend zijn, maar slechts met een precisie δ , dwz. we kennen slechts f_{δ} en g_{δ} met

$$\|f_{\delta} - f\| \leq \delta, \quad \|g_{\delta} - g\| \leq \delta,$$

dan weten we

- 1^o. dat $u_{\tau\delta} = B_{\tau}(f_{\delta}, g_{\delta})$ op stabiele wijze verkregen is; d.w.z. dat $u_{\tau\delta}$ maar weinig van u_{τ} verschilt, als δ klein is;
- 2^o. dat als we (τ, δ) op een slimme manier in het τ - δ -vlak naar de oorsprong laten gaan, $u_{\tau\delta}$ naar de ware oplossing u convergeert in de maximumnorm.

Willen we een expliciete schatting hebben voor $u - u_{\tau\delta}$, d.w.z. weten we welke orde in δ dit verschil naar 0 gaat, dan moeten we $\varepsilon_2(\tau, \delta)$ even uitrekenen.

$$\begin{aligned} |u_\tau(x, b) - u_{\tau\delta}(x, b)| &\leq |R_\tau(x, b) * (f - f_\delta)| + |T_\tau(x, b) * (g - g_\delta)| \\ &\leq \delta \{ |R_\tau(x, b)| + |T_\tau(x, b)| \}, \end{aligned}$$

waaruit, in verband met de integralen (7.17) en (7.18) volgt dat we voor $\varepsilon_2(\tau, \delta)$ kunnen nemen

$$\varepsilon_2(\tau, \delta) = \delta \left\{ \exp \frac{b^2}{4\tau^2} + \int_0^{b/2\tau} e^{t^2} dt \right\}.$$

Deze schatting is, wat de orde betreft, niet beter te krijgen, immers, als we bijvoorbeeld kiezen

$$f - f_\delta = \delta \cos \frac{bx}{4\tau^2}, \quad g - g_\delta = 0,$$

dan zouden we in verband met (7.16') vinden

$$\begin{aligned} u_\tau(0, b) - u_{\tau\delta}(0, b) &= \frac{\delta}{2\tau\pi^{\frac{1}{2}}} \exp \frac{b^2}{4\tau^2} \int_{-\infty}^{\infty} \cos^2 \frac{bt}{2\tau^2} \\ &= \frac{\delta}{2} \exp \left(\frac{b^2}{4\tau^2} + 1 \right). \end{aligned}$$

Bij gegeven δ kunnen we dus bepalen welke τ we moeten nemen, om ervoor te zorgen dat ε_1 en ε_2 van dezelfde orde van grootte zijn. De expliciete uitdrukking voor $u - u_{\tau\delta}$ is

$$\|u - u_{\tau\delta}\| \leq \left(\frac{2}{\pi^{\frac{1}{2}}} + 1 \right) \omega(\tau) + \delta \left\{ \exp \frac{b^2}{4\tau^2} + \int_0^{b/2\tau} e^{t^2} dt \right\}.$$

Dat deze schatting niet zo best is, hangt samen met het feit dat we slechts zeer weinig informatie over de oplossing op de lijn $y = b$ hebben gebruikt.

Voorbeeld 7.6. (zie ook [10], p. 18 e.v.). Het Cauchyprobleem voor de potentiaalvergelijking in een begrensde gebied $\Omega \subset \mathbb{R}^3$, waarbij op een deel $\partial_1 \Omega$ $u = f_0, \partial u / \partial \nu = g_0$ gegeven zijn, en op $\partial_2 \Omega = \partial \Omega \setminus \partial_1 \Omega$ niets. We nemen weer aan dat existentie en uniciteit (via Cauchy-Kowalewski) verzekerd zijn voor een keuze van f_0 en g_0 .

Een Carlemanfunctie G voor Ω en $\partial_1 \Omega$ zal zijn een functie van de driedimensionale vectoren x, ξ , en van de scalaire parameter $\tau > 0$, die voldoet aan

$$(7.21) \quad G(x, \xi, \tau) = \frac{1}{|x - \xi|} + G_1(x, \xi, \tau)$$

en

$$(7.22) \quad \frac{1}{4\pi} \int_{\partial_2 \Omega} \{ |G(x, \xi, \tau)| + \left| \frac{\partial G}{\partial \nu_\xi}(x, \xi, \tau) \right| \} d\sigma_\xi \leq \tau \text{ voor } x \in \Omega.$$

Hierbij is $|x - \xi| = \left(\sum_{i=1}^3 (x_i - \xi_i)^2 \right)^{\frac{1}{2}}$, $G_1(x, \xi, \tau)$ een functie die harmonisch is in Ω , en $\partial / \partial \nu_\xi$ de afgeleide m.b.t. ξ langs de naar binnen gerichte normaal, en $d\sigma_\xi$ een oppervlakte-element m.b.t. ξ . We willen nu het probleem

$$(7.23) \quad \begin{cases} \Delta u = 0 & \text{in } \Omega, \\ u = f & \text{op } \partial_1 \Omega, \\ \frac{\partial u}{\partial \nu} = g & \text{op } \partial_2 \Omega, \\ |u| + \left| \frac{\partial u}{\partial \nu} \right| \leq M & \text{op } \partial_2 \Omega, \end{cases}$$

behandelen.

Als we aannemen dat de Carlemanfunctie bekend is, dan kunnen we de operatoren $B_\tau (\tau > 0) : F \times G \rightarrow U$ definiëren als

$$B_\tau(f, g) = u_\tau = \frac{1}{4\pi} \int_{\partial_1 \Omega} \left\{ f \frac{G(x, \xi, \tau)}{\partial \nu_\xi} - G(x, \xi, \tau) g \right\} d\sigma_\xi.$$

Als we voor F en G enigszins redelijke klassen van Cauchydata nemen, dan vormen de B_τ een regularisator; immers B_τ is lineair en continu, en er geldt

$$(7.24) \quad \lim_{\tau \rightarrow 0} B_{\tau}(f, g) = \lim_{\tau \rightarrow 0} \frac{1}{4\pi} \left[\int_{\partial_1 \Omega} \left(f \frac{\partial G}{\partial \nu_{\xi}} - G g \right) d\sigma_{\xi} + \int_{\partial_2 \Omega} \left(u \frac{\partial G}{\partial \nu_{\xi}} - G \frac{\partial u}{\partial \nu} \right) d\sigma_{\xi} - \int_{\partial_2 \Omega} \left(u \frac{\partial G}{\partial \nu_{\xi}} - G \frac{\partial u}{\partial \nu} \right) d\sigma_{\xi} \right].$$

Daar u en $\partial u / \partial \nu$ begrensd zijn op $\partial_2 \Omega$ gaat de laatste term naar 0. Als we weten dat de oplossing bestaat, zoals voor $f = f_0$, $g = g_0$, dan volgt dat de limiet (7.24) gelijk is aan

$$\lim_{\tau \rightarrow 0} \frac{1}{4\pi} \int_{\partial \Omega} \left(u \frac{\partial G}{\partial \nu_{\xi}} - G \frac{\partial u}{\partial \nu} \right) d\sigma_{\xi} = u(x).$$

Opmerking 7.2. Er is een sterk verband tussen de functies van Carleman en van Green en Neumann. Wanneer $G(x, \xi)$ in $\xi = x$ een enkelvoudige pool heeft, dan geldt altijd

$$(7.25) \quad u(x) = \frac{1}{4\pi} \left\{ \int_{\partial_1 \Omega} + \int_{\partial_2 \Omega} \right\} \left[u \frac{\partial G}{\partial \nu_{\xi}} - G \frac{\partial u}{\partial \nu} \right] d\sigma_{\xi}, \quad \partial_1 \Omega \cup \partial_2 \Omega = \partial \Omega$$

Als we voor G de functie van Green nemen, d.w.z. $G = 0$ op $\partial \Omega$, dan vinden we u in termen van zijn randwaarden. Als we $\partial G / \partial \nu_{\xi} = 0$ op de rand nemen, dan vinden we u in termen van de waarden van zijn normale afgeleide op de rand. Bij het Cauchyprobleem zouden we er voor willen zorgen dat zowel G als $\partial G / \partial \nu_{\xi} = 0$ op het deel $\partial_1 \Omega$ van de rand, maar dit is onmogelijk. Daarom willen we dat de integraal

$$\int_{\partial_2 \Omega} \left(|G| + \left| \frac{\partial G}{\partial \nu_{\xi}} \right| \right) d\sigma_{\xi}$$

klein is.

Opmerking 7.3. Een expliciete uitdrukking voor het verschil $\|u_{\tau\delta} - u\|$ is niet moeilijk te vinden:

$$(7.26) \quad |u(x) - u_{\tau\delta}(x)| \leq \frac{1}{4\pi} \int_{\partial_2 \Omega} \left| \left[u \frac{\partial G}{\partial \nu_{\xi}} - G \frac{\partial u}{\partial \nu} \right] d\sigma_{\xi} \right| +$$

$$\begin{aligned}
& + \left| \frac{1}{4\pi} \int_{\partial_1 \Omega} \left[(f - f_\delta) \frac{\delta G}{\delta v_\xi} - G(g - g_\delta) \right] d\sigma_\xi \right| \leq \\
& \leq 2M\tau + \frac{1}{4\pi} \delta \cdot A(\partial_1 \Omega) \max_{\xi \in \partial_1 \Omega} (|G| + \left| \frac{\partial G}{\partial v_\xi} \right|),
\end{aligned}$$

waarin $A(\partial_1 \Omega)$ het oppervlak van $\partial_1 \Omega$ is.

Opmerking 7.4. In de uitdrukking (7.26) is de eerste term van de afschatting te identificeren met $\epsilon_1(\tau)$, en de tweede term met $\epsilon_2(\tau, \delta)$, d.w.z.

$$\begin{cases} \epsilon_1(\tau) = M\tau \\ \epsilon_2(\tau, \delta) = \frac{1}{4\pi} \delta \cdot A(\partial_1 \Omega) \cdot \max_{\xi \in \partial_1 \Omega} (|G| + \left| \frac{\partial G}{\partial v_\xi} \right|). \end{cases}$$

Volgen we ons algemeen principe, dan kiezen we $\epsilon_1(\tau)$ en $\epsilon_2(\tau, \delta)$ gelijk, zo dat kan, of van dezelfde orde. Bij gegeven δ is dan τ te bepalen.

Opmerking 7.5. Voor het tweedimensionale Cauchy-probleem gaat een en ander analoog aan het driedimensionale probleem.

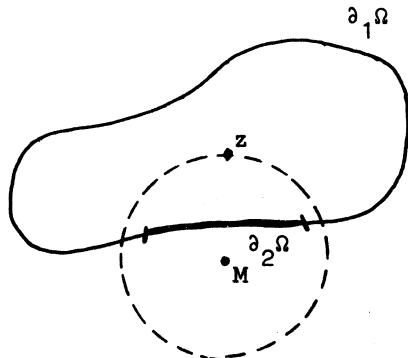
De Carlemanfunctie heeft dan de gedaante

$$(7.27) \quad G(x, \xi, \tau) = \ln |x - \xi| + G_1(x, \xi, \tau),$$

waarin $x = (x_1, x_2)$, $\xi = (\xi_1, \xi_2)$, $|x| = (x_1^2 + x_2^2)^{\frac{1}{2}}$.

Voorbeeld 7.5. Constructie van een Carlemanfunctie in het tweedimensionale probleem, voor bepaalde punten van Ω . We gaan over op de complete notatie: $z = x_1 + i x_2$, met $x = (x_1, x_2)$, $\xi = \xi_1 + i \xi_2$, $\xi = (\xi_1, \xi_2)$ willekeurig punt.

Laat $z \in \Omega$ een zodanig punt zijn, dat er een cirkel bestaat door z , waar het deel $\partial_2 \Omega$ van de rand geheel binnen ligt. Het coördinatenstelsel is zo te kiezen, dat de oorsprong in het middelpunt M van de cirkel terechtkomt, namelijk door een translatie $\zeta \rightarrow \zeta - \zeta_M$. Bovendien kunnen we ervoor zorgen dat $\arg z = 0$ door een rotatie $\zeta \rightarrow \zeta e^{-i \arg z}$.



Figuur 7.3

De straal van de cirkel is dan $|z|$. Er geldt dan ($z = |z|$):

$$\ln(z-\zeta) = \ln|z| + \ln\left(1 - \frac{\zeta}{z}\right).$$

Voor $|\zeta| < |z|$ geldt echter

$$\ln\left(1 - \frac{\zeta}{z}\right) = - \sum_{k=1}^{\infty} \frac{1}{k} \left(\frac{\zeta}{z}\right)^k.$$

Termsgewijs differentiëren levert

$$\left(1 - \frac{\zeta}{z}\right)^{-1} = - \sum_{k=1}^{\infty} \left(\frac{\zeta}{z}\right)^{k-1} \cdot \frac{1}{z}.$$

De functie

$$(7.28) \quad G(z, \zeta, n) = \ln(z-\zeta) - \ln|z| + \sum_{k=1}^{n-1} \frac{1}{k} \left(\frac{\zeta}{z}\right)^k$$

is dan een Carlemanfunctie voor z , Ω en $\partial_1\Omega$. Immers

$$|G(z, \zeta, n)| = \left| \sum_{k=n}^{\infty} \frac{1}{k} \left(\frac{\zeta}{z}\right)^k \right| \leq \frac{1}{n} \frac{|z|}{|z| - |\zeta|} \cdot \left|\frac{\zeta}{z}\right|^n$$

en

$$\left| \frac{\partial G}{\partial \zeta}(z, \zeta, n) \right| = \left| \sum_{k=n}^{\infty} \left(\frac{\zeta}{z}\right)^{k-1} \frac{1}{z} \right| \leq \frac{1}{|z| - |\zeta|} \left|\frac{\zeta}{z}\right|^{n-1}$$

in het inwendige van de cirkel. Er geldt nu

$$(7.29) \quad \int_{\partial_2\Omega} \left(|G(z, \zeta, n)| + \left| \frac{\partial G}{\partial \zeta}(z, \zeta, n) \right| \right) |d\zeta| \leq$$

$$\left(\frac{|\zeta|}{n} + 1 \right) \frac{1}{|z| - |\zeta|} \left|\frac{\zeta}{z}\right|^{n-1} \cdot l(\partial_2\Omega) \leq$$

$$\left(\frac{B}{n} + 1 \right) \cdot \frac{1}{|z| - B} \left(\frac{B}{|z|}\right)^{n-1} \cdot l(\partial_2\Omega),$$

waarin

$$B = \max_{\zeta \in \partial_2\Omega} |\zeta|, \quad l(\partial_2\Omega) = \text{lengte van } \partial_2\Omega.$$

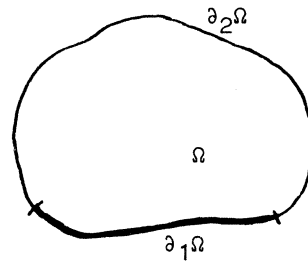
Door n voldoende groot te kiezen, is dit zo klein te krijgen als we maar willen.

Opmerking 7.6. Het karakteristieke van de methode van de regularisatie is, dat de oplossing in een enigszins gedeformeerde vorm wordt gevonden ten gunste van de stabiliteit. In de volgende sectie zullen we een andere methode behandelen, waarbij we niet te maken zullen hebben met een oplossingsmethode die uitgaande van de oorspronkelijke vergelijking de oplossing benadert, maar waarbij het oorspronkelijke probleem zodanig wordt vervormd, dat het verkregen probleem wel goed gesteld is.

7.3. Quasi-reversibiliteit

De nu volgende aanpak van onjuist gestelde problemen komt van Lions en Lattès [9]. Om de naam "quasi-reversibiliteit" (afgekort tot QR) te verklaren, moeten we terug naar voorbeeld 7.3. We hadden daar te maken met een irreversibel fysisch proces, en met de vraag: hoe kom ik achter een vroegere situatie, als de situatie nu bekend is. Met een foefje weten we in het wiskundig model de tijd wel om te keren: vandaar de naam quasi-reversibiliteit. Meer over de relatie tussen mathematische modellen en irreverzibele processen is te vinden in [2].

In deze sectie beschouwen we het volgende probleem. Laat Ω een open gebied in $\mathbb{R}^n$ zijn, met rand $\partial\Omega = \partial_1\Omega \cup \partial_2\Omega$, en laat L een sterk elliptische operator van de tweede orde in Ω zijn:



Figuur 7.4

$$(7.31) \quad Lu = - \sum_{i,j=1}^n D_i(a_{ij}(x)D_j u) + a_0 u$$

waarin

$$(7.32) \quad a_{ij} \in C^\infty(\bar{\Omega}), \quad a_0 \in C^\infty(\bar{\Omega}), \text{ met reële waarden, voor het gemak}$$

en

$$(7.33) \quad \sum_{i,j=1}^n a_{ij}(x) \xi_i \xi_j \geq \alpha_1 \sum_{i=1}^n \xi_i^2 \quad (\alpha_1 > 0),$$

$$(7.34) \quad a_0(x) \geq \alpha_0 > 0.$$

Het probleem is nu de functie u te bepalen, die voldoet aan

$$(7.35) \quad \begin{cases} Lu = h, h \in C^\infty(\bar{\Omega}), \text{ voor het gemak,} \\ u|_{\partial_1 \Omega} = f \quad (\text{gegeven}), \\ \frac{\partial u}{\partial \nu}|_{\partial_1 \Omega} = g \quad (\text{gegeven}), \end{cases}$$

wat in het algemeen een onjuist gesteld probleem is. Een directe numerieke benadering van (7.35) is daarom niet mogelijk. We merken op dat niet voor alle paren (f, g) het probleem (7.35) een oplossing u heeft; we zullen aannemen dat f en g zodanig zijn, dat het bestaan van u verzekerd is. Verder is u eenduidig bepaald door (7.35). Als er twee oplossingen zijn, u_1 en u_2 , neem dan het verschil $w = u_1 - u_2$. Hiervoor geldt

$$\begin{cases} Lw = 0 & \text{in } \Omega, \\ w|_{\partial_1 \Omega} = \frac{\partial w}{\partial \nu}|_{\partial_1 \Omega} = 0. \end{cases}$$

De stelling van Cauchy-Kowalewski geeft $w = 0$, als de coëfficiënten van L analytisch zijn. Dit geldt echter ook onder minder stringente voorwaarden. We zullen hier niet verder op ingaan.

Definieer nu

$$d(x, \partial_2 \Omega) = \text{afstand van } x \text{ tot } \partial_2 \Omega, x \in \Omega.$$

Laat ε een ("klein") positief getal zijn, dan definiëren we

$$(7.36) \quad M_\varepsilon(x) = \begin{cases} 1 & \text{als } d(x, \partial_2 \Omega) \geq 2\varepsilon \\ 0 & \text{als } d(x, \partial_2 \Omega) \leq \varepsilon \\ \in C^\infty(\bar{\Omega}), & \text{met waarden tussen 0 en 1.} \end{cases} \quad (x \in \Omega)$$

en

$$(7.37) \quad \Omega_\varepsilon = \Omega \setminus \{x \mid d(x, \partial_2 \Omega) \leq \varepsilon\}.$$

De QR-methode bestaat er uit dat aan het probleem (7.32) een familie van gemodificeerde problemen wordt toegevoegd, die wel goedgesteld zijn.

We definiëren u als de oplossing van

$$(7.38) \quad \begin{cases} L^*(M_\varepsilon^2 L u) = L^* M_\varepsilon^2 h \\ u_\varepsilon|_{\partial_1 \Omega} = f, \quad \frac{\partial u_\varepsilon}{\partial \nu}|_{\partial_1 \Omega} = g, \end{cases}$$

waarin L^* de geadjungeerde operator van L voorstelt (zie hoofdstuk 4 van [1]). We zullen verderop bewijzen dat dit probleem goed gesteld is. De operator die in (7.38) voorkomt, is van de vierde orde in Ω_ε , en die op het "onbekende" deel $\partial_2 \Omega$ van de rand zodanig degenereert, dat we daar geen randvoorwaarden nodig hebben.

We zullen eveneens laten zien dat

$$(7.39) \quad u_\varepsilon \rightarrow u \quad \text{als } \varepsilon \rightarrow 0,$$

waarbij de convergentie geschiedt in een zin, die we dan zullen aangeven. We hebben dan het onjuist gestelde probleem (7.35) benaderd door een familie goed gestelde problemen.

Opmerking 7.8. Willen we een numerieke oplossing vinden van het oorspronkelijke Cauchyprobleem (7.35), dan kan dit nu via een gekozen gemodificeerd probleem (d.w.z. voor een gekozen ε) gebeuren. We hebben dan namelijk een vierde orde elliptisch probleem, waarop de beschouwingen van hoofdstuk 6 over numerieke oplossingsmethoden van toepassing zijn.

Om te bewijzen dat (7.38) een goed gesteld probleem is, hebben we enkele begrippen en stellingen uit de functionaal-analytische benadering van randwaardeproblemen nodig. Deze Hilbertruimte-methoden voor de behandeling van elliptische randwaardeproblemen zijn in het eerste deel van dit colloquium, in hoofdstuk 5, uitvoerig aan de orde geweest, zie [1]. We zullen hier enkele begrippen en stellingen even herhalen, en maken daarbij gebruik van dezelfde nummering als in hoofdstuk 5, opdat het eventueel terugzoeken in [1] gemakkelijker zal zijn.

Stelling 4.11. De ruimte van alle oneindig vaak differentieerbare functies met compacte drager in Ω $C_c^\infty(\Omega)$ ligt dicht in $L^2(\Omega)$.

Definitie 5.1. Zij m een geheel getal ≥ 0 , en Ω een open gebied in $\mathbb{R}^n$.
Dan is de m^e Sobolevruimte

$$H^m(\Omega) = \{u \mid u \in \mathcal{D}'(\Omega), D^p u \in L^2(\Omega) \text{ voor } |p| \leq m\},$$

met inproduct $(\cdot, \cdot)_{m, \Omega}$ (of: $(\cdot, \cdot)_m$) en norm $\|\cdot\|_{m, \Omega}$:

$$(u, v)_{m, \Omega} = \sum_{|p| \leq m} \int_{\Omega} D^p u \overline{D^p v} \, dt.$$

Merk op dat $H^0(\Omega) = L^2(\Omega)$, en dat de norm $\|\cdot\|_0$ de bekende L^2 -norm is.

Uit deze definitie volgt

$$C_c^\infty(\Omega) \subset H^k(\Omega) \subset H^m(\Omega) \subset H^0(\Omega) = L^2(\Omega) \quad (k \geq m \geq 0).$$

Men gaat gemakkelijk na dat de injecties

$$\mathcal{D}(\Omega) \hookrightarrow H^k(\Omega) \hookrightarrow H^m(\Omega) \hookrightarrow \mathcal{D}'(\Omega) \quad k \geq m \geq 0$$

continu zijn. Hier stelt $\mathcal{D}'(\Omega)$ voor de ruimte van alle distributies in Ω .

Stelling 5.1. $H^m(\Omega)$ is een Hilbertruimte.

Definitie 5.2. $H_0^m(\Omega)$ is de afsluiting van $C_c^\infty(\Omega)$ in $H^m(\Omega)$. (Hier stelt $C_c^\infty(\Omega)$ de ruimte van alle oneindig vaak differentieerbare functies met compacte drager in Ω voor.) Deze ruimte kan men interpreteren als de deelruimte van alle elementen u uit $H^m(\Omega)$ waarvoor

$$u|_{\partial\Omega} = \frac{\partial u}{\partial \nu}|_{\partial\Omega} = \dots = \frac{\partial^{m-1} u}{\partial \nu^{m-1}}|_{\partial\Omega} = 0.$$

We zullen regulariteitseisen moeten opleggen aan het gebied Ω . We eisen het volgende:

Definitie 5.4. De rand $\partial\Omega$ van een open gebied $\Omega \subset \mathbb{R}^n$ heet glad, als er voor elke $x \in \partial\Omega$ een open omgeving W_x van x en een afbeelding κ_x van W_x op de bol $B^n = \{y \mid y \in \mathbb{R}^n, |y| < 1\}$ bestaan zodanig dat

- (i) κ_x is 1-1 en op en zowel κ_x als κ_x^{-1} zijn C^∞ -afbeeldingen;
- (ii) $\kappa_x(W_x \cap \Omega) = B_+^n (= \{y | y \in B^n, y_n > 0\})$;
- (iii) $\kappa_x(W_x \cap \partial\Omega) = B_0^n (= \{y | y \in B^n, y_n = 0\} = B^{n-1})$.

Het komt er dus op neer dat plaatselijk de omgeving van de rand moet lijken op een deel van de bovenhalfruimte $\mathbb{R}_+^n$. Een open gebied $\Omega \subset \mathbb{R}^n$ zullen we regulier noemen, als

- (i) Ω begreind is;
- (ii) $\partial\Omega$ glad is;
- (iii) Ω aan één kant van $\partial\Omega$ ligt.

In hoofdstuk 5 is bewezen: wanneer Ω regulier is, en $\bar{\Omega} \subset \Omega_1$, dan geldt: een functie u behoort tot $H^m(\Omega) \iff u$ is de restrictie tot Ω van een functie $v \in H_0^m(\Omega_1)$.

Stelling 5.7. (Sobolev). Wanneer $\Omega \subset \mathbb{R}^n$ regulier is, en $m > \frac{n}{2} + k$, dan is

$$H^m(\Omega) \subset C^k(\bar{\Omega}).$$

Stelling 5.11. Zij Ω regulier. Als $u \in C^{m-1}(\bar{\Omega}) \cap H_0^m(\Omega)$, dan geldt:

$$u|_{\partial\Omega} = \frac{\partial u}{\partial \nu}|_{\partial\Omega} = \dots = \frac{\partial^{m-1} u}{\partial \nu^{m-1}}|_{\partial\Omega} = 0.$$

In datzelfde hoofdstuk 5 hebben we aan elliptische operatoren bilineaire vormen toegevoegd. Het zou te ver voeren om hier precies aan te geven hoe dat allemaal gegaan is. We zullen hier daarom een min of meer heuristische beschouwing geven. Bekijken we eens de vergelijking

$$(7.40) \quad Lu = f \text{ in } \Omega,$$

en laat $u \in \mathcal{D}'(\Omega)$ een distributionele oplossing van deze vergelijking zijn, d.w.z. voor alle $\phi \in C_c^\infty(\Omega)$ geldt

$$\langle Lu, \phi \rangle = \langle f, \phi \rangle.$$

Als $f \in L^2(\Omega)$ en als u voldoende glad is, dan volgt hieruit

$$(7.41) \quad \sum_{i,j=1}^n (a_{ij} D^i u, D^j \phi)_{0,\Omega} + (a_0 u, \phi)_{0,\Omega} = (f, \phi)_{0,\Omega}.$$

Definiëren we voor u en $v \in H^1(\Omega)$ nu $E(u, v)$ door

$$E(u, v) = \sum_{i,j=1}^n (a_{ij} D^i u, D^j v)_{0,\Omega} + (a_0 u, v)_{0,\Omega},$$

dan is (7.41) ook te schrijven als

$$E(u, \phi) = (f, \phi)_{0,\Omega} \text{ voor alle } \phi \in C_c^\infty(\Omega).$$

En ook

$$E(u, v) = (f, v)_{0,\Omega} \text{ voor alle } v \in H_0^1(\Omega),$$

immers, de functionaal $v \rightarrow L(u, v)$ is bij vaste u begrensd op $H_0^1(\Omega)$ en $C_c^\infty(\Omega)$ ligt dicht in $H_0^1(\Omega)$. Dit bracht ons ertoe het probleem vergelijking (7.40) op te lossen te vervangen door

$$\text{gegeven: } f \in L^2(\Omega)$$

$$\text{gevraagd: } u \in H_0^1(\Omega) \text{ zodanig dat}$$

$$E(u, v) = (f, v)_{0,\Omega} \text{ voor alle } v \in H_0^1(\Omega).$$

Door te eisen dat $u \in H_0^1(\Omega)$ moet zijn, eisen we (zie stelling 5.11) dat u in een gegeneraliseerde zin aan de randvoorwaarde $u|_{\partial\Omega} = 0$ moet voldoen. Door i.p.v. $u \in H_0^1(\Omega)$ te eisen $u \in V$ met V een **gesloten** deelruimte van $H^1(\Omega)$ die $H_0^1(\Omega)$ omvat konden we de randvoorwaarden variëren. Voor details zie men sectie 5.2 van [1]. We zullen de bilineaire vorm $E(u, v)$ niet nodig hebben; de invoering ervan diende ervoor te memoreren hoe aan elliptische differentiaalvergelijkingen bilineaire vormen kunnen worden toegevoegd. De volgende stelling (uit hoofdstuk 5) vertelt ons wanneer zo'n probleem oplosbaar is.

Stelling 5.16. Zij H een Hilbertruimte, en V een lineaire deelruimte V van H die dicht ligt in H . Laten we het inproduct in H aangeven met $(\cdot, \cdot)_H$, en de norm met $||\cdot||_H$. Laat verder op V een inproduct $(\cdot, \cdot)_V$ gedefinieerd

zijn, en V in dit inproduct een Hilbertruimte zijn, waarvan we de norm aangeven met $||\cdot||_V$. Verder nemen we aan dat er een $k_0 \geq 0$ bestaat zó dat

$$||v||_H \leq k_0 ||v||_V.$$

Laat verder gegeven zijn een bilineaire vorm B op V , waarvoor geldt

- (i) B is begrensd op V , d.w.z. er bestaat een $k \geq 0$ zó dat

$$|B(u,v)| \leq k ||u||_V ||v||_V \text{ voor alle } u,v \in V;$$

- (ii) B is sterk coërcief, d.w.z. er bestaat een constante $c > 0$ zó dat

$$B(u,u) \geq c ||u||_V^2 \quad \text{voor alle } u \in V.$$

Dan geldt: er bestaat bij iedere $f \in H$ precies één $u \in V$ zodanig dat

$$B(u,v) = (f,v)_H \quad \text{voor alle } v \in V.$$

Tenslotte citeren we een stelling die ons zegt wanneer een gegeneraliseerde oplossing tevens klassieke oplossing is.

Stelling 5.31. Laat Ω regulier zijn, $f \in C^\infty(\bar{\Omega})$ en $u \in H^1(\Omega)$ een gegeneraliseerde oplossing zijn van de vergelijking $Lu = f$, d.w.z. $\langle Lu, \phi \rangle = (f, \phi)_{0,\Omega}$ voor alle $\phi \in C_c^\infty(\Omega)$. Dan geldt $u \in C^\infty(\bar{\Omega})$.

Na zo genoemd te hebben welke stellingen we uit de Hilbertruimtheorie nodig zullen hebben, beginnen we met het eigenlijke bewijs dat (7.38) een goed gesteld probleem is, en dat $u_\epsilon \rightarrow u$ in een nog nader aan te geven zin als $\epsilon \rightarrow 0$.

Lemma 7.2. Voor de operator L zoals gedefinieerd in (7.31) geldt dat er een constante $\alpha > 0$ bestaat zodanig dat

$$(7.42) \quad (L\phi, \phi)_{0,\Omega} \geq \alpha ||\phi||_{1,\Omega}^2 \quad \text{voor alle } \phi \in C_c^\infty(\Omega).$$

Bewijs. Er geldt, wegens de eigenschappen (7.33) en (7.34)

$$\begin{aligned}
(L\phi, \phi)_{0,\Omega} &= - \int_{\Omega} \sum_{i,j=1}^n (D_i(a_{ij} D_j \phi)) \phi \, dx + \int_{\Omega} a_0 \phi^2 \, dx \\
&= \int_{\Omega} \sum_{i,j=1}^n a_{ij} (D_j \phi) (D_i \phi) \, dx + \int_{\Omega} a_0 \phi^2 \, dx \\
&\geq \alpha_1 \int_{\Omega} \sum_{i=1}^n (D_i \phi)^2 \, dx + \alpha_0 \int_{\Omega} \phi^2 \, dx \\
&\geq \min(\alpha_1, \alpha_0) \cdot \|\phi\|_{1,\Omega}^2.
\end{aligned}$$

Met behulp van Cauchy-Schwarz volgt er uit (7.42)

$$\text{const.} \|L\phi\|_{0,\Omega} \|\phi\|_{0,\Omega} \geq \|\phi\|_{1,\Omega}^2 \geq \|\phi\|_{0,\Omega}^2 \text{ voor alle } \phi \in C_c^\infty(\Omega).$$

Lemma 7.3. De operator L is zelfgeadjungeerd, d.w.z. $L = L^*$, en verder geldt dat er een constante $\beta > 0$ bestaat zó dat

$$(7.43) \quad (L^* L \phi, \phi)_{0,\Omega} \geq \beta \|\phi\|_{2,\Omega}^2 \text{ voor alle } \phi \in C_c^\infty(\Omega).$$

Bewijs. Volgt via partiële integraties en gebruikmaking van eigenschappen (7.33) en (7.34).

Het bewijs dat (7.38) een unieke oplossing heeft verloopt nu als volgt. We voeren eerst een ruimte V^ε in van functies v die (in een gegeneraliseerde zin) voldoen aan homogene randvoorwaarden op het deel $\partial_1 \Omega$ van de rand $\partial\Omega$, en die aan geen enkele voorwaarde op het andere deel $\partial_2 \Omega$ van de rand hoeven te voldoen. We voegen verder aan de operator $L M_\varepsilon^* L$ een bilineaire vorm π^ε op V^ε toe, en bewijzen dat deze voldoet aan de eisen van stelling 5.16, met $V = V^\varepsilon$, $H = L^2(\Omega_\varepsilon)$, zodat we de existentie en uniciteit bewezen hebben van een oplossing v_ε van

$$\pi^\varepsilon(v_\varepsilon, v) = 0 \text{ voor alle } v \in V^\varepsilon.$$

Tenslotte bewijzen we dat $u_\varepsilon = u - v_\varepsilon$ voldoet aan (7.38).

Definitie 7.1. Met de ruimte V^ϵ zullen we aangeven

$$V^\epsilon = \{v | M_\epsilon v \in H_0^2(\Omega), M_\epsilon Lv \in L^2(\Omega)\},$$

met de norm

$$|||v||| = ||M_\epsilon Lv||_{0,\Omega}.$$

Opmerking 7.9. We zien dat inderdaad aan v in de buurt van $\partial_2\Omega$ geen eisen worden opgelegd.

Opmerking 7.10. Wanneer $v \in C(\overline{\Omega_\delta}) \cap V^\epsilon$ voor alle $\delta > \epsilon$, dan zijn v en $\frac{\partial v}{\partial \nu}$ op dat deel van $\partial_1\Omega$ dat tevens rand van Ω_δ is, gelijk aan nul. Dit volgt uit stelling 5.11.

Lemma 7.4. De ruimte V^ϵ voorzien van de norm $|||.|||$ is een Hilbertruimte.

Bewijs. Laat $\delta > \epsilon$ willekeurig zijn. Daar $M_\epsilon \geq K > 0$ op Ω_δ , geldt er

$$||M_\epsilon Lv||_{0,\Omega} \geq K ||Lv||_{0,\Omega_\delta}.$$

Zij nu $\eta > \delta$ willekeurig, en beschouw functies $\phi \in C_c^\infty(\Omega_\delta)$. Voor zulke functies ϕ geldt i.v.m. lemma 7.3

$$\langle L^*L\phi, \phi \rangle \geq \text{const.} ||\phi||_{2,\Omega_\delta}^2$$

ofwel

$$||L\phi||_{0,\Omega_\delta}^2 = (L\phi, L\phi)_{0,\Omega_\delta} \geq \text{const.} ||\phi||_{2,\Omega_\delta}^2.$$

Hetzelfde geldt dan als $\phi \in H_0^2(\Omega_\delta)$ daar $C_c^\infty(\Omega_\delta)$ dicht in $H_0^2(\Omega_\delta)$. Omdat elke functie $v \in H_0^2(\Omega_\eta) \cap V^\epsilon$ te zien is als een beperking van een functie uit $H_0^2(\Omega_\delta)$, geldt dus

$$||Lv||_{0,\Omega_\delta}^2 \geq \text{const.} ||v||_{2,\Omega_\eta}^2, \text{ voor alle } v \in V^\epsilon$$

zodat we hebben gevonden

$$|||v||| \geq \text{const.} ||v||_{2,\Omega_\eta}.$$

Hieruit volgt het lemma.

Definitie 7.2. We voeren nu een bilineaire vorm in, t.w.

$$\pi^\varepsilon(u, v) = (M_\varepsilon Lu, M_\varepsilon Lv)_{0, \Omega},$$

voor alle u en $v \in V^\varepsilon$.

Lemma 7.5. V^ε en π^ε voldoen aan de volgende eigenschappen:

- (i) V^ε ligt dicht in $H^\varepsilon = \{v | M_\varepsilon v \in L^2(\Omega)\}$ (met $\|v\|_{H^\varepsilon} = \|M_\varepsilon v\|_{0, \Omega}$);
- (ii) de injectie $V^\varepsilon \hookrightarrow H^\varepsilon$ is begrensd;
- (iii) $\pi^\varepsilon(u, v)$ is begrensd op V^ε ;
- (iv) $\pi^\varepsilon(u, v)$ is sterk coërcief over V^ε .

Bewijs.

(i) Daar $C_c^\infty(\Omega)$ reeds dicht in $L^2(\Omega)$ ligt, volgens stelling geldt hetzelfde voor V^ε .

(ii) $\|v\|_{0, \Omega_\eta} \leq \text{const.} \|Lv\|_{0, \Omega_\delta}$ voor alle $\eta > \delta > \varepsilon$, zie lemma 7.2. We passen hier eenzelfde truc toe als in het bewijs van lemma 7.4.

$$(iii) |\pi^\varepsilon(u, v)| = (M_\varepsilon Lu, M_\varepsilon Lv)_{0, \Omega} \leq \|u\| \cdot \|v\|.$$

$$(iv) \pi^\varepsilon(v, v) = \|M_\varepsilon Lv\|_{0, \Omega}^2 = \|v\|_{H^\varepsilon}^2$$

Wegens dit lemma kunnen we de Hilbertruimte-theorie van hoofdstuk 5 van deel 1 van deze syllabus [1] toepassen. Er volgt, i.v.m. stelling 5.16, dat er precies één $v_\varepsilon \in V^\varepsilon$ bestaat die voldoet aan

$$(7.44) \quad \pi^\varepsilon(v_\varepsilon, v) = 0.$$

Dit betekent dat v_ε voldoet aan de vergelijking

$$(M_\varepsilon Lv_\varepsilon, M_\varepsilon L\phi)_{0, \Omega} = (L^* M_\varepsilon^2 Lv_\varepsilon, \phi)_{0, \Omega} = 0$$

voor alle $\phi \in C_c^\infty(\Omega)$, zodat v_ε in gegeneraliseerde zin voldoet aan

$$(7.45) \quad L^* M_\varepsilon^2 Lv_\varepsilon = 0.$$

Daar het rechterlid een element is van $C^\infty(\bar{\Omega})$, en $v_\varepsilon \in H^2(\Omega_\varepsilon)$, volgt i.v.m. stelling 5.31 dat $u \in C^\infty(\bar{\Omega}_\delta)$ voor alle $\delta > \varepsilon$, zodat aan (7.45) ook in klassieke zin voldaan is.

We zijn nu in staat de volgende stelling te formuleren.

Stelling 7.1. Laat L gegeven zijn zoals in (7.31) t/m (7.34) en Ω een regulier gebied zijn. Laten f en g verder gegeven analytische functies zijn, zodanig dat er een unieke $u \in C^\infty(\bar{\Omega})$ bestaat die voldoet aan

$$(7.35) \quad \begin{cases} Lu = h & \text{in } \Omega, \\ u|_{\partial_1 \Omega} = f, \quad \frac{\partial u}{\partial \nu}|_{\partial_1 \Omega} = g. \end{cases}$$

Dan geldt: er bestaat een unieke $u_\epsilon \in C^\infty(\bar{\Omega}_\epsilon) \cap V^\epsilon$ die voldoet aan

$$(7.38) \quad \begin{cases} L^* M_\epsilon^2 Lu_\epsilon = L^* M_\epsilon^2 h, \\ u_\epsilon|_{\partial_1 \Omega} = f, \quad \frac{\partial u_\epsilon}{\partial \nu}|_{\partial_1 \Omega} = g. \end{cases}$$

Bewijs. Laat v_ϵ gedefinieerd zijn als in (7.44). Beschouw

$$(7.46) \quad u_\epsilon = u - v_\epsilon.$$

Er geldt dan $u_\epsilon \in C^\infty(\bar{\Omega}_\epsilon)$, en verder $u_\epsilon|_{\partial_1 \Omega} = f$, $\frac{\partial u_\epsilon}{\partial \nu}|_{\partial_1 \Omega} = g$. Uit (7.46) volgt verder

$$L^* M_\epsilon^2 Lu_\epsilon = L^* M_\epsilon^2 Lu - L^* M_\epsilon^2 Lv_\epsilon = L^* M_\epsilon^2 h.$$

Nu de existentie en de uniciteit van de oplossing van het gemodificeerde probleem (7.38) is bewezen, rest ons te bewijzen dat (7.38) een goed gesteld probleem is, d.w.z. dat u_ϵ continu afhangt van f en g ; verder moeten we nog het verband aangeven tussen u en u_ϵ . We beginnen met het laatste.

Stelling 7.2. Laat u de oplossing zijn van het oorspronkelijke probleem (7.35), en u_ϵ van het gemodificeerde probleem (7.38). Dan geldt

$$u_\epsilon = u \text{ in } \Omega_\epsilon.$$

Bewijs. De functie v_ϵ uit de vorige stelling voldoet aan

$$L^* M_\epsilon^2 L v_\epsilon = 0$$

(7.47)

$$v_\epsilon \Big|_{\partial_1 \Omega} = \frac{\partial v_\epsilon}{\partial \nu} \Big|_{\partial_1 \Omega} = 0.$$

Er geldt echter

$$\int_{\Omega} (L^*(M_\epsilon^2 L v_1)) v_2 dx = \int_{\Omega} (M_\epsilon L v_1)(M_\epsilon L v_2) dx$$

voor functies v_1 en $v_2 \in V^\epsilon$, zodat geldt

$$\|M_\epsilon L v_\epsilon\|_0^2 = \int_{\Omega} (M_\epsilon L v_\epsilon)^2 dx = 0,$$

waaruit volgt dat $L v_\epsilon = 0$ in Ω_ϵ . Tezamen met (7.47) volgt nu het resultaat.

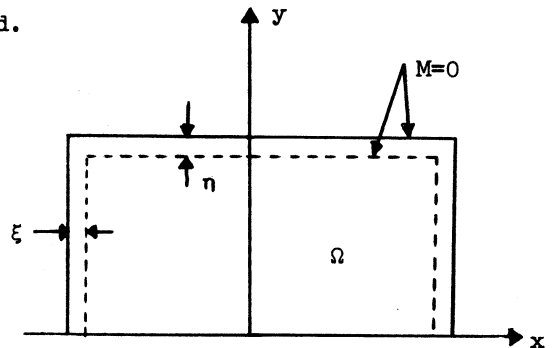
Opmerking 7.11. We hebben het bewijs gegeven voor het geval dat de coëfficiënten van L en randfuncties analytisch zijn, en $a_0(x) \geq \alpha_0 > 0$. De stellingen blijven doorgaan voor het geval dat de coëfficiënten en randfuncties minder glad zijn, en ook voor $a_0(x) = 0$. Hiervoor is men verwezen naar [9]. Het bewijs voor deze gevallen is in essentie hetzelfde.

We geven nu een numeriek voorbeeld.

Voorbeeld 7.6. Beschouw de rechthoek (figuur 7.5)

$$\begin{cases} -1 < x < 1, \\ 0 < y < 1. \end{cases}$$

We zoeken de oplossing van



Figuur 7.5

$$(7.48) \quad \left\{ \begin{array}{l} \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \\ u(x, 0) = f(x) \text{ gegeven,} \\ u_y(x, 0) = g(x) \text{ gegeven.} \end{array} \right.$$

Op dit probleem willen we de zojuist beschreven methode toepassen. We worden ertoe geleid het volgende probleem op te lossen

$$(7.49) \quad \left\{ \begin{array}{l} \Delta M_\epsilon^2 u_\epsilon = 0, \\ u_\epsilon(x, 0) = f(x), \\ \frac{\partial u_\epsilon}{\partial y}(x, 0) = g(x). \end{array} \right.$$

Om (7.49) numeriek aan te kunnen pakken leggen we een rooster over de rechthoek Ω met roosterafstanden $\xi = 1/J$ langs de x-as en $\eta = 1/L$ langs de y-as, waarin J en L gehele positieve getallen zijn. We kiezen M_ϵ zo, dat hij gelijk is aan nul op de randen die niet samenvallen met de x-as, en ook op de roosterlijnen parallel met deze randen. We kiezen voor de operator Δ een discretisering, zoals bijvoorbeeld in sectie 6.1 gebeurd is:

$$(6.4) \quad D = \begin{bmatrix} L_1 & L_3 & L_1 \\ L_2 & L_4 & L_2 \\ L_1 & L_3 & L_1 \end{bmatrix},$$

waarin

$$(6.5) \quad \left\{ \begin{array}{l} L_1 = \frac{1}{2} \xi^{-2} (1 + \rho^2 - \gamma), \quad L_2 = \xi^{-2} (\gamma - \rho^2), \\ L_3 = \xi^{-2} (\gamma - 1) \quad , \quad L_4 = -2\gamma \xi^{-2}, \\ \rho = \xi/\eta \quad , \quad \gamma \text{ reële parameter.} \end{array} \right.$$

We maken hier gebruik van de in hoofdstuk 6 ingevoerde notatie. Geven we met $\tilde{u}[j,1]$ aan $u(j\xi,1\eta)$, dan worden we geleid tot het matrixprobleem

$$(7.50) \quad \left\{ \begin{array}{l} DM_{\varepsilon}^2 \tilde{u}[j,1] = 0, \quad -J \leq j \leq J, \quad 2 \leq 1 \leq L, \\ \tilde{u}[j,0] = \tilde{f}[j], \quad -J \leq j \leq J, \\ \frac{1}{2\eta} (\tilde{u}[j,1] - \tilde{u}[j,-1]) = \tilde{g}[j], \quad -J \leq j \leq J. \end{array} \right.$$

Dit probleem moet nu met een of andere numerieke techniek worden opgelost.

Opmerking 7.12. De schrijver van dit hoofdstuk was bij het ter perse gaan van dit boek nog in een correspondentie gewikkeld met J.L. Lions ([9]) over de juistheid van een bewijs van de eigenschap dat u_{ε} continu afhangt van f en g . Het is dus helaas niet mogelijk het bewijs hier op te nemen.

Literatuur

- [1] Coolen, T.M.T.,
Förch, G.J.R.,
Jager, E.M. de,
Pijls, H.G.J. Colloquium elliptische differentiaal-
vergelijkingen, deel 1.
Math. Centrum Amsterdam, 1969.
- [2] Coolen, T.M.T. De wiskunde gezien vanuit haar toepassing.
Rapport ZW 1969-015. Mathematisch Centrum
Amsterdam, 1969.
- [3] Garabedian, P.R. Partial differential equations.
Wiley, New York enz. 1964.
- [4] Grant, F.S.,
West, G.F. Interpretation theory in applied geophysics.
MacGraw Hill, New York, 1965
- [5] Hadamard, J. Le problème de Cauchy.
Hermann, Paris, 1932.
- [6] Hellwig, G. Partial differential equations.
Blaisdell, New York, 1964.
- [7] Ivanov, V.K. The Cauchyproblem for the Laplace equation
in an infinite strip.
Differential equations 1(1965), 98-102,
Eng. vert. uit Differensial'nye Uravneniya
1(1965), 131-136.
- [8] Korovkin, P.P. Linear operators and approximation theory
Hindustan Publ. Co., Delhi, 1960.
Vertaling uit het Russisch.
- [9] Lattès, R.,
Lions, J.L. Méthode de quasi-reversibilité et
applications.
Dunod, Paris, 1967.
- [10] Lavrentiev, M.M. Some improperly posed problems of mathe-
matical physics.
Springer, Berlin enz. 1967.
Vert. uit het Russisch door R.J. Sacker.

- [11] Tychonov, A.N. Solution of incorrectly posed problems
and the regularization method.
Soviet Mathematics 4(1963), 1035-1039.
- [12] Tychonov, A.N. Regularization of incorrectly posed
problems.
Soviet Mathematics 4(1963), 1624-1627.

8. OVER DE INVLOED VAN NUMERIEKE FOUTEN BIJ VARIATIEBEREKENINGEN

Vele randwaardeproblemen uit de mathematische physica zijn ook te formuleren als variatieproblemen (zie o.a. [1], [3], [4], [5], [6] en [7]). De laatste vorm heeft het voordeel dat vaak met vrij weinig rekenwerk al vrij goede benaderingen voor de oplossing van het probleem zijn te vinden; in het bijzonder wanneer de rand een mooie vorm heeft en de randvoorwaarden van eenvoudige aard zijn, omdat dan gemakkelijk goed hanteerbare, bij het probleem aangepaste, variatiefuncties te construeren zijn. Voor de berekening maakt het dan weinig verschil of men met eindige of oneindige gebieden te doen heeft, terwijl ook de dimensie van het probleem geen grote rol speelt.

In dit opzicht is de variatiemethode in het voordeel tegenover de differentiemethode (vgl. hoofdstuk 6), omdat bij deze laatste methode het rekenwerk bij groter wordende dimensie aanzienlijk toeneemt. Bij gebieden met gecompliceerde rand is de differentiemethode vaak weer beter. Een nadeel van de variatiemethode is bovendien, dat de nauwkeurigheid van de resultaten, vooral bij numerieke berekeningen, nogal sterk afhangt van de keuze van de variatiefuncties. Dit laatste aspect, dat van de numerieke stabiliteit, is uitvoerig onderzocht door Michlin en medewerkers (zie [6], waarin talrijke voorbeelden zijn gegeven).

In dit hoofdstuk kiezen we als voorbeeld om de invloed van numerieke fouten aan te demonstreren de Rayleigh-Ritz methode voor het oplossen van eigenwaardeproblemen. Daarbij volgen we niet de methode van Michlin, maar, in navolging van Delves [2], een methode die wat directer aansluit bij de rekenpraktijk.

In principe levert de Rayleigh-Ritz methode voor elk van de eigenwaarden een bovengrens, die met het toenemen van het aantal variatieparameters monotoon convergeert naar de werkelijke waarde. Numerieke benaderingsfouten kunnen echter leiden tot aanzienlijke afwijkingen, zelfs zodanig dat de berekende Rayleigh-Ritz waarden kleiner worden dan de werkelijke waarden, terwijl ook de monotonie niet meer gegarandeerd is.

Om een betrouwbare bovengrens te verkrijgen moet een schatting voor de numerieke fout in rekening worden gebracht. Dit leidt in het algemeen tot onnodig hoge bovengrenzen. Door Delves [2] is een correctie op de Rayleigh-Ritz methode voorgesteld, waardoor het effect van de afrondingsfouten verkleind kan worden.

8.1. De Rayleigh-Ritz methode

We beschouwen het eigenwaardeprobleem

$$(8.1) \quad (\mathcal{H} - \lambda \mathcal{N}) \psi = 0,$$

waarbij $\mathcal{H}$ en $\mathcal{N}$ zelfgadjungeerde elliptische differentiaaloperatoren zijn, gedefinieerd op een gebied $\Omega \subset \mathbb{R}^n$, $\mathcal{N}$ positief definit, terwijl de oplossing ψ moet voldoen aan zekere homogene randvoorwaarden op de rand $\delta\Omega$ (Neumann-, Dirichlet- of gemengde randvoorwaarden). Zij nu W de verzameling van alle functies χ die voldoen aan de gestelde randvoorwaarden, die bovendien continu zijn op $\Omega + \delta\Omega$, terwijl zowel $\mathcal{H}\chi$ als $\mathcal{N}\chi$ continu zijn op Ω . Dan zijn volgens het variatieprincipe ([1], [3], [4], [7]) de eigenwaarden van (8.1) gelijk aan de stationaire waarden van de functionaal

$$(8.2) \quad \Lambda[\chi] = \int_{\Omega} \chi \mathcal{H} \chi \, dx \bigg/ \int_{\Omega} \chi \mathcal{N} \chi \, dx, \quad \chi \in W.$$

Bij de Rayleigh-Ritz methode worden nu van te voren een aantal lineair onafhankelijke basisfuncties $\phi_i \in W$ ($i = 1, \dots, m$) uitgekozen. Als variatiefuncties worden dan toegelaten de functies χ_m van de vorm

$$(8.3) \quad \chi_m = \sum_{i=1}^m a_i \phi_i, \quad \phi_i \in W, \quad i = 1, \dots, m,$$

waarbij de m getallen a_i de variatieparameters zijn. Men varieert dus slechts binnen de m -dimensionale lineaire deelruimte $W_m \subset W$ opgespannen door de m functies ϕ_i .

Substitutie van χ_m voor χ maakt van de beide integralen in (8.2) kwadratische vormen in de parameters a_i . De stationaire waarden van $\Lambda[\chi_m]$ worden gevonden door de afgeleiden naar a_i gelijk aan nul te stellen. Dit resulteert in m homogene lineaire vergelijkingen voor de parameters a_i , te schrijven als een matrix-eigenwaardevergelijking van de vorm

$$(8.4) \quad [H - \lambda_\mu N] \alpha = 0.$$

Hierin zijn H en N symmetrische $m \times m$ -matrices, N positief definit, met elementen

$$(8.5) \quad H_{ij} = \int_{\Omega} \phi_i \mathcal{H} \phi_j \, dx, \quad N_{ij} = \int_{\Omega} \phi_i \mathcal{N} \phi_j \, dx$$

en is $\alpha = (a_1, \dots, a_m)$ de m -vector van variatieparameters. Wanneer het probleem (8.1) een kleinste eigenwaarde λ_0 heeft, dan zijn de eigenwaarden λ_μ van (8.4) steeds groter dan of hoogstens gelijk aan de corresponderende eigenwaarden λ van (8.1), omdat de variatieruimte W groter is dan de Rayleigh-Ritz ruimte W_m . In het algemeen is de afwijking bij de lagere eigenwaarden kleiner dan bij de hogere. Heeft het probleem (8.1) tenminste $m + 1$ discrete eigenwaarden λ_k dan geldt echter steeds

$$(8.6) \quad \lambda_k \leq \lambda_{\mu k} \leq \lambda_{k+1} \quad (k=0, 1, \dots, m-1),$$

waarbij beide series van eigenwaarden in volgorde van grootte genummerd zijn.

Voegt men nu aan de m basisfuncties ϕ_i een nieuwe basisfunctie ϕ_{m+1} toe, dan zullen de eerste m eigenwaarden λ_μ van dit nieuwe $(m+1)$ -dimensionale Rayleigh-Ritz probleem, omdat het variatiegebied weer uitgebreid is, kleiner zijn dan of hoogstens gelijk aan de corresponderende eigenwaarden van het m -dimensional probleem.

Op deze wijze verkrijgt men met toenemende m voor elke $k = 0, 1, 2, \dots$ een monotoon dalende rij van getallen $\lambda_{\mu k}$, die convergeert naar de gezochte eigenwaarde λ_k van het probleem (8.1), terwijl onder bepaalde condities (zie bv. [3]) de bijbehorende rij functies χ_{mk} convergeert naar de bijbehorende eigenfunctie ψ_k .

In het algemeen verkrijgt men al vrij gauw een redelijke benadering voor de laagste eigenwaarden, ook wanneer de eigenfunctie nog vrij slecht benaderd is. De reden daarvan is, dat de functionaal $\Lambda[\chi]$ in feite kwadratisch is in χ , zodat een eerste-orde fout in χ in de buurt van een stationair waarde λ_k van Λ slechts resulteert in een tweede-orde fout voor λ_k . Daar men meestal meer geïnteresseerd is in de eigenwaarden dan in de eigenfuncties, wordt dit niet als een bezwaar ondervonden.

8.2. Afrondingsfouten

De in de voorgaande paragraaf vermelde resultaten, in het bijzonder de ongelijkheid (8.6), zijn alleen geldig wanneer de berekeningen exact uitgevoerd kunnen worden. In de numerieke praktijk is dat niet meer het geval. Allereerst hebben we daar te maken met fouten bij de berekening van de matrixelementen H_{ij} en N_{ij} uit (8.5), ten dele veroorzaakt door het rekenen met eindige precisie (bepaald door de woordlengte van de computer), ten dele veroorzaakt door het gebruik van approximatietechnieken bij het numerieke integreren. Deze laatste foutenbron is meestal overheersend. Tengevolge van deze fouten is het matrixprobleem, dat we gaan oplossen, niet het probleem (8.4), maar in feite het probleem

$$(8.7) \quad [\underline{H} + h - \lambda_c(N+n)] b = 0$$

waarbij h en n de fouten zijn in de berekende matrices. Omdat dit een iets gewijzigd probleem is, hebben we de eigenwaarde nu λ_c genoemd en de eigenvector b .

Een tweede foutenbron is gelegen in de numerieke berekening van de eigenwaarden λ_c van (8.8). De daarbij optredende fout is echter in het algemeen klein t.o.v. de fout veroorzaakt door het niet nul zijn van h en n . Bovendien kan de eigenwaardeberekening gemakkelijk verbeterd worden door met meervoudige precisie te werken, zonder dat dit teveel extra reken-tijd kost.

Ook de matrixelementen kunnen in principe nauwkeuriger berekend worden, maar dat kost (waar men gewoonlijk met meerdimensionale integraties te doen heeft) aanzienlijk meer tijd. We zullen daarom in het vervolg aannemen, dat de fouten in de variatieberekening alleen veroorzaakt worden door het niet nul zijn van de foutenmatrices h en n .

Opmerking: De matrices h en n zijn te beschouwen als symmetrische matrices, omdat wegens de onderstelde symmetrie van H en N slechts de integralen van (8.5) met $i \geq j$ (of omgekeerd) uitgerekend worden.

Een schatting van de fout in de berekende eigenwaarden

We zullen nu een bovengrens bepalen voor de afwijking van de uit (8.7) berekende eigenwaarden λ_c van de eigenwaarden λ_μ van (8.4), die bij exacte berekeningen gevonden zouden zijn. Door vermenigvuldiging met $N^{-\frac{1}{2}}$ brengen we de vergelijkingen (8.4) en (8.7) in de vorm

$$(8.8) \quad [\hat{H} - \lambda_\mu \bar{I}] \hat{a} = 0,$$

$$(8.9) \quad [\hat{H} + \hat{h} - \lambda_c \hat{n} - \lambda_c \bar{I}] \hat{b} = 0$$

waarbij I de eenheidsmatrix is en

$$(8.10) \quad \begin{aligned} \hat{H} &= N^{-\frac{1}{2}} H N^{-\frac{1}{2}}, \quad \hat{h} - \lambda_c \hat{n} = N^{-\frac{1}{2}} (h - \lambda_c n) N^{-\frac{1}{2}}, \\ \hat{a} &= N^{\frac{1}{2}} a, \quad \hat{b} = N^{\frac{1}{2}} b \end{aligned}$$

Uit een theorema van Gerschgorin volgt dan [8, p. 93] voor het verschil van de k^{de} eigenwaarden:

$$(8.11) \quad |\lambda_{\mu k} - \lambda_{ck}| \leq \|\hat{h} - \lambda_{ck} \hat{n}\| \leq \|N^{-\frac{1}{2}}\|^2 (\|h\| + |\lambda_{ck}| \cdot \|n\|),$$

waarbij voor de norm gekozen kan worden uit de drie mogelijkheden

$$\|A\|_1 = \max_j \sum_i |A_{ij}|$$

$$\|A\|_\infty = \max_i \sum_j |A_{ij}| = \|A\|_1,$$

$$\|A\|_2 = \max_{x \neq 0} \{ \|Ax\|_2 / \|x\|_2 \} = |\lambda_{\max}|$$

In de laatste formule is $\|x\|_2$ de gewone euclidische norm van de vector x en $\lambda_{\max}$ de in absolute waarde grootste eigenwaarde van de matrix A . Voor de beide andere normen geldt $|\lambda_{\min}| \leq \|A\| \leq |\lambda_{\max}|$, waarbij $\lambda_{\min}$ de in absolute waarde kleinste eigenwaarde is van A . De in (8.11) vermelde bovengrens voor de afwijkingen kan zeer groot worden, als de matrix N kleine eigenwaarden heeft. Het kan daardoor voorkomen, dat wanneer de afwijking naar beneden uitvalt, de berekende eigenwaarde λ_{ck} zelfs nog kleiner wordt dan de werkelijke eigenwaarde λ_k van het probleem (8.1), zodat λ_{ck} zelf beslist geen betrouwbare bovengrens is.

De bovengrens (8.11) voor de afwijking is duidelijk het kleinst voor $\|n\| = 0$. Dat is onder meer het geval wanneer de basisfuncties ϕ_i zo gekozen zijn dat ze orthogonaal zijn t.o.v. de operator $\mathcal{N}$. (Het verdient daarom aanbeveling bij numerieke berekeningen de basisfuncties, als het enigszins kan, steeds orthogonaal, of tenminste bijna orthogonaal te kiezen. De voorbeelden gegeven in [6] spreken in dit opzicht duidelijke taal). Door een simpele hernormering kan de matrix N dan tot de eenheidsmatrix gemaakt worden. Is de maximale fout bij de integratie gelijk aan ϵ , dan vinden we in dit geval

$$(8.12) \quad |\lambda_{uk} - \lambda_{ck}| \leq \|h\| \leq m\epsilon, \quad N + n = I,$$

een schatting, die er in de praktijk meestal niet ver naast is.

De schatting (8.11) bevat nog de onbekende matrix $N^{-\frac{1}{2}}$ en is daarom in de praktijk moeilijk te hanteren. Een wat bruikbaarere schatting, die bovendien nauwer aansluit bij de in de praktijk gevolgde rekenmethode, wordt verkregen door inplaats van met $N^{\frac{1}{2}}$, zoals bij (8.8) en (8.9), te transformeren met de triangulaire matrix L verkregen uit de Cholesky decompositie

$$(8.13) \quad N + n = L^T L$$

Dan krijgen we uit (8.4) en (8.7) de vergelijkingen

$$(8.14) \quad [\underline{H} + \lambda_{\mu} n - \lambda_{\mu} \underline{I}] a = 0,$$

$$(8.15) \quad [\underline{H} + h - \lambda_c \underline{I}] b = 0,$$

waarbij $\tilde{H} = L^{-1T} H L^{-1}$ en analoog voor $\tilde{h}$ en $\tilde{n}$, terwijl $\tilde{a} = La$ en $\tilde{b} = Lb$. $\tilde{H}$, $\tilde{h}$ en $\tilde{n}$ zijn weer symmetrisch.

Voor het verschil van de eigenwaarden krijgen we nu de ongelijkheid

$$(8.16) \quad |\lambda_{\mu} - \lambda_{ck}| \leq \|\tilde{h} - \lambda_{\mu k} \tilde{n}\| \leq \|L^{-1T}\| \cdot \|L^{-1}\| (\|h\| + |\lambda_{\mu k}| \cdot \|n\|).$$

Daar $\lambda_{\mu k} - \lambda_{ck} \leq |\lambda_{\mu k} - \lambda_{ck}|$, volgt hieruit (met $\omega = \|L^{-1T}\| \cdot \|L^{-1}\|$):

$$(8.17) \quad \lambda_{\mu k} \leq \frac{\lambda_{ck} + \omega \|h\|}{1 - \omega \|n\|} \text{ als } \lambda_{\mu k} > 0, \omega \|n\| < 1,$$

$$\lambda_{\mu k} \leq \frac{\lambda_{ck} + \omega \|h\|}{1 + \omega \|n\|} \text{ als } \lambda_{\mu k} < 0.$$

De bovengrens (8.17) voor $\lambda_{\mu k}$, die wegens (8.6) ook een bovengrens is voor de werkelijke eigenwaarde λ_k van (8.1) (dit in tegenstelling tot λ_{ck} zelf), kan op vrij eenvoudige wijze uitgerekend worden, omdat ook bij de numerieke methode om het probleem (8.7) op te lossen matrix L gebruikt wordt om (8.7) te herleiden tot de standaard vorm (8.15), zij het op iets efficiëntere wijze dan door voor- en navermenigvuldiging met L^{-1T} en L^{-1} .

Omdat de schatting (8.16) nogal grof is, ligt deze bovengrens (8.17) meestal hoger dan nodig is. We zullen daarom voor de kleinste eigenwaarde λ_0 , waarin we meestal het meest geïnteresseerd zijn, een lagere bovengrens proberen te berekenen.

8.3. Een andere schatting voor de bovengrens

De kleinste eigenwaarde $\lambda_{\mu 0}$ van de vergelijking (8.4) is de beste bovengrens voor de gezochte kleinste eigenwaarde λ_0 van (8.1) binnen de Rayleigh-Ritz klasse (8.3) van functies χ_m , omdat $\lambda_{\mu 0}$ het minimum is van de functionaal $\Lambda[\chi_m]$. Met de notatie ingevoerd in (8.5) kan deze functionaal ook geschreven worden als

$$(8.18) \quad \Lambda[\chi_m] = \lambda[a] = a H a / a N a.$$

De eigenvector a_0 behorende bij de eigenwaarde $\lambda_{\mu 0}$ is dat stel variatieparameters, waarvoor (8.18) precies zijn minimumwaarde $\lambda_{\mu 0}$ aanneemt. In die zin is de vector a_0 dus optimaal. Evenzo is de eigenvector b_0 gevonden bij de laagste eigenwaarde $\lambda_{c 0}$ van het probleem (8.7) optimaal in die zin, dat daardoor het quotient

$$(8.19) \quad \lambda_c[b] = b^T(H + h) b / b^T(N + n) b$$

geminimaliseerd wordt. Daar echter $\lambda_{c 0}$, tengevolge van de fouten h en n , geen betrouwbare bovengrens is voor de werkelijke eigenwaarde λ_0 , hebben we aan dit optimum niet zoveel. Een betere waarde voor de vector b zou een waarde zijn waardoor het getal (8.19) wel wat groter wordt, maar de fout $|\lambda_{\mu} - \lambda_c|$ zoveel kleiner dat een echte bovengrens, zoals bijv. (8.17), maar dan met de waarde (8.19) voor λ_c gesubstitueerd, zo klein mogelijk wordt. Nog beter is, het als b toch gevarieerd wordt, een bovengrens niet via (8.14) en (8.15), maar direct uit (8.18) en (8.19) af te leiden. Dat zullen we nu doen.

Zij b een willekeurige m -vector en het getal λ_b gedefinieerd door

$$(8.20) \quad \lambda_b = \lambda[b] = b^T H b / b^T N b.$$

Dan geldt, omdat $\lambda_{\mu 0}$ het minimum is van de functionaal (8.18):

$$(8.21) \quad \lambda_{\mu 0} \leq \lambda_b.$$

Voor het verschil in waarden berekend met b uit (8.19) en (8.20) geldt nu:

$$(8.22) \quad \lambda_c - \lambda_b = b^T (h - \lambda_b n) b / b^T P b,$$

waarbij we gemakshalve $\lambda_c = \lambda_c[b] = P = N + n$ geschreven hebben. Uit (8.22) volgt nu als schatting voor de fout

$$(8.23) \quad |\lambda_c - \lambda_b| \leq \|b\|^2 (\|h\| + |\lambda_b| \cdot \|n\|) / {}^T.$$

Voor λ_b zelf vinden we dan de bovengrens

$$(8.24) \quad \lambda_b \leq \frac{\lambda_c + \beta \|b\|}{1 - \beta \|n\|} \quad \text{als } \lambda_b > 0 \text{ en } \beta \|n\| < 1,$$

$$\lambda_b \leq \frac{\lambda_c + \beta \|h\|}{1 + \beta \|n\|} \quad \text{als } \lambda_b < 0,$$

$$(8.25) \quad \beta = \|b\|^2 / b^T P b$$

De bovengrens (8.24), die wegens (8.21) ook een bovengrens is voor λ_0 , lijkt veel op die afgeleid in (8.7). Een voordeel is echter, dat deze bovengrens niet vast ligt, omdat we niet zijn uitgegaan van de uit (8.7) berekende eigenwaarde λ_{c0} . Bovendien zijn de getallen λ_c en β gemakkelijk uit de formule (8.19) en (8.25) te berekenen.

Door b te variëren kan nu getracht worden de bovengrens (8.24) zo klein mogelijk te maken.

Door Delves [2] is een iteratief schema aangegeven om (8.24) te minimaliseren. De resultaten van zijn numerieke berekeningen laten zien dat de uit (8.24) bepaalde bovengrens inderdaad beter is dan de bovengrens (8.17).

Bovenstaand, door Delves voorgestelde, variatiemethode heeft het voordeel dat, inplaats van de onbetrouwbare schatting uit de Rayleigh-Ritz berekening, een echte bovengrens wordt verkregen voor de kleinste eigenwaarde λ_0 van het probleem (8.1) (en door een kleine wijziging ook van de andere eigenwaarden). Het belangrijkste voordeel van de Rayleigh-Ritz methode lag echter in het feit dat het minimum van $\Lambda[\chi_m]$ direct berekend kon worden als de kleinste eigenwaarde van vergelijking (8.4). Bij de methode van Delves moet het minimum weer bepaald worden door variatie van de parameters en dat kost in het algemeen veel meer rekentijd dan een eigenwaardebepaling. Voor het voordeel van de betere bovengrens van Delves moet dus toch weer betaald worden.

Literatuur.

- [1] Courant, R. en Hilbert, P. Methods of mathematical physics,
Vol. I, ch. VI, Interscience,
New York, 1962.

- [2] Delves, L.M. J. Comp. Physics 3(1968), 17-28.

- [3] Gould, S.H. Variational methods for eigenvalue
problems, University of Toronto Press,
Toronto, 1957.

- [4] Kantorowitsch, L.W. en Krylow, W.I. Näherungsmethoden der höheren Analysis
Deutscher Verlag der Wissenschaften,
Berlin, 1956.
Engelse vertaling:
Approximate methods of higher Analysis,
Interscience, New York, 1958.

- [5] Michlin, S.G., Variationsmethoden der Mathematischen
Physik, Akademie-Verlag, Berlin, 1962.

- [6] Michlin, S.G. Numerische Realisierung von Variations-
methoden, Akademie-Verlag, Berlin, 1969.

- [7] Morse, P.M. en Feshbach, H. Methods of Theoretical Physics,
Vol. II, ch. 9.4,
McGraw-Hill, New York etc., 1953.

- [8] Wilkinson, J.H. The algebraic eigenvalue problem,
Clarendon Press, Oxford, 1965.

9. SINGULIERE STORINGSPROBLEMEN VOOR ELLIPTISCHE DIFFERENTIAAL- VERGELIJKINGEN

9.1. Probleemstelling

We beschouwen het volgende randwaardeprobleem:

In een begrensd gebied G van het (x,y) vlak voldoet de functie $\phi_\varepsilon(x,y)$ aan de elliptische differentiaalvergelijking

$$(9.1) \quad L_\varepsilon[\phi_\varepsilon] = \varepsilon L_2[\phi_\varepsilon] + L_1[\phi_\varepsilon] = h(x,y), \quad 0 < \varepsilon \ll 1,$$

waarin de operatoren L_2 en L_1 gedefinieerd zijn door

$$(9.2) \quad L_2 = a(x,y) \frac{\partial^2}{\partial x^2} + 2b(x,y) \frac{\partial^2}{\partial x \partial y} + c(x,y) \frac{\partial^2}{\partial y^2} + d(x,y) \frac{\partial}{\partial x} + \\ + e(x,y) \frac{\partial}{\partial y} + f(x,y)$$

en

$$(9.3) \quad L_1 = p(x,y) \frac{\partial}{\partial x} + q(x,y) \frac{\partial}{\partial y} + g(x,y).$$

Langs de rand Γ van het gebied G voldoet $\phi_\varepsilon(x,y)$ aan de randvoorwaarde

$$(9.4) \quad \phi_\varepsilon(x,y)|_\Gamma = \phi(x,y)|_\Gamma.$$

We maken de volgende veronderstellingen aangaande de coëfficiënten en het rechterlid van (9.1):

- 1^e. $a, b, \dots, f, p, q, g$ en h behoren tot de klasse $C^\alpha(\bar{G})$, d.w.z. deze functies zijn Hölder-continu met exponent α in $G \cup \Gamma$;
- 2^e. voor alle $(x,y) \in G$ in alle $(\xi, \eta) \in \mathbb{R}_2$ geldt

$$a(x,y)\xi^2 + 2b(x,y)\xi\eta + c(x,y)\eta^2 \geq m(\xi^2 + \eta^2)$$
 met m een positieve constante (d.w.z. L_2 is uniform elliptisch);
- 3^e. $g(x,y) < 0$ in $\bar{G}$.

We maken vervolgens de volgende veronderstellingen aangaande de randvoorwaarde:

- 1^e. de rand Γ is glad, d.w.z. er bestaat een eindig aantal cirkels die de

rand Γ bedekken en in elke cirkel kan de rand gerepresenteerd worden door de vergelijking

$$x = g_1(y) \text{ of } y = g_2(x),$$

waarin $g_1(y)$ en $g_2(x)$ Hölder-continue afgeleiden van de tweede orde met exponent α bezitten;

2^e. $\phi(x,y)$ bezit in elke cirkel op de rand Γ als functie van x of y Hölder-continue afgeleiden van de tweede orde met exponent α .

Volgens het existentiebewijs van Schauder ([1], pag. 46 of [2], Hoofdstuk IV, §7.2) bezit dit randwaardeprobleem voor ε positief en voldoende klein ($g + \varepsilon f$ mag niet positief worden in $\bar{G}$) een ondubbelzinnig bepaalde oplossing $\phi_\varepsilon(x,y) \in C_{2+\alpha}(\bar{G})$.

Opmerking 9.1. De eis dat de randfunctie ϕ op Γ glad moet zijn kan afgezwakt worden tot de eis dat ϕ op Γ continu is, indien men gebruik maakt van "interne" Schauder schattingen (zie [1], pag. 50 of [2], Hoofdstuk IV, §7.2).

Na nog enkele beperkingen op te leggen aan de parameters van het probleem (coëfficiënten, rechterleden, vorm van Γ) zullen we in dit hoofdstuk een expliciete asymptotische oplossing van het randwaardeprobleem construeren; we zullen namelijk een functie $\phi_1(x,y;\varepsilon)$ geven met de eigenschap:

$$(9.5) \quad \phi_\varepsilon(x,y) - \phi_1(x,y;\varepsilon) = R_1(x,y;\varepsilon),$$

waarin de restterm $R_1(x,y;\varepsilon)$ "bijna overal" in $\bar{G}$ uniform van de orde $O(\varepsilon)$ is. De uitdrukking "bijna overal" zal later nader verklaard worden.

Dit vraagstuk is voorheen behandeld door Levinson [6], die een minder scherpe schatting voor de restterm aangaf, door Wasow [11], die voor de elliptische differentiaaloperator L_2 de operator van Laplace gebruikte, door Višík en Lyusternik [9] en door Eckhaus en de Jager [3], die het vraagstuk voor algemene elliptische differentiaaloperatoren L_2 behandelden en tot de schatting $O(\varepsilon)$ voor de restterm kwamen.

Naast puntsgewijze locale schattingen voor de restterm beschouwden Višík

en Lyusternik ook globale norm afschattingen voor de norm $\|R\|_{L^2(G)}$ van de restterm; deze laatste zijn echter van geen waarde voor de puntgewijze schattingen van de oplossing $\Phi_\epsilon(x,y)$.

Eckhaus en de Jager hebben de analyse van Visik en Lyusternik op enkele belangrijke punten verbeterd en tevens uitvoerig randwaardeproblemen voor gebieden bestudeerd, die door Visik en Lyusternik slechts zijdelings aangegeven waren (parabolische grenslagen, zie [3], pag. 57-84).

We zullen nu het probleem (9.1) t/m (9.4) formuleren als een singulier storingsprobleem.

De differentiaaloperator $L_1 + \epsilon L_2$ kan men opvatten als een differentiaaloperator L_1 van de eerste orde, die gestoord wordt door een differentiaaloperator ϵL_2 van de tweede orde ($0 < \epsilon \ll 1$).

We nemen aan, dat de differentiaaloperator L_1 nergens in $\bar{G}$ ontaard is in een differentiaaloperator van de orde 0, d.w.z.

$$(9.6) \quad \{p(x,y)\}^2 + \{q(x,y)\}^2 > 0 \text{ in } \bar{G},$$

$p(x,y)$ en $q(x,y)$ zijn gedefinieerd in een gebied $\Omega \supset \bar{G}$ en we nemen voortaan aan, dat zij behoren tot de klasse $C^{1+\alpha}(\Omega)$.

Door overgang op nieuwe orthogonale coördinaten (ξ, η) gedefinieerd door

$$(9.7) \quad \begin{aligned} \frac{\partial \xi}{\partial x} &= -q(x,y), & \frac{\partial \xi}{\partial y} &= +p(x,y), \\ \frac{\partial \eta}{\partial x} &= +p(x,y), & \frac{\partial \eta}{\partial y} &= +q(x,y), \end{aligned}$$

gaat (9.1) over in

$$\epsilon \tilde{L}_2 [\tilde{\Phi}_\epsilon(\xi, \eta)] + \tilde{L}_1 [\tilde{\Phi}_\epsilon(\xi, \eta)] = \tilde{h}(\xi, \eta)$$

met

$$\begin{aligned} \tilde{L}_2 &= \tilde{a}(\xi, \eta) \frac{\partial^2}{\partial \xi^2} + 2\tilde{b}(\xi, \eta) \frac{\partial^2}{\partial \xi \partial \eta} + \tilde{c}(\xi, \eta) \frac{\partial^2}{\partial \eta^2} + \\ &+ \tilde{d}(\xi, \eta) \frac{\partial}{\partial \xi} + \tilde{e}(\xi, \eta) \frac{\partial}{\partial \eta} + \tilde{f}(\xi, \eta) \end{aligned}$$

en

$$\tilde{L}_1 = \frac{\partial}{\partial \eta} + \tilde{g}(\xi, \eta),$$

waarin $\tilde{a}, \dots, \tilde{f}, \tilde{g}$ en $\tilde{h}$ wederom aan de voorwaarden 1^e , 2^e en 3^e van paragraaf 9.1 voldoen.

Men kan nagaan dat de getransformeerde randkromme $\tilde{\Gamma}$ en de getransformeerde randfunctie $\tilde{\phi}(\xi, \eta)|_{\tilde{\Gamma}}$ glad blijven na transformatie en dus ook aan de voorwaarden van de vorige paragraaf blijven voldoen.

Als we (ξ, η) wederom door (x, y) vervangen en de slangetjes weglaten is ons oorspronkelijk probleem zonder de algemeenheid van het probleem te zeer te beperken teruggebracht tot het randwaardeprobleem

$$(9.8) \quad L_\varepsilon[\phi_\varepsilon] = \varepsilon L_2[\phi_\varepsilon] + L_1[\phi_\varepsilon] = h(x, y) \quad \text{voor } (x, y) \in G$$

met

$$L_2 = a \frac{\partial^2}{\partial x^2} + 2b \frac{\partial^2}{\partial x \partial y} + c \frac{\partial^2}{\partial y^2} + d \frac{\partial}{\partial x} + e \frac{\partial}{\partial y} + f$$

en

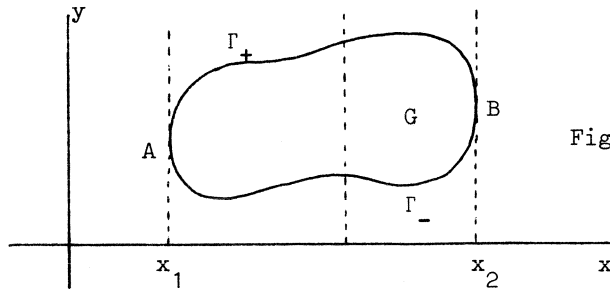
$$L_1 = \frac{\partial}{\partial y} + g,$$

terwijl langs de rand Γ van G geldt

$$(9.9) \quad \phi_\varepsilon(x, y)|_\Gamma = \phi(x, y)|_\Gamma.$$

De coëfficiënten, de rechterleden en de randkromme zijn aan dezelfde condities onderworpen als bij het randwaardeprobleem van §9.1.

Om een asymptotische oplossing van (9.8) en (9.9) te vinden leggen we nu een belangrijke beperking aan het gebied G op. We eisen namelijk, dat er op de rand Γ twee punten (x_1, y_1) en (x_2, y_2) liggen, die de kromme Γ verdeelt in een bovenste helft Γ_+ en een benedenste helft Γ_- en de lijnen $x = C$ (karakteristieken van L_1) met $x_1 < C < x_2$ snijden Γ_- en Γ_+ in slechts één punt (zij mogen niet aan Γ_- of Γ_+ raken). Verder mogen de karakteristieken $x = x_1$ en $x = x_2$ geen deel uitmaken van de rand Γ , de lijnen $x = x_1$ en $x = x_2$ raken aan Γ . (Zie figuur).



Figuur 9.1

De vergelijking van Γ_- zij $y = \gamma_-(x)$ en van Γ_+ zij $y = \gamma_+(x)$. Het is voor de hand liggend om voor een benaderde oplossing van (9.8) - (9.9) te proberen de oplossing van de zogenaamde gereduceerde vergelijking

$$(9.10) \quad L_1[w] = h(x, y) \quad \text{voor } (x, y) \in G,$$

terwijl w voldoet aan de randvoorwaarde langs Γ_- of langs Γ_+ ; w kan niet voldoen aan de randvoorwaarde langs de gehele rand Γ omdat L_1 een differentiaaloperator van de eerste orde is.

Voor

$$(9.11) \quad w(x, \gamma_-(x)) = \phi(x, \gamma_-(x)) \stackrel{\text{def}}{=} \phi_-(x)$$

vindt men

$$(9.12) \quad w^{(1)}(x, y) = \phi_-(x) + \int_{\gamma_-(x)}^y \exp\left[-\int_{\eta}^y g(x, \zeta) d\zeta\right] \{h(x, \eta) - g(x, \eta)\phi_-(x)\} d\eta.$$

Voor

$$(9.13) \quad w(x, \gamma_+(x)) = \phi(x, \gamma_+(x)) \stackrel{\text{def}}{=} \phi_+(x)$$

vindt men

$$(9.14) \quad w^{(2)}(x, y) = \phi_+(x) - \int_y^{\gamma_+(x)} \exp\left[-\int_{\eta}^y g(x, \zeta) d\zeta\right] \{h(x, \eta) - g(x, \eta)\phi_+(x)\} d\eta.$$

Het is duidelijk dat $w^{(1)}$ en $w^{(2)}$ stellig niet bij benadering kunnen voldoen aan het randwaardeprobleem in de omgeving van Γ_+ respectievelijk Γ_- . We konden ook niet verwachten dat $w^{(1)}$ of $w^{(2)}$ een overal in G geldige be-

nadering zou geven voor de oplossing ϕ_ε van ons randwaardeprobleem: immers we hebben $\varepsilon L_2[\phi]$ verwaarloosd t.o.v. $L_1[\phi_\varepsilon]$; dit is echter alleen geoorloofd in dat deel van G waarin $L_1[\phi_\varepsilon]$ groot is t.o.v. $\varepsilon L_2[\phi_\varepsilon]$, maar het is denkbaar dat $L_1[\phi_\varepsilon]$ en $\varepsilon L_2[\phi_\varepsilon]$ ergens in G van dezelfde orde zijn en in dat deel van G kan $w^{(1)}$ of $w^{(2)}$ geen benadering zijn voor ϕ_ε .

We zullen in de volgende paragrafen het deel van G bepalen, waarin $w^{(1)}$ of $w^{(2)}$ een goede benadering is voor ϕ_ε , en tevens in dat deel van G waar $w^{(1)}$ of $w^{(2)}$ geen goede benadering is voor ϕ_ε een correctiefunctie $v(x,y;\varepsilon)$ construeren zodat $w^{(1)} + v(x,y;\varepsilon)$ of $w^{(2)} + v(x,y;\varepsilon)$ daar wel een goede benadering voor ϕ_ε levert.

Het essentiële hulpmiddel dat we hierbij nodig hebben is een consequentie van het z.g.n. maximumprincipe voor elliptische differentiaalvergelijkingen (zie [1], Hoofdstuk 3, §3.1, pag. 35 e.v.; zie ook [7]). Deze consequentie van het maximumprincipe is stelling 3.7 uit [1] (Hoofdstuk 3, §3.1).

Deze stelling luidt als volgt:

Zij u een functie gedefinieerd in een willekeurig begrensde gebied G en behorend tot de klasse $C^2(G)$, die voldoet aan

$$L[u] = M[u] + a(x,y)u$$

met

$$\begin{aligned} M[u] = & a_{11}(x,y) \frac{\partial^2 u}{\partial x^2} + 2a_{12}(x,y) \frac{\partial^2 u}{\partial x \partial y} + a_{22}(x,y) \frac{\partial^2 u}{\partial y^2} + a_1(x,y) \frac{\partial u}{\partial x} + \\ & + a_2(x,y) \frac{\partial u}{\partial y} \end{aligned}$$

waarin a_{11} , a_{12} , a_{22} , a_1 , a_2 en a behoren tot $C^0(G)$,

$$a_{11}(x,y)\xi^2 + 2a_{12}(x,y)\xi\eta + a_{22}(x,y)\eta^2 \geq m(\xi^2 + \eta^2), \quad m > 0, \quad \forall (x,y) \in G$$

en $a(x,y) \leq 0$ in G .

Onder deze voorwaarden volgt voor functies $v \in C^2(G)$ uit

$$|L[u]| < L[-v] \text{ in } G \text{ en } |u| \leq v \text{ op } \partial G$$

de relatie

$$|u| \leq v \text{ in } \bar{G}.$$

De functie $v(x,y)$ heet een barrièrefunctie voor de functie $u(x,y)$.

9.2. Toepassingen van het maximumprincipe op singuliere storingsproblemen

Stelling 9.1. De oplossing van het randwaardeprobleem (9.8) - (9.9) is voor voldoende kleine waarde van ε begrensd in $\bar{G}$ en wel uniform met betrekking tot ε .

Bewijs. Er bestaat een getal $0 < \alpha^2 < 1$ en een getal $\varepsilon_0 > 0$ met de eigenschap, dat

$$1 + \varepsilon e(x,y) > \alpha^2 \quad \text{in } \bar{G},$$

$$g(x,y) + \varepsilon f(x,y) \leq 0 \quad \text{in } \bar{G},$$

geldig voor alle ε met $\varepsilon < \varepsilon_0$.

Verder bestaat er een positief getal p onafhankelijk van ε met de eigenschap

$$|h(x,y)| \leq p \quad \text{in } \bar{G},$$

$$|\phi(x,y)| \leq p \quad \text{langs } \Gamma.$$

De functie

$$v(x,y) = \frac{p}{\alpha^2} (y_0 - y) + p$$

met $y_0 = \max_{\Gamma} y$ is een barrièrefunctie van $\phi_{\varepsilon}(x,y)$ voor $\varepsilon < \varepsilon_0$, zoals gemakkelijk is na te gaan. Derhalve

$$(9.15) \quad |\phi_{\varepsilon}(x,y)| \leq v(x,y) \leq Kp \quad \text{voor } \varepsilon < \varepsilon_0 \text{ en } (x,y) \in \bar{G},$$

met $K = \frac{1}{\alpha^2} \max_{\Gamma} (y_0 - y) + 1$.

We beschouwen wederom het randwaardeprobleem (9.8) - (9.9). We veronderstellen dat de bovenrand $y = \gamma_+(x)$ de eigenschap bezit, dat er een rechter omgeving Δ_1 en een linker omgeving Δ_2 van $x = x_1$ respectievelijk $x = x_2$ bestaan, zodanig dat

$$(9.16) \quad \frac{d\gamma_+}{dx} \geq 0 \quad \text{voor } x \in \Delta_1, \quad \frac{d\gamma_+}{dx} < 0 \quad \text{voor } x \in \Delta_2$$

en

$$(9.17) \quad \frac{d^2\gamma_+}{dx^2} \leq 0 \quad \text{voor } x \in \Delta_1 \cup \Delta_2$$

$$\Delta_1 \cap \Delta_2 = \emptyset.$$

Stelling 9.2. De oplossing ϕ_ε van het randwaardeprobleem (9.8) - (9.9), waarbij de bovenrand voldoet aan de relaties (9.16) en (9.17), heeft de eigenschap dat voor voldoende kleine waarden van ε geldt

$$(9.18) \quad |\phi_\varepsilon(x, y) - \phi_+(x)| \leq M\{\gamma_+(x) - y\}, \quad \text{voor } (x, y) \in \bar{G},$$

waarin M een constante is onafhankelijk van ε .

Bewijs. We beschouwen de functie

$$\phi_\varepsilon^*(x, y) = \phi_\varepsilon(x, y) - \phi_+(x).$$

Derhalve voldoet $\phi_\varepsilon^*(x, y)$ aan het randwaardeprobleem

$$(9.19) \quad L_\varepsilon[\phi_\varepsilon^*(x, y)] = h(x, y) - L_\varepsilon[\phi_+(x)] = h^*(x, y)$$

met de randvoorwaarden

$$\phi_\varepsilon^*(x, y)|_{\Gamma_+} = 0$$

en

$$\phi_\varepsilon^*(x, y)|_{\Gamma_-} = \phi_-(x) - \phi_+(x).$$

We voeren nu in de volgende functie

$$(9.20) \quad \Psi(x, y) = M_0 \{ \gamma_+(x) - y \} \chi(x),$$

waarin M_0 een positieve constante is en $\chi(x)$ een nader te bepalen positieve tweemaal continu differentieerbare functie op $[x_1, x_2]$; M_0 en $\chi(x)$ zijn onafhankelijk van ε en we zullen aantonen dat M_0 en $\chi(x)$ zodanig gekozen kunnen worden dat $\Psi(x, y)$ een barrièrefunctie is voor $\phi_\varepsilon^*(x, y)$, indien ε voldoende klein genomen wordt.

Door $\chi(x) > 1$ en door M_0 voldoende groot te kiezen, namelijk

$$(9.21) \quad M_0 > \max_{x_1 \leq x \leq x_2} \left| \frac{\phi(x, \gamma_+(x)) - \phi(x, \gamma_-(x))}{\gamma_+(x) - \gamma_-(x)} \right|$$

kan men er voor zorgen dat geldt

$$(9.22) \quad |\phi_\varepsilon^*(x, y)|_\Gamma \leq \Psi(x, y)|_\Gamma, \quad \varepsilon > 0.$$

We passen nu L_ε toe op $-\Psi(x, y)$:

$$\begin{aligned} L_\varepsilon [-\Psi(x, y)] &= \varepsilon M_0 [-a(x, y) \{ \gamma_+(x) - y \} \frac{d^2 \chi}{dx^2} - 2a(x, y) \frac{d\gamma_+}{dx} \frac{d\chi}{dx} \\ &\quad - a(x, y) \frac{d^2 \gamma_+}{dx^2} \chi(x) + 2b(x, y) \frac{d\chi}{dx} - d(x, y) \{ \gamma_+(x) - y \} \frac{d\chi}{dx} \\ &\quad - d(x, y) \frac{d\gamma_+}{dx} \chi(x) + e(x, y) \chi(x) - f(x, y) \{ \gamma_+(x) - y \} \chi(x)] \\ &\quad + M_0 \chi(x) - M_0 g(x, y) \{ \gamma_+(x) - y \} \chi(x). \end{aligned}$$

Alle termen tussen de vierkante haken zijn continu in $\bar{G}$, behalve

$$-2a \frac{d\gamma_+}{dx} \frac{d\chi}{dx}, \quad -a \frac{d^2 \gamma_+}{dx^2} \chi(x) \quad \text{en} \quad -d \frac{d\gamma_-}{dx} \chi(x).$$

Ons herinnerend dat $g(x, y) < 0$ is in G vinden we voor voldoende kleine waarden

van ε (stel $\varepsilon < \varepsilon_1$)

$$(9.23) \quad L_\varepsilon[-\psi] > \frac{1}{2} M_0 \chi(x) - \varepsilon M_0 \left(2a \frac{d\chi}{dx} + d \chi \right) \frac{d\gamma_+}{dx} \\ - \varepsilon M_0 a \frac{d^2 \gamma_+}{dx^2} \chi(x).$$

In $\bar{H} = \overline{G - \{(x,y) \in G; x \in \Delta_1\}}$ zijn alle termen in het rechterlid continu en dus begreind; derhalve geldt in $\bar{H}$ voor ε voldoende klein ($\varepsilon < \varepsilon_2 \leq \varepsilon_1$)

$$(9.24) \quad L_\varepsilon[-\psi] > \frac{1}{3} M_0 \chi(x) > \frac{1}{3} M_0, \quad (x,y) \in \bar{H}.$$

We kiezen nu $\chi(x)$ zodanig dat geldt

$$(9.25) \quad 2(\min a(x,y)) \frac{d\chi}{dx} + \max |d(x,y)| \chi = 0 \quad \text{in } \bar{\Delta}_1 = [x_1, x_1 + \delta_1]$$

met $\chi(x_1 + \delta_1) = 2$ en het minimum en maximum is genomen over $(x,y) \in G$ met $x_1 \leq x \leq x_1 + \delta_1$.

Uit (9.25) volgt $\frac{d\chi}{dx} \leq 0$ in $\bar{\Delta}_1$ en de eis $\chi(x) > 1$ blijft gehandhaafd in $\bar{\Delta}_1$. Tevens geldt er

$$(9.26) \quad 2a(x,y) \frac{d\chi}{dx} + d(x,y) \chi \leq 0 \quad \text{in } \bar{\Delta}_1.$$

Analoog wordt $\chi(x)$ in $\bar{\Delta}_2$ zodanig gekozen, dat

$$(9.27) \quad 2(\min a(x,y)) \frac{d\chi}{dx} - \max |d(x,y)| \chi = 0 \quad \text{in } \bar{\Delta}_2 = [x_2 - \delta_2, x_2]$$

met $\chi(x_2 - \delta_2) = 2$.

Uit (9.27) volgt weer $\frac{d\chi}{dx} \geq 0$ in $\bar{\Delta}_2$ en de eis $\chi(x) > 1$ blijft gehandhaafd in $\bar{\Delta}_2$. Tevens geldt er

$$(9.28) \quad 2a(x,y) \frac{d\chi}{dx} + d(x,y) \chi(x) \geq 0 \quad \text{in } \bar{\Delta}_2.$$

Substitueren we tenslotte (9.26) en (9.28) in (9.23) en maken we gebruik van de voorwaarden (9.16) en (9.17) dan vinden we

$$(9.29) \quad L_\varepsilon[-\psi] > \frac{1}{2} M_0 \chi(x) > \frac{1}{2} M_0 \quad \text{voor } (x,y) \in \Delta_i \quad (i = 1,2).$$

Combinatie van (9.24) en (9.29) geeft tenslotte

$$(9.30) \quad L_\varepsilon[-\psi] > \frac{1}{3} M_0, \quad \text{voor alle } (x,y) \in G \text{ en voor } 0 < \varepsilon < \varepsilon_2.$$

Vergroten we eventueel M_0 tot een waarde met

$$M_0 > 3 \sup_{\substack{(x,y) \in G \\ \varepsilon < \varepsilon_2}} |h_\varepsilon^*(x,y)|$$

dan vinden we uiteindelijk

$$(9.31) \quad L_\varepsilon[-\psi] > |L_\varepsilon[\phi_\varepsilon^*(x,y)]| \quad \text{in } G \text{ voor } \varepsilon < \varepsilon_2.$$

Dit resultaat gevoegd bij (9.22) maakt $\psi(x,y)$ tot een barrièrefunctie voor $\phi_\varepsilon^*(x,y)$ en daarmee is de stelling bewezen.

Opmerking 9.2. Een formule analoog aan (9.18) met

$$(9.32) \quad |\phi_\varepsilon(x,y) - \phi_-(x)| \leq M\{y-\gamma_-(x)\}, \quad \forall (x,y) \in \bar{G}$$

kan niet op soortgelijke wijze, als boven geschied is, bewezen worden. Immers voor

$$\psi(x,y) = M_0\{y-\gamma_-(x)\} \chi(x)$$

wordt het analogon van (9.23)

$$(9.33) \quad L_\varepsilon[-\psi] > \frac{1}{2} M_0 \chi(x) + \varepsilon M_0 \left\{ 2a(x,y) \frac{d\chi}{dx} + d(x,y) \chi(x) \right\} \frac{d\gamma_-}{dx} \\ + \varepsilon M_0 a(x,y) \frac{d^2 \gamma_-}{dx^2} \chi(x).$$

Op grond van het - teken in de term $-\frac{1}{2} M_0 \chi(x)$ van het rechterlid kan $L_\varepsilon[-\psi]$ niet door een positieve ondergrens afgeschat worden en derhalve kan

niet meer voldaan worden aan de relatie:

$$L_{\varepsilon}[-\Psi] > |L_{\varepsilon}[\phi_{\varepsilon}(x,y) - \phi_{-}(x)]|,$$

waarmee de conclusie vervalst dat $\Psi(x,y)$ een barrièrefunctie is voor $\phi_{\varepsilon}(x,y) - \phi_{-}(x)$.

9.3. De benadering voor de functie $\phi_{\varepsilon}(x,y)$

Uit de afschatting (9.18) volgt dat $\frac{\partial \phi_{\varepsilon}}{\partial y}$ voor voldoende kleine waarden van ε uniform begrensd is langs de bovenrand Γ_{+} van het gebied G . We kunnen echter niet de conclusie trekken dat $\frac{\partial \phi_{\varepsilon}}{\partial y}$ voor voldoende kleine waarden van ε ook uniform begrensd is langs de benedenrand Γ_{-} van het gebied G (zie opmerking aan het slot van de vorige paragraaf). Het is dus heel goed denkbaar, dat $\frac{\partial \phi_{\varepsilon}}{\partial y}$ willekeurig groot wordt in een "bovenomgeving" van Γ_{-} naarmate ε kleiner wordt; in dit geval is het onmogelijk dat $\phi_{\varepsilon}(x,y)$ in een "bovenomgeving" van Γ_{-} bij benadering voldoet aan de gereduceerde vergelijking $L_1[w] = h(x,y)$ en de conditie dat $\phi_{\varepsilon}(x,y)$ uniform begrensd is in G (stelling 9.1). Derhalve zal in het geval, dat $\frac{\partial \phi_{\varepsilon}}{\partial y}$ willekeurig groot wordt in een "bovenomgeving" van Γ_{-} voor $\varepsilon \rightarrow 0$, de onbekende functie $\phi_{\varepsilon}(x,y)$ niet benaderd worden door de oplossing $w^{(1)}(x,y)$ van de gereduceerde vergelijking (zie formules (9.10) en (9.12)).

We stellen daarom

$$(9.34) \quad \phi_{\varepsilon}(x,y) = w(x,y) + v_{\varepsilon}(x,y),$$

waarin

$$(9.35) \quad w(x,y) = w^{(2)}(x,y) = \\ = \phi_{+}(x) - \int_y^{\gamma_{+}(x)} \exp\left[-\int_n^y g(x,\zeta) d\zeta\right] \{h(x,n) - g(x,n)\phi_{+}(x)\} dn ;$$

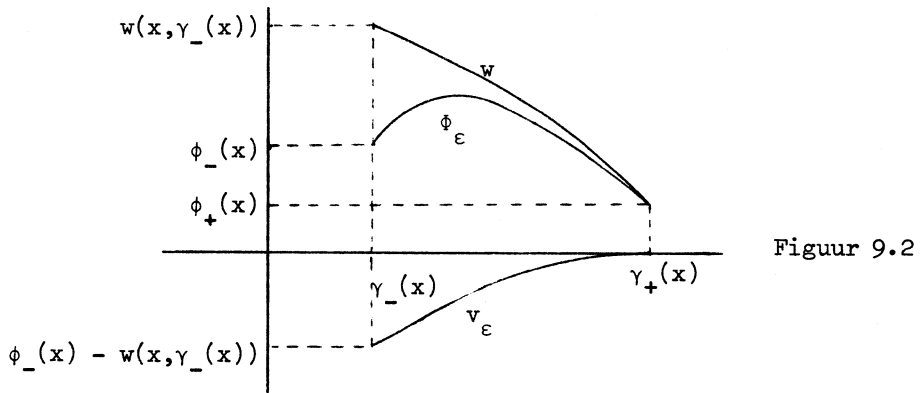
$v_{\varepsilon}(x,y)$ is een correctie op $w(x,y)$, die in G bij benadering moet voldoen aan de differentiaalvergelijking

$$(9.36) \quad L_\varepsilon [v(x,y)] = L_\varepsilon [\phi_\varepsilon] - L_\varepsilon [w] = -\varepsilon L_2 [w(x,y)]$$

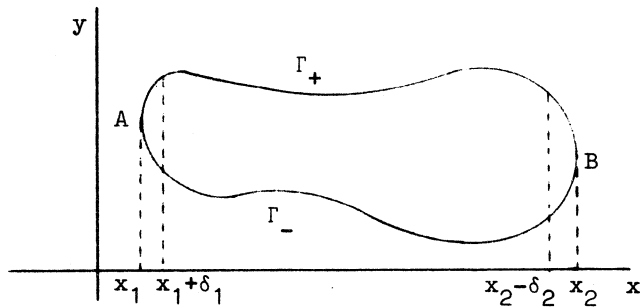
met randvoorwaarden

$$(9.37) \quad \begin{cases} v_\varepsilon(x,y)|_{\Gamma_+} = 0, \\ v_\varepsilon(x,y)|_{\Gamma_-} = \phi_-(x) - w(x, \gamma_-(x)). \end{cases}$$

Beschouwen we de functies $\phi_\varepsilon(x,y)$, $w(x,y)$ en $v_\varepsilon(x,y)$ voor $x = \text{constant}$ met $x_1 < x < x_2$, dan krijgen we ongeveer het volgende beeld:



Het rechterlid van (9.36) is singulier in een rechter omgeving $x = x_1$, en een linker omgeving van $x = x_2$ ($\gamma_+(x)$ is niet differentieerbaar in $x = x_1$ en $x = x_2$); daarom zullen we van nu af aan vaste willekeurig kleine omgevingen van de randpunten $A(x_1, y_1)$ en $B(x_2, y_2)$ van onze analyse uitsluiten. We beschouwen daarom voortaan het gebied $\tilde{G} = G - V(A) - V(B) \stackrel{\text{def}}{=} \{(x,y); (x,y) \in G \text{ en } x_1 + \delta_1 \leq x \leq x_2 - \delta_2\}$ (zie figuur 9.3).



Figuur 9.3

We weten nu: $w(x,y)$ en $v_\epsilon(x,y)$ voldoen in $\tilde{G}$ aan de partiële differentiaalvergelijkingen

$$(9.38) \quad \begin{cases} L_\epsilon [w(x,y)] = h(x,y) + O(\epsilon), \\ L_\epsilon [v_\epsilon(x,y)] = O(\epsilon) \end{cases}$$

en aan de randvoorwaarden

$$(9.39) \quad \begin{cases} w(x,y)|_{\tilde{\Gamma}_+} = \phi_+(x) ; \tilde{\Gamma}_+ = \Gamma_+ \cap \tilde{G}, \\ v_\epsilon(x,y)|_{\tilde{\Gamma}_+} = 0, \\ v_\epsilon(x,y)|_{\tilde{\Gamma}_-} = \phi_-(x) - w(x, \gamma_-(x)) ; \tilde{\Gamma}_- = \Gamma_- \cap \tilde{G}. \end{cases}$$

Onze taak is nu het randwaardeprobleem voor $v_\epsilon(x,y)$ in voldoende benadering op te lossen.

Omdat de mogelijkheid bestaat, dat de functie $\frac{\partial \phi_\epsilon}{\partial y}$ niet uniform in ϵ begrensd is in de omgeving van de onderrand, bestaat deze mogelijkheid ook voor de functie $\frac{\partial v_\epsilon}{\partial y}$; we zullen daarom bij de oplossing van het randwaardeprobleem voor $v_\epsilon(x,y)$ bijzondere aandacht besteden aan het gedrag van $v_\epsilon(x,y)$ in de "bovenomgeving" van de onderrand. Voor dit doel is het dienstig de volgende coördinatentransformatie in te voeren:

$$(9.40) \quad \begin{cases} \xi = x, \\ \eta = y - \gamma_-(x). \end{cases}$$

De partiële differentiaalvergelijking $L_\epsilon [v_\epsilon(x,y)] = O(\epsilon)$ gaat nu over in

$$(9.41) \quad \begin{cases} \epsilon [\alpha(\xi, \eta) \frac{\partial^2 v_\epsilon}{\partial \xi^2} + 2\beta(\xi, \eta) \frac{\partial^2 v_\epsilon}{\partial \xi \partial \eta} + \gamma(\xi, \eta) \frac{\partial^2 v_\epsilon}{\partial \eta^2} + x(\xi, \eta) \frac{\partial v_\epsilon}{\partial \xi} + \\ \quad + \lambda(\xi, \eta) \frac{\partial v_\epsilon}{\partial \eta} + \mu(\xi, \eta) v_\epsilon] + \frac{\partial v_\epsilon}{\partial \eta} + v(\xi, \eta) v_\epsilon = O(\epsilon) \end{cases}$$

met

$$(9.42) \left\{ \begin{array}{l} \alpha(\xi, \eta) = a\{\xi, \eta + \gamma_-(\xi)\}, \\ 2\beta(\xi, \eta) = 2b\{\xi, \eta + \gamma_-(\xi)\} - a\{\xi, \eta + \gamma_-(\xi)\}\gamma'_+(\xi), \\ \gamma(\xi, \eta) = a\{\xi, \eta + \gamma_-(\xi)\}(\gamma'_-(\xi))^2 - 2b\{\xi, \eta + \gamma_-(\xi)\}\gamma'_-(\xi) + c\{\xi, \eta + \gamma_-(\xi)\}, \\ x(\xi, \eta) = d\{\xi, \eta + \gamma_-(\xi)\}, \\ \lambda(\xi, \eta) = e\{\xi, \eta + \gamma_-(\xi)\} - d\{\xi, \eta + \gamma_-(\xi)\}\gamma'_-(\xi) - a\{\xi, \eta + \gamma_-(\xi)\}\gamma''_-(\xi), \\ \mu(\xi, \eta) = f\{\xi, \eta + \gamma_-(\xi)\}, \\ v(\xi, \eta) = g\{\xi, \eta + \gamma_-(\xi)\}. \end{array} \right.$$

Opmerking: onderscheid goed de coëfficiënt $\gamma(\xi, \eta)$ en de functies $\gamma_{\pm}(\xi)$, die de boven- en onderrand representeren. Voor $x_1 + \delta_1 = \xi_1 + \delta_1 \leq \xi \leq \xi_2 - \delta_2 = x_2 - \delta_2$ zijn de coëfficiënten van de partiële differentiaalvergelijking (9.41) continu in $\bar{G}$ en de transformatie mocht veilig toegepast worden. Om een scherp inzicht te verkrijgen van het gedrag van $v_{\varepsilon}(\xi, \eta)$ in de boven-omgeving van de onderrand $\eta = 0$ stellen we

$$(9.43) \quad \eta = \varepsilon t.$$

Verder veronderstellen we dat in $\bar{G}$ de coëfficiënten $a(x, y)$, $b(x, y)$ en $c(x, y)$ tweemaal continu differentieerbaar zijn naar y en dat de coëfficiënten $d(x, y)$, $e(x, y)$ en $g(x, y)$ in $\bar{G}$ éénmaal continu differentieerbaar zijn naar y . Uit (9.42) volgt dan dat $\beta(\xi, \eta)$, $\lambda(\xi, \eta)$ en $v(\xi, \eta)$ in $\bar{G}$ éénmaal continu differentieerbaar zijn naar η en dat $\gamma(\xi, \eta)$ in $\bar{G}$ tweemaal continu differentieerbaar is naar η .

De operator L_{ε} kan nu m.b.v. Taylor-ontwikkeling van de coëfficiënten ontwikkeld worden als

$$(9.44) \quad L_{\varepsilon} = \frac{1}{\varepsilon} M_{-1} + M_0 + \varepsilon M_1$$

met

$$(9.45) \quad \left\{ \begin{aligned} M_{-1} &= \gamma(\xi, 0) \frac{\partial^2}{\partial t^2} + \frac{\partial}{\partial t}, \\ M_0 &= 2\beta(\xi, 0) \frac{\partial^2}{\partial \xi \partial t} + t \frac{\partial \gamma}{\partial \eta}(\xi, 0) \frac{\partial^2}{\partial t^2} + \lambda(\xi, 0) \frac{\partial}{\partial t} + v(\xi, 0), \\ M_1 &= \alpha(\xi, \epsilon t) \frac{\partial^2}{\partial t^2} + 2t \frac{\partial \beta}{\partial \eta}(\xi, \theta_1 \epsilon t) \frac{\partial^2}{\partial \xi \partial t} + t^2 \frac{\partial^2}{\partial \eta^2}(\xi, \theta_2 \epsilon t) \frac{\partial^2}{\partial t^2} \\ &\quad + x(\xi, \epsilon t) \frac{\partial}{\partial \xi} + t \frac{\partial \lambda}{\partial \eta}(\xi, \theta_3 \epsilon t) \frac{\partial}{\partial t} + \mu(\xi, \epsilon t) + t \frac{\partial v}{\partial \eta}(\xi, \theta_4 \epsilon t), \end{aligned} \right.$$

waarin $0 < \theta_i < 1$ ($i = 1, 2, 3, 4$).

We voeren nu in de volgende functie

$$(9.46) \quad v(\xi, t; \epsilon) = v_0(\xi, t; \epsilon) + \epsilon v_1(\xi, t; \epsilon)$$

met

$$(9.47) \quad M_{-1}[v_0] = \gamma(\xi, 0) \frac{\partial^2 v_0}{\partial t^2} + \frac{\partial v_0}{\partial t} = 0$$

en

$$(9.48) \quad M_{-1}[v_1] = \gamma(\xi, 0) \frac{\partial^2 v_1}{\partial t^2} + \frac{\partial v_1}{\partial t} = -M_0[v_0],$$

en de randvoorwaarden

$$(9.49) \quad v_0(\xi, 0; \epsilon) = \phi_-(\xi) - w(\xi, \gamma_-(\xi)), \quad \xi_1 + \delta_1 \leq \xi \leq \xi_2 - \delta_2,$$

$$(9.50) \quad \lim_{t \rightarrow \infty} v_0(\xi, t; \epsilon) = 0, \quad \xi_1 + \delta_1 \leq \xi \leq \xi_2 - \delta_2$$

en

$$(9.51) \quad v_1(\xi, 0; \epsilon) = 0, \quad \xi_1 + \delta_1 \leq \xi \leq \xi_2 - \delta_2,$$

$$(9.52) \quad \lim_{t \rightarrow \infty} v_1(\xi, t; \epsilon) = 0, \quad \xi_1 + \delta_1 \leq \xi \leq \xi_2 - \delta_2.$$

Uit (9.44), (9.46), (9.47) en (9.48) volgt

$$\begin{aligned}
L_\varepsilon [v(\xi, t; \varepsilon)] &= L_\varepsilon [v_0 + \varepsilon v_1] = \frac{1}{\varepsilon} M_{-1} [v_0] + M_{-1} [v_1] + M_0 [v_0] + \\
(9.53) \quad &+ \varepsilon M_0 [v_1] + \varepsilon M_1 [v_0] + \varepsilon^2 M_1 [v_1] = \\
&= \varepsilon \{M_0 [v_1] + M_1 [v_0] + \varepsilon M_1 [v_1]\},
\end{aligned}$$

terwijl voor de randvoorwaarden van $v(\xi, t; \varepsilon)$ geldt:

$$(9.54) \quad v(\xi, 0; \varepsilon) = \phi_-(\xi) - w(\xi, \gamma_-(\xi)), \quad \xi_1 + \delta_1 \leq \xi \leq \xi_2 - \delta_2$$

en

$$(9.55) \quad \lim_{t \rightarrow \infty} v(\xi, 0; \varepsilon) = 0, \quad \xi_1 + \delta_1 \leq \xi \leq \xi_2 - \delta_2.$$

Indien v_0 en v_1 voldoende vaak continu differentieerbaar zijn, dan voldoet $v(\xi, t; \varepsilon)$ tot op $O(\varepsilon)$ aan de differentiaalvergelijking, geldig voor $v_\varepsilon(x, y)$ (zie (9.38)), terwijl de randvoorwaarde voor $v(\xi, t; \varepsilon)$ langs de benedenrand ($t = 0$) dezelfde is als die voor $v_\varepsilon(x, y)$ langs de benedenrand ($y = \gamma_-(x)$); zie (9.39).

De randvoorwaarde $\lim_{t \rightarrow \infty} v(\xi, t; \varepsilon) = 0$ is een goede benadering voor de randvoorwaarde $v(\xi, \frac{\gamma_+(\xi) - \gamma_-(\xi)}{\varepsilon}; \varepsilon) = 0$, welke dezelfde is als die voor $v_\varepsilon(x, y)$ langs de bovenrand $y = \gamma_+(x)$; zie (9.39).

We kunnen dus verwachten dat de functie

$$v(\xi, t; \varepsilon) = v_0(\xi, t; \varepsilon) + v_1(\xi, t; \varepsilon)$$

een goede benadering zal leveren voor de gezochte correctieterm $v_\varepsilon(x, y)$. Bij deze analyse wordt het gedrag van $v_\varepsilon(x, y)$ in de omgeving van de onder- rand nauwkeurig onderzocht. We zullen nu achtereenvolgens de functies $v_0(\xi, t; \varepsilon)$ en $v_1(\xi, t; \varepsilon)$ berekenen.

Uit (9.47), (9.49) en (9.50) volgt onmiddellijk

$$(9.56) \quad v_0(\xi, t; \varepsilon) = [\phi_-(\xi) - w(\xi, \gamma_-(\xi))] \exp [-\sigma(\xi) \cdot t]$$

met $\sigma(\xi) = \frac{1}{\gamma(\xi, 0)}$. Er zij opgemerkt dat de functie $\gamma(\xi, \eta)$ uit (9.42) groter dan nul is in $\bar{G}$ omdat de differentiaaloperator L_2 uniform elliptisch is in

G (zie §9.1) en derhalve $\gamma(\xi, 0) > 0$.

De functie

$$(9.57) \quad v_0(\xi, \frac{\eta}{\epsilon}; \epsilon) = [\phi_-(\xi) - w(\xi, \gamma_-(\xi))] \exp[-\sigma(\xi) \frac{\eta}{\epsilon}]$$

is dus overal in $\tilde{G}$ asymptotisch gelijk aan nul behalve langs de benedenrand $\eta = 0$. De functie $v_0(\xi, \eta; \epsilon)$ heeft een zogenaamd typisch grenslaag karakter en heet een "grenslaag"-functie.

Stellen we

$$(9.58) \quad \phi_-(\xi) - w(\xi, \gamma_-(\xi)) = A(\xi)$$

dan vinden we na een weinig rekenwerk

$$M_0[v_0] = \{B(\xi) + tC(\xi)\} \exp[-\sigma(\xi) \cdot t]$$

met

$$(9.59) \quad \begin{aligned} B(\xi) = & -2B(\xi, 0) \left\{ \frac{dA}{d\xi} \frac{1}{\gamma(\xi, 0)} - \frac{A(\xi)}{\{\gamma(\xi, 0)\}^2} \frac{d}{d\xi} \gamma(\xi, 0) \right\} \\ & - \lambda(\xi, 0) \frac{A(\xi)}{\gamma(\xi, 0)} + v(\xi, 0) A(\xi) \end{aligned}$$

en

$$(9.60) \quad C(\xi) = -2B(\xi, 0) \frac{A(\xi)}{\{\gamma(\xi, 0)\}^3} \frac{d}{d\xi} \gamma(\xi, 0) + \frac{\partial \gamma}{\partial \eta}(\xi, 0) \frac{A(\xi)}{\{\gamma(\xi, 0)\}^2},$$

en derhalve met behulp van (9.48), (9.51) en (9.52)

$$(9.61) \quad \begin{cases} v_1(\xi, t; \epsilon) = [\{\gamma(\xi, 0) C(\xi) + B(\xi)\}t + \frac{1}{2} C(\xi) t^2] \exp[-\sigma(\xi)t] \\ \text{ofwel} \\ v_1(\xi, \frac{\eta}{\epsilon}; \epsilon) = [\{\gamma(\xi, 0) C(\xi) + B(\xi)\} \frac{\eta}{\epsilon} + \frac{1}{2} C(\xi) \frac{\eta^2}{\epsilon^2}] \exp[-\sigma(\xi) \frac{\eta}{\epsilon}]. \end{cases}$$

Uit (9.61) blijkt dat $v_1(\xi, \eta; \epsilon)$ in $\tilde{G}$ uniform begrensd is en dat ook $v_1(\xi, \eta; \epsilon)$ het grenslaagkarakter bezit.

We verwachten, dat $v_0 + \epsilon v_1$ een goede benadering is voor de correctiefunctie $v_\epsilon(x, y)$, die tezamen met de functie $w(x, y)$ een goede benadering voor de

gezochte functie $\phi_\varepsilon(x, y)$ moet opleveren. We stellen daarom nu

$$(9.62) \quad \phi_\varepsilon(x, y) = w(x, y) + \tilde{v}_0(x, y; \varepsilon) + \varepsilon \tilde{v}_1(x, y; \varepsilon) + R(x, y; \varepsilon),$$

met

$$\tilde{v}_0(x, y; \varepsilon) = v_0\left(x, \frac{y - \gamma_-(x)}{\varepsilon}; \varepsilon\right)$$

en

$$\tilde{v}_1(x, y; \varepsilon) = v_1\left(x, \frac{y - \gamma_-(x)}{\varepsilon}; \varepsilon\right).$$

De restterm $R(x, y; \varepsilon)$ voldoet aan de volgende voorwaarden:

$$(9.63) \quad \begin{aligned} L_\varepsilon[R] &= L_\varepsilon[\phi_\varepsilon - w(x, y)] - L_\varepsilon[\tilde{v}_0 + \varepsilon \tilde{v}_1] = \\ &= -\varepsilon\{L_2[w] + M_0[v_1] + M_1[v_0] + \varepsilon M_1[v_1]\} \end{aligned}$$

geldig in $\tilde{G}$, (zie (9.53)), en

$$(9.64) \quad R(x, y; \varepsilon)|_{\tilde{\Gamma}_+} = O(\varepsilon^n), \text{ (n willekeurig groot natuurlijk getal),}$$

$$(9.65) \quad R(x, y; \varepsilon)|_{\tilde{\Gamma}_-} = 0.$$

De eerste drie termen van het rechterlid (9.62) leveren een goede benadering voor $\phi_\varepsilon(x, y)$ in het gebied $\tilde{G}$, indien we uit (9.63), (9.64) en (9.65) kunnen afleiden dat $R(x, y; \varepsilon)$ in $\tilde{G}$ uniform naar nul nadert voor $\varepsilon \rightarrow 0$. Dit zal in de volgende paragraaf bewezen worden.

9.4. Afschatting van de restterm

De restterm wordt afgeschat door weer gebruik te maken van een geschikte barrièrefunctie voor $R(x, y; \varepsilon)$; hiertoe is het evenwel noodzakelijk eerst de condities op te stellen, waaronder het rechterlid van (9.63) van de orde $O(\varepsilon)$ is. Hiervoor is voldoende de uniforme begrensdsheid in $\tilde{G}$ van $L_2[w]$, $M_0[v_1]$, $M_1[v_0]$ en $M_1[v_1]$.

Deze uniforme begrensdsheid zal ons voeren tot extra condities voor de coëf-

ficiënten van de oorspronkelijke differentiaalvergelijking, de randwaardefuncties $\phi_+(x)$ en de vergelijkingen $y = \gamma_+(x)$ voor boven- en benedenrand. Tot nu toe hadden we voor de parameters van het randwaardeprobleem geëist:

- $$\left. \begin{array}{ll} \text{(i)} & a(x,y), b(x,y), \dots, f(x,y), g(x,y) \text{ en } h(x,y) \text{ behoren tot } C^\alpha(\bar{G}), \\ \text{(ii)} & a(x,y)\xi^2 + 2b(x,y)\xi\eta + c(x,y)\eta^2 \geq m(\xi^2 + \eta^2), \quad m > 0 \\ \text{(iii)} & g(x,y) < 0 \text{ in } \bar{G} \\ \text{(iv)} & \gamma_+(x) \in C^{2+\alpha}([x_1, x_2]) \\ \text{(v)} & \phi_+(x) \in C^{2+\alpha}([x_1, x_2]) \end{array} \right\} \text{ zie 9.1.}$$
- (vi) $a(x,y)$, $b(x,y)$ en $c(x,y)$ tweemaal continu differentieerbaar naar y in $\bar{G}$ (noodzakelijk voor de ontwikkeling van $L_\varepsilon = \frac{1}{\varepsilon} M_{-1} + M_0 + \varepsilon M_1$),
- (vii) $d(x,y)$, $e(x,y)$ en $g(x,y)$ éénmaal continu differentieerbaar naar y in $\bar{G}$ (noodzakelijk voor de ontwikkeling van $L_\varepsilon = \frac{1}{\varepsilon} M_{-1} + M_0 + \varepsilon M_1$; zie (9.44)).

- (a) $L_2[w]$ is uniform begrensd in $\bar{G}$, indien $g(x,y)$ en $h(x,y)$ in $\bar{G}$ tweemaal continu differentieerbaar zijn naar x en éénmaal continu differentieerbaar naar y ; verder dienen $\phi_+(x)$ en $\gamma_+(x)$ tweemaal continu differentieerbaar te zijn voor $x_1 < x < x_2$.
- (b) $M_1[v_0]$ is uniform begrensd in $\bar{G}$, indien behalve de bovengenoemde eisen (i) - (vii) tevens geldt, dat $\frac{d^2 A}{d\xi^2}$ en $\frac{d^2}{d\xi^2} \{\gamma(\xi, 0)\}$ continu zijn voor $\xi_1 < \xi < \xi_2$.
- (c) $M_0[v_1]$ is uniform begrensd in $\bar{G}$, indien behalve de eisen (i) - (vii) tevens geldt, dat $\frac{dB}{d\xi}$, $\frac{dC}{d\xi}$ en $\frac{d}{d\xi} \{\gamma(\xi, 0)\}$ continu zijn voor $\xi_1 < \xi < \xi_2$.
- (d) $M_1[v_1]$ is uniform begrensd in $\bar{G}$, indien behalve de eisen (i) - (vii) tevens geldt, dat $\frac{d^2 B}{d\xi^2}$, $\frac{d^2 C}{d\xi^2}$ en $\frac{d^2}{d\xi^2} \{\gamma(\xi, 0)\}$ continu zijn voor $\xi_1 < \xi < \xi_2$.

Derhalve zijn $M_1[v_0]$, $M_0[v_1]$ en $M_1[v_1]$ uniform begrensd in $\bar{G}$ indien behalve de voorwaarden (i) - (vii) geldt, dat $\frac{d^3 A}{d\xi^3}$ en $\frac{d^3}{d\xi^3} \{\gamma(\xi, 0)\}$ continu zijn voor

$$\xi_1 < \xi < \xi_2.$$

De continuïteit van $\frac{d^3 A}{d\xi^3}$ brengt mede (zie (9.58) en (9.35)), dat de functies $\phi_+(\xi)$ en $\gamma_+(\xi)$ driemaal continu differentieerbaar zijn in $\xi_1 < \xi < \xi_2$ en dat $g(x,y)$ en $h(x,y)$ in $\bar{G}$ driemaal continu differentieerbaar zijn naar x en tweemaal continu differentieerbaar naar y .

De continuïteit van $\frac{d^3}{d\xi^3} \gamma(\xi, 0)$ (zie (9.42)) brengt mede dat de coëfficiënten $a(x,y)$, $b(x,y)$ en $c(x,y)$ in $\bar{G}$ driemaal continu differentieerbaar zijn naar x en y en dat $\gamma_-(\xi)$ viermaal continu differentieerbaar is voor $\xi_1 < \xi < \xi_2$.

Resumerend zullen we nu voortaan eisen:

- (i*) De coëfficiënten $a(x,y), b(x,y) \dots g(x,y)$ van de differentiaaloperator L_ϵ en het rechterlid $h(x,y)$ behoren tot de klasse $C^3(\bar{G})$;
- (ii*) $a(x,y) \xi^2 + 2b(x,y) \xi \eta + c(x,y) \eta^2 \geq m(\xi^2 + \eta^2)$, $m \geq 0$ voor alle $(x,y) \in G$ en alle getallenparen (ξ, η) ;
- (iii*) $g(x,y) < 0$ in $\bar{G}$;
- (iv*) $\gamma_+(x) \in C^3((x_1, x_2))$;
 $\gamma_-(x) \in C^4((x_1, x_2))$;
- (v*) $\phi_+(x) \in C^3([x_1, x_2])$.

Deze voorwaarden omvatten alle eerder gemaakte veronderstellingen betreffende de parameters van het randwaardeprobleem (9.8) - (9.9). Er geldt nu: $L_2[\bar{w}]$, $M_0[\bar{v}_1]$, $M_1[\bar{v}_0]$ en $M_1[\bar{v}_1]$ zijn uniform begrensd in G en op grond van (9.63) bestaat er een getal K , onafhankelijk van ϵ , zodanig dat

$$(9.66) \quad |L_\epsilon[\bar{R}]| < K\epsilon \quad \text{voor } (x,y) \in \bar{G}.$$

Verder gelden voor $R(x,y;\epsilon)$ de randvoorwaarden

$$(9.64) \quad R(x,y;\epsilon)|_{\tilde{\Gamma}_+} = O(\epsilon^n), \quad n \text{ willekeurig groot},$$

en

$$(9.65) \quad R(x,y;\epsilon)|_{\tilde{\Gamma}_-} = 0.$$

Om voor de restterm $R(x,y;\varepsilon)$ een barrièrefunctie te vinden, dienen we $R(x,y;\varepsilon)$ nog af te schatten langs de randen $x = x_1 + \delta_1$ en $x = x_2 - \delta_2$ van het gebied $\tilde{G}$.

In $\tilde{G}$ zijn de functies ϕ_ε , $w(x,y)$, $\tilde{v}_0(x,y;\varepsilon)$ en $\varepsilon\tilde{v}_1(x,y;\varepsilon)$ uniform begrensd (zie stelling 9.1 en de formules voor w , $\tilde{v}_0$ en $\tilde{v}$) en dus is ook volgens (9.62) de functie $R(x,y;\varepsilon)$ uniform begrensd in $\tilde{G}$. Derhalve bestaat er een getal M , onafhankelijk van ε , zodanig dat

$$(9.67) \quad \begin{cases} |R(x_1 + \delta_1, y; \varepsilon)| < M, \\ |R(x_2 - \delta_2, y; \varepsilon)| < M. \end{cases}$$

Zoals bij het bewijs van stelling 9.1 is verklaard bestaat er een getal $0 < \alpha^2 < 1$ en een getal $\varepsilon_0 > 0$ met de eigenschap, dat

$$1 + \varepsilon e(x,y) > \alpha^2 \quad \text{in } \tilde{G},$$

$$g(x,y) + \varepsilon f(x,y) \leq 0 \quad \text{in } \tilde{G},$$

geldig voor ε met $\varepsilon < \varepsilon_0$.

We construeren nu de volgende barrièrefunctie voor $R(x,y;\varepsilon)$:

$$(9.68) \quad \Psi(x,y) = \frac{2K\varepsilon}{\alpha^2} (y_0 - y) + 2K\varepsilon \quad \text{voor } (x,y) \in \tilde{G} \text{ en } x_1 + 2\delta_1 \leq x \leq x_2 - 2\delta_2,$$

$$(9.69) \quad \Psi(x,y) = \frac{2K\varepsilon}{\alpha^2} (y_0 - y) + 2K\varepsilon + \rho(x) \quad \text{voor } (x,y) \in \tilde{G} \text{ en} \\ x_1 + \delta_1 \leq x \leq x_1 + 2\delta_1 \text{ of} \\ x_2 - 2\delta_2 \leq x \leq x_2 - \delta_2,$$

waarin $y_0 = \max_{\Gamma} y$ en $\rho(x)$ is een niet-negatieve functie behorende tot

$C^\infty([x_1 + \delta_1, x_2 - \delta_2])$ met $\rho(x_1 + \delta_1) = \rho(x_2 - \delta_2) = M$ en $\rho(x) \equiv 0$ voor $x_1 + 2\delta_1 \leq x \leq x_2 - 2\delta_2$.

Er geldt nu $|R(x,y;\varepsilon)| < \Psi(x,y)$ langs de gehele rand van het gebied $\tilde{G}$.

Verder geldt in $\tilde{G}$

$$L_\varepsilon[-\Psi] = -\varepsilon \left[a \frac{d^2 \rho}{dx^2} + d \frac{d\rho}{dx} - \frac{2K\varepsilon}{\alpha^2} e + f \Psi \right] + \frac{2K\varepsilon}{\alpha^2} - g\Psi.$$

Derhalve voor $\varepsilon < \varepsilon_0$

$$L_\varepsilon [-\psi] > 2K\varepsilon - \varepsilon \left[a \frac{d^2 \rho}{dx^2} + a \frac{d\rho}{dx} \right].$$

Omdat $\rho(x) \in C^\infty$ en $\rho(x)$ onafhankelijk is van ε kunnen we K eventueel verhogen zodanig, dat geldt

$$L_\varepsilon [-\psi] > K\varepsilon > |L_\varepsilon [R(x,y;\varepsilon)]|.$$

Derhalve is volgens §9.1 $\psi(x,y)$ in $\tilde{G}$ een barrièrefunctie voor de restterm $R(x,y;\varepsilon)$.

Er geldt dus in $\tilde{G}$

$$R(x,y;\varepsilon) \leq \psi(x,y).$$

Dit geldt a priori in $\tilde{\tilde{G}}$ met $\tilde{\tilde{G}} = \{(x,y); (x,y) \in \tilde{G} \text{ en } x_1 + 2\delta_1 \leq x \leq x_2 - 2\delta_2\}$.

Derhalve

$$R(x,y;\varepsilon) = O(\varepsilon) \text{ in } \tilde{\tilde{G}}.$$

Dus in $\tilde{\tilde{G}}$ geldt volgens (9.62)

$$\phi_\varepsilon(x,y) = w(x,y) + \tilde{v}_0(x,y;\varepsilon) + \varepsilon \tilde{v}_1(x,y;\varepsilon) + O(\varepsilon).$$

De functie $\tilde{v}_1(x,y;\varepsilon)$ is uniform begrensd in $\tilde{\tilde{G}}$ en we mogen dus ook schrijven

$$\phi_\varepsilon(x,y) = w(x,y) + \tilde{v}_0(x,y;\varepsilon) + O(\varepsilon),$$

geldig in het gebied $\tilde{\tilde{G}}$, ofwel

$$(9.70) \quad |\phi_\varepsilon(x,y) - w(x,y) - \tilde{v}_0(x,y;\varepsilon)| < C.\varepsilon \text{ in } \tilde{\tilde{G}}$$

waarin C een constante is, onafhankelijk van (x,y) en ε .

We zijn nu gekomen tot de volgende stelling:

Stelling 9.3. Indien de parameters van het randwaardeprobleem (9.8) - (9.9) voldoen aan de bovengenoemde voorwaarden (i^*) , (ii^*) , ..., (v^*) en indien

de boven- en de benedenrand van het gebied G door elke verticale rechte $x = \text{constant}$ met $x_1 < x < x_2$ in slechts één punt gesneden wordt (geen raakpunten) dan wordt $\phi_\varepsilon(x, y)$ in elk gebied $\tilde{G} = \{(x, y); (x, y) \in \bar{G} \text{ en } x_1 - \delta_1 \leq x \leq x_2 - \delta_2\}$ uniform benaderd door:

$$\phi_\varepsilon(x, y) = w(x, y) + \{\phi_-(x) - w(x, \gamma_-(x))\} \exp\left[-\left\{\frac{y - \gamma_-(x)}{\varepsilon}\right\} \sigma(x, 0)\right] + R(x, y; \varepsilon),$$

waarin $|R(x, y; \varepsilon)| < C\varepsilon$ met C een constante onafhankelijk van (x, y) en ε .

9.5. Plaatselijk niet differentieerbare randvoorwaarden

In deze paragraaf zal een uitbreiding van stelling 9.3 gegeven worden. Het zal blijken dat aan de differentieerbaarheid van de parametrisering van de bovenrand en van de randvoorwaarde aan de bovenrand in geïsoleerde punten niet voldaan behoeft te worden om een eerste benadering van de oplossing te geven. De prijs die we hiervoor moeten betalen is echter dat de orde van benadering achteruit gaat.

Vanwege een veelvuldig gebruik van ordesymbolen met gebroken exponenten gebruiken we om typografische redenen de volgende notatie:

$$O(\varepsilon; \mu) = O(\varepsilon^\mu)$$

Het belangrijkste hulpmiddel voor deze uitbreiding is de "regularisatie", d.w.z. het differentieerbaar maken van de rand en/of randfunctie in de punten waar deze het niet is met behulp van navolgend procedé.

$$|x|^\alpha \exp(-|x|^\beta)$$

is uniform begrensd op $\mathbb{R}$ voor alle positieve getallen α en β .

Bijgevolg geldt

$$(9.71) \quad |x|^\alpha \exp\left(-\frac{|x|^\beta}{\varepsilon^\gamma}\right) = \varepsilon^{\frac{\alpha\gamma}{\beta}} \left(\frac{|x|}{\varepsilon^{\gamma/\beta}}\right)^\alpha \exp\left(-\left|\frac{x}{\varepsilon^{\gamma/\beta}}\right|^\beta\right) = O(\varepsilon; \frac{\alpha\gamma}{\beta})$$

uniform voor $x \in \mathbb{R}$ en α, β en γ positief.

Evenzo is

$$|x|^\alpha (1 - \exp(-|x|^\beta))$$

uniform begrensd op $\mathbb{R}$ voor alle $\beta > 1$ en $0 \leq \alpha > 1 - \beta$ en zelfs continu differentieerbaar als de afgeleide in het punt nul apart wordt gedefinieerd als nul, immers

$$(9.72) \quad |x|^\alpha (1 - \exp - |x|^\beta) = |x|^{\beta+\alpha} - \frac{|x|^{2\beta+\alpha}}{2!} + \frac{|x|^{3\beta+\alpha}}{3!} = o(x; \beta+\alpha)$$

en hieruit volgt de differentieerbaarheid in het punt $x = 0$.

Evenals bij (9.71) geldt nu voor β en $\gamma > 0$ en $0 \leq \alpha > 1 - \beta$ de schatting

$$(9.73) \quad |x|^\alpha (1 - \exp - \frac{|x|^\beta}{\epsilon^\gamma}) = o(\epsilon; \frac{\alpha\gamma}{\beta})$$

We beschouwen evenals boven op het gebied G , met onderrand de x -as en bovenrand $\{(x, \gamma_+(x)) | x \in [\bar{x}_1, \bar{x}_2]\}$, het probleem

$$(9.74) \quad \left\{ \begin{array}{l} L_\epsilon \phi(x, y) = h(x, y), \\ \phi|_{\Gamma_+} = \phi_+(x), \\ \phi|_{\Gamma_-} = \phi_-(x), \end{array} \right.$$

waarin de coëfficiënten van L_ϵ evenals h en $\phi_\pm$ van klasse C^3 zijn en waarin $\phi_+(x_i) = \phi_-(x_i)$ ($i = 1, 2$).

De operator L_ϵ ontwikkelen we, als boven, in een omgeving van de onderrand na substitutie $y = \epsilon\eta$ naar ϵ :

$$L(x, \epsilon\eta) = \frac{1}{\epsilon} M_0(x, \eta) + M_1(x, \eta) + \epsilon M_2(x, \eta, \epsilon).$$

Ontwikkeling van coëfficiënten van L_ϵ geeft

$$\left\{ \begin{array}{l} b(x, \epsilon\eta) = b_0(x) + \epsilon\eta b_1(x, \epsilon\eta), \\ c(x, \epsilon\eta) = c_0(x) + \epsilon\eta c_1(x) + \epsilon^2 \eta^2 c_2(x, \epsilon\eta), \\ e(x, \epsilon\eta) = e_0(x) + \epsilon\eta e_1(x, \epsilon\eta), \\ g(x, \epsilon\eta) = g_0(x) + \epsilon\eta g_1(x, \epsilon\eta), \end{array} \right.$$

en de operatoren zijn dan

$$(9.75) \quad \left\{ \begin{array}{l} M_0 = c_0(x) \frac{\partial^2}{\partial \eta^2} + \frac{\partial}{\partial \eta} \\ M_1 = \eta c_1(x) \frac{\partial^2}{\partial \eta^2} + 2 b_0(x) \frac{\partial^2}{\partial x \partial \eta} + e_0(x) \frac{\partial}{\partial \eta} + g_0(x) \\ M_2 = a(x, \eta) \frac{\partial^2}{\partial x^2} + 2\eta b_1(x, \eta) \frac{\partial^2}{\partial x \partial \eta} + \eta^2 c_2(x, \eta) \frac{\partial^2}{\partial \eta^2} + \\ + d(x, \eta) \frac{\partial}{\partial x} + \eta e_1(x, \eta) \frac{\partial}{\partial \eta} + \eta g_1(x, \eta). \end{array} \right.$$

Omdat L_ϵ een elliptische operator is en $a(x, \eta) > 0$, is ook $c(x, \eta) > 0$ en daarmee ook $c_0(x)$; verder zijn c_0 , b_0 , e_0 en g_0 van klasse C^3 en c_1 van klasse C^2 . De coëfficiënten in M_2 zijn alle minstens C^1 .

Stelling 9.4. Als γ_+ een C^3 functie is op (x_1, x_2) en ϕ_+ is C^3 uitgezonderd in het punt $p \in (x_1, x_2)$ en in p geldt

$$(9.76) \quad \left\{ \begin{array}{l} \phi_+(x) = O(|x-p|; \alpha) \text{ met } 0 < \alpha \leq 3 \\ \phi'_+(x) = O(|x-p|; \beta) \\ \phi''_+(x) = O(|x-p|; \gamma) \\ \phi'''_+(x) = O(|x-p|; \delta) \text{ met } \delta \leq 0 \end{array} \right.$$

en als $\rho = \max(3-\alpha, 2-\beta, 1-\gamma, -\delta)$.

Dan kunnen we een functie ϕ_0 konstrueren, zo dat

$$\phi - \phi_0 = O(\epsilon; \frac{\alpha}{\alpha+\rho})$$

uniform op $\tilde{G}$ (= G afgezien van willekeurig kleine omgevingen van x_1 en x_2).

Voorbeeld 9.1. Een voorbeeld van een functie die aan de eisen van stelling 9.4 voldoet is $x^2 \sin \frac{1}{x}$, waarvoor dan $\beta = 0$, $\gamma = -2$ en $\delta = -4$.

Opmerking 9.3. Het is geen verlies van algemeenheid, dat we $\phi_+(p) = 0$ gesteld hebben en zijn afgeleiden ook indien ze bestaan want omdat het probleem lineair is kunnen we altijd een C^∞ functie van ϕ_+ aftrekken en deze onafhankelijk met stelling 9.3 behandelen.

Bewijs. We regulariseren de randwaarde functie, rekenen hiermede analoog aan stelling 9.3 de benaderende ϕ_0 uit en bewijzen vervolgens met behulp van stelling 9.3 dat deze een goede benadering is van de echte oplossing.

Definieer nu

$$(9.77) \quad \phi_0(x) = \phi_+(x) \left(1 - \exp - \frac{(x-p)^{2k}}{\varepsilon^\lambda}\right),$$

met $\lambda > 0$ en $k \in \mathbb{N}$ en $2k > p+1$,

$$\text{en} \quad \phi_0^{(i)}(p) = 0 \text{ met } i = 1, 2, 3.$$

$$\begin{aligned} \text{Dan volgt: } \phi_0'(x) &= \phi_+'(x) \left(1 - \exp - \varepsilon^{-\lambda}(x-p)^{2k}\right) + \\ &+ 2k\varepsilon^{-\lambda}(x-p)^{2k-1} \phi_+(x) \exp - \varepsilon^{-\lambda}(x-p)^{2k}, \end{aligned}$$

en analoog aan (9.72) is ϕ_0 differentieerbaar in p en met (9.73) en (9.71) volgt

$$(9.78) \quad \phi_0'(x) = O(\varepsilon; \frac{\lambda\beta}{2k}) + O(\varepsilon; \frac{\lambda(\alpha-1)}{2k}) \quad \text{als } \beta \leq 0,$$

$$\text{en} \quad \phi_0'(x) = O(\varepsilon; \frac{\lambda(\alpha-1)}{2k}) + O(1) \quad \text{als } \beta > 0.$$

Op dezelfde wijze volgt dat

$$\begin{aligned} \phi_0''(x) &= \phi_+''(1 - \exp - \varepsilon^{-\lambda}(x-p)^{2k}) + 4k\varepsilon^{-\lambda}(x-p)^{2k-1} \phi_+'(x) \exp - \varepsilon^{-\lambda}(x-p)^{2k} \\ &+ [2k(2k-1)\varepsilon^{-\lambda}(x-p)^{2k-2} \phi_+(x) - 4k^2\varepsilon^{-2\lambda}(x-p)^{4k-2} \phi_+(x)] \exp - \varepsilon^{-\lambda}(x-p)^{2k} \end{aligned}$$

continuïes in p en afgeschat kan worden door

$$(9.79) \quad \phi'' = O(\varepsilon; \frac{\lambda}{2k} \min(\gamma, \beta-1, \alpha-2)) \quad \text{als } \gamma \leq 0,$$

en

$$\phi''(x) = O(\varepsilon; \frac{\lambda}{2k} \min(\beta-1, \alpha-2, 0)) \quad \text{als } \gamma > 0,$$

en dat ook ϕ''' continu is in nul en afgeschat kan worden door

$$(9.80) \quad \phi''' = O(\varepsilon; \frac{\lambda}{2k} \min(\delta, \gamma-1, \beta-2, \alpha-3)) = O(\varepsilon; -\frac{\lambda\rho}{2k}).$$

Verder vinden we dat

$$(9.81) \quad \phi_0(x) - \phi_+(x) = -\phi_+(x) \exp - \varepsilon^{-\lambda}(x-p)^{2k} = O(\varepsilon; \frac{\lambda\alpha}{2k}) \text{ met (1).}$$

Los nu op:

$$\left\{ \begin{array}{l} \left[\frac{\partial}{\partial y} + g(x,y) \right] w(x,y) = h(x,y) \\ \qquad \qquad \qquad \text{met } w(x, \gamma_+(x)) = \phi_0(x); \\ \\ M_0 v_0(x,\eta) = 0 \\ \qquad \qquad \qquad \text{met } v_0(x,0) = \phi_+(x) - w(x,0) \\ \qquad \qquad \qquad \text{en } \lim_{\eta \rightarrow \infty} v_0(x,\eta) = 0; \\ \\ M_0 v_1(x,\eta) = -M_1 v_0(x,\eta) \\ \qquad \qquad \qquad \text{met } v_1(x,0) = 0 \\ \qquad \qquad \qquad \text{en } \lim_{\eta \rightarrow \infty} v_1(x,\eta) = 0. \end{array} \right.$$

Hieruit volgt achtereenvolgens

$$(9.82) \quad w(x, y) = \phi_-(x) - \int_y^{\gamma_+(x)} \{h(x, n) - g(x, n) \phi_0(x)\} \left\{ \exp \int_y^n g(x, \zeta) d\zeta \right\} dn,$$

$$(9.83) \quad v_0(x, n) = \{\phi_-(x) - w(x, 0)\} \exp - \frac{n}{c_0(x)}$$

en als $w_0(x) \stackrel{\text{def}}{=} w(x, 0)$, dan

$$(9.84) \quad v_1(x, n) = \left\{ \frac{1}{2} (c_0 c_1 - 2 b_0 c_0') (\phi_0 - w_0) c_0^{-3} n \right. + \\ \left. + (g_0 c_0^2 - c_0 c_1 - c_0 e_0) (\phi_0 - w_0) c_0^{-2} - 2 b_0 c_0^{-1} (\phi_0' - w_0') \right\} + n \exp - \frac{n}{c_0}.$$

Nu geldt voor aldus geconstrueerde ϕ_0

$$(9.85) \quad \phi_0(x, y) \stackrel{\text{def}}{=} w(x, y) + v_0(x, \frac{y}{\varepsilon}) + v_1(x, \frac{y}{\varepsilon})$$

dat

$$L_\varepsilon [\phi - \phi_0] = O(\varepsilon; 1 - \frac{\lambda \rho}{2k}).$$

We zien dit als volgt

$$L_\varepsilon [\phi - \phi_0] = L_1 w(x, y) + \\ + \varepsilon M_2 v_0(x, n) + \varepsilon (M_1 + \varepsilon M_2) v_1(x, n).$$

Uit de formules (9.78) t/m (9.80) zien we dat we kunnen schrijven

$$(9.86) \quad \begin{cases} w(x,y) = r_1(x,y) \phi_0(x) + r_2(x,y) \quad , \\ v_0(x,\eta) = s_1(x,\eta) \phi_0(x) + s_2(x,\eta) \quad , \\ v_1(x,\eta) = t_1(x,\eta) \phi_0(x) + t_2(x,\eta) \phi_0'(x) + t_3(x,\eta) \quad , \end{cases}$$

waarin r_i , s_i en t_i minstens C^2 zijn en onafhankelijk van ε en dus uniform begrensde eerste en tweede partiele afgeleiden hebben op $\tilde{G}$.

Bijgevolg is

$$(9.87) \quad |L_\varepsilon[\Phi - \phi_0]| \leq \varepsilon K(|\phi_0| + |\phi_0'| + |\phi_0''| + |\phi_0'''|) .$$

Volgens de schattingen (9.76) geldt dus

$$L_\varepsilon[\Phi - \phi_0] = O(\varepsilon; 1 - \frac{\lambda\rho}{2k}) .$$

Op de onderrand geldt

$$\phi_0(x,0) = \phi_-(x) .$$

Op de bovenrand geldt

$$\begin{aligned} (9.88) \quad \phi_0(x, \gamma_+(x)) &= \phi_0(x) + v_0(x, \frac{\gamma_+(x)}{\varepsilon}) + \varepsilon v_1(x, \frac{\gamma_+(x)}{\varepsilon}) \\ &= \phi_0(x) + O(\varepsilon^N) \text{ buiten willekeurig kleine} \\ &\quad \text{vaste omgevingen van } x_1 \text{ en } x_2 \\ &= \phi_+(x) + (\phi_0(x) - \phi_+(x)) + O(\varepsilon^N) \\ &= \phi_+(x) + O(\varepsilon; \frac{\lambda\alpha}{2k}) \text{ volgens (9.81)} \end{aligned}$$

op de bovenrand van $\tilde{G}$.

Toepassing van stelling 9.3 geeft dan

$$\phi - \phi_0 = O(\varepsilon; \min(\frac{\lambda\alpha}{2k}, 1 - \frac{\lambda\rho}{2k})) \text{ uniform op } \tilde{G}.$$

We zijn nu nog geheel vrij in de keuze van $\frac{\lambda}{2k}$ en kiezen deze dus optimaal, zodat

$$\phi - \phi_0 = O(\varepsilon; \frac{\alpha}{\alpha+\rho}) \text{ uniform op } \tilde{G}.$$

Opmerkingen 9.4.

(i) Omdat v_1 uniform begrensd is op $\tilde{G}$ is dus $w(x,y) + v_0(x, \frac{y}{\varepsilon})$ ook een

benadering van ϕ tot op $O(\varepsilon; \frac{\alpha}{\alpha+\rho})$.

(ii) Omdat $\phi_- - \phi_0 = O(\varepsilon; \frac{\alpha}{\alpha+\rho})$ is dus ook $\tilde{w}(x,y) + \tilde{v}(x, \frac{y}{\varepsilon})$

een benadering van ϕ tot op $O(\varepsilon; \frac{\alpha}{\alpha+\rho})$, waarbij

$$\tilde{w}(x,y) = \phi_+(x) - \int_y^{\gamma_+(x)} \{h(x,\eta) - g(x,\eta) \phi_+(x)\} \left\{ \exp \int_y^\eta g(x_1 \zeta) d\zeta \right\} d\eta$$

en

$$\tilde{v}(x,\eta) = \{\phi_+(x) - \tilde{w}(x,0)\} \exp - \frac{\eta}{c_0(x)}.$$

(iii) Deze stelling stelt ons in staat de oplossing te benaderen tot op $O(\varepsilon; 1/3)$ als de bovenrandfunctie is:

$$\phi_+(x) = (x-p)^2 \sin \frac{1}{x-p}.$$

Als aan de parametrisering van de bovenrand, $\gamma_+(x)$, de voorwaarden (9.75) wordt opgelegd, dan kunnen we een resultaat, analoog aan de bovenstaande stelling, afleiden als we γ_+ regulariseren.

Daar dit echter betekent dat we rekenen met een iets andere bovenrand, moeten we eisen dat de coëfficiënten van L_ε en ook $h(x,y)$ gedefinieerd zijn tot op een omgeving van de bovenrand buiten het gebied G en daar C^3 functies zijn. Een ander verschil is dat w , v_0 en v_1 niet lineair van γ_+ en γ'_+ afhankelijk zijn zoals in (9.86) van ϕ_0 en ϕ'_0 ; $w(x,y)$ is immers een C^4 - functie van γ_- , zodat ook v_0 een C^4 - functie is van γ_- en v_1 een C^3 - functie van γ_- (via w_0). Een afschatting als in (9.87) zal dan ook niet mogelijk zijn; we zullen dit dan ook op de volgende manier doen.

Laat $p : \mathbb{R} \rightarrow \mathbb{R}$ een driemaal continu differentieerbare functie zijn en laat γ_0 de geregulariseerde functie zijn bij γ_+ zoals in (9.77), dan geldt:

$$(9.89) \left\{ \begin{aligned} p(\gamma_0(x)) &= O(1) \\ \frac{d}{dx}[p(\gamma_0(x))] &= p'(\gamma_0(x)) \gamma'_0(x) = O(\varepsilon; \frac{\lambda}{2k} \cdot \min(\beta, \alpha-1, 0)) \\ \frac{d^2}{dx^2}[p(\gamma_0(x))] &= p''(\gamma_0(x))(\gamma'_0(x))^2 + p'(\gamma_0(x)) \gamma''_0(x) \\ &= O(\varepsilon; \frac{\lambda}{2k} \cdot \min(2\beta, 2\alpha-2, \gamma, \beta-1, \alpha-2, 0)) \\ \frac{d^3}{dx^3}[p(\gamma_0(x))] &= p'''(\gamma_0(x))(\gamma'_0(x))^3 + 3 p''(\gamma_0(x)) \gamma'_0(x) \gamma''_0(x) \\ &\quad + p'(\gamma_0(x)) \gamma'''_0(x) \\ &= O(\varepsilon; \frac{\lambda}{2k} \cdot \min(\beta+\gamma, 2\beta-1, \beta+\alpha-2, \gamma+\alpha-1, 2\alpha-3, \gamma, \beta-1, \alpha-2, \\ &\quad \alpha-3, \beta-2, \gamma-1, \delta, 3\beta, 3\alpha-3, 0)) \\ &= O(\varepsilon; \frac{\lambda}{2k} \cdot \min(-\rho, \beta+\gamma, 2\beta-1, 3\beta)). \end{aligned} \right.$$

Definieer nu $\rho_0 = \max(\rho, -\beta-\gamma, 1-2\beta, -3\beta)$ dan volgt analoog aan (9.87) als ϕ_0 een benadering is als (9.85) voor het gebied met geregulariseerde rand:

$$(9.90) \quad \begin{aligned} |L_\varepsilon[\Phi - \phi_0]| &\leq \varepsilon K(|\gamma_0| + |\gamma'_0| + |\gamma''_0| + |\gamma'''_0| + |\gamma_0|^3 + |\gamma'_0 \gamma''_0|) \\ &= O(\varepsilon; 1 - \frac{\lambda \rho_0}{2k}) \quad (\text{volgens (9.89)}). \end{aligned}$$

Het resultaat is dus de volgende stelling:

Stelling 9.5. Als ϕ de oplossing is van het probleem (9.74), als de coëfficiënten van $L_\varepsilon C^3$ zijn op een omgeving van G en als γ_+ aan de voorwaarden (9.76) voldoet, dan kunnen we een functie ϕ_0 construeren, zodat

$$\phi - \phi_0 = O(\varepsilon; \frac{\alpha}{\alpha + \rho_0})$$

uniform op $\tilde{G}$, waarbij $\rho_0 = \max(\rho, -\beta - \gamma, 1 - 2\beta, -3\beta)$.

Bewijs. We behoeven alleen nog op de rand te verifiëren dat

$$(\phi - \phi_0)|_\Gamma = O(\varepsilon; \frac{\alpha}{\alpha + \rho}) .$$

$$\begin{aligned} \text{Nu is } \phi_0|_{\Gamma_-} &= \phi_- \quad \text{en } \phi_0(x, \gamma_+(x)) = \phi_0(x, \gamma_0(x)) + O(\gamma_- - \gamma_0) \\ &= \phi_- + O(\varepsilon; \frac{\lambda\alpha}{2k}) . \end{aligned}$$

Met behulp van (9.90) en dit laatste volgt dan, evenals de vorige keer, het gestelde.

Voorbeeld 9.2. Als γ_+ lokaal gelijk is aan $(x-p)^{1/3}$, dan is de benadering ϕ_0 goed tot op orde $O(\varepsilon; 1/9)$.

Opmerking 9.5. We zien hier dat, in tegenstelling tot het aanvankelijke vermoeden van het ontstaan van een vrije grenslaag van gelijke grootte orde als de oplossing zelf langs de karakteristiek $x = p$ van L_2 , er tot op zekere orde van ε niets te bespeuren is van een dergelijke vrije grenslaag.

Opmerking 9.6. Regularisatie d.m.v.

$$\phi(x) \mapsto \tilde{\phi}(x) = \frac{2}{3\sqrt{\pi\varepsilon}} \int_{x_1}^{x_2} \phi(t) \left(1 - \frac{2}{3}t^2\right) \exp - \frac{(x-t)^2}{\varepsilon} dt$$

voor $\phi \in C^{k+\alpha} [x_1, x_2]$ met $k = 0, 1, 2$ en $0 < \alpha \leq 1$ geeft op analoge wijze veel betere schattingen. Als ϕ_+ , ϕ_- en $\gamma_+ \in C^{k+\alpha}$ dan is te bewijzen

$$\phi - \phi_0 = O(\varepsilon; \min(\frac{k+\alpha}{2}, 1)).$$

Zie hiervoor bijvoorbeeld [5].

9.6. Asymptotische oplossing van een singulier storingsprobleem voor een kwartvlak

In dit hoofdstuk zijn tot nu toe partiële differentiaalvergelijkingen beschouwd van het type

$$(9.91) \quad L_{\epsilon} [\bar{\phi}_{\epsilon}] = h(x,y),$$

waarbij het definitiegebied G eindig verondersteld is en de karakteristieken van de gereduceerde vergelijking

$$L_1 [\bar{\phi}_{\epsilon}] = h(x,y)$$

niet samenvallen met gedeelten van de rand van G . In het nu volgende onderzoeken wij een probleem waarbij dit juist wel het geval is. Wij zullen een konkreet geval behandelen, waarin (9.91) en G zodanig gekozen zijn dat de oplossing als integraal gegeven kan worden, waarna wij bepaalde methoden uit de asymptotiek te hulp roepen om het gedrag van de oplossing $\bar{\phi}_{\epsilon}(x,y)$ te bepalen voor $\epsilon \rightarrow 0$. Er zal vooral aandacht besteed worden aan het gedrag in de grenslaag, waarbij onze aanpak in vele aspecten verschilt met die van Grasman [4], die dit probleem eerder behandelde, en met die van Eckhaus en de Jager [3].

Wij beschouwen het volgende randwaardeprobleem

$$(9.92) \quad \begin{cases} \epsilon \Delta \bar{\phi}_{\epsilon}(x,y) - \frac{\partial}{\partial y} \bar{\phi}_{\epsilon}(x,y) = 0, & x > 0, y > 0, \\ \bar{\phi}_{\epsilon}(x,0) = 0, & \bar{\phi}_{\epsilon}(0,y) = \phi(y), \end{cases}$$

waarin $\phi(y)$ aan zekere voorwaarden voldoet, die later duidelijk zullen worden. Zoals in [3] is aangetoond zal bij $x \sim 0$, $y > 0$ de grenslaag ontstaan als $\epsilon \rightarrow 0$, een parabolische grenslaag. Aan de hand van de keuze $\phi(y) = y^{-\frac{1}{2}}$ kan dit direkt geverifieerd worden, aangezien in dit geval de oplossing luidt

$$(9.93) \quad \phi_{\epsilon}(x,y) = \sqrt{\frac{2}{r}} \sin \frac{1}{2} \theta \exp\left(\frac{r(\sin\theta-1)}{2\epsilon}\right),$$

waarin r en θ de poolcoördinaten zijn ($x = r\cos\theta$, $y = r\sin\theta$). Het argument van de exponentiële functie is negatief voor $0 \leq \theta < \frac{\pi}{2}$, zodat voor deze θ -waarden

$$\phi_{\epsilon}(x,y) = O(\epsilon^N),$$

als $\epsilon \rightarrow 0$, voor elke positieve N . Deze schatting is echter niet uniform in θ , want voor die waarden van θ en r waarvoor

$$\frac{r(1-\sin\theta)}{2\epsilon} = O(1),$$

als $\epsilon \rightarrow 0$, is de invloed van de exponentiële functie in (9.93) verdwenen. De verzameling in het x - y -vlak waarvoor geldt

$$\frac{r(1-\sin\theta)}{2\epsilon} = c$$

is de parabool

$$y = (x^2 - 4\epsilon^2 c^2)/4\epsilon c,$$

buiten deze parabool begint de e -macht pas aan invloed te winnen. De oplossing van (9.92) met algemenere randfunctie $\phi(y)$ zal dit verschijnsel ook vertonen, er ontstaat eveneens een parabolische grenslaag.

Wij zullen nu de oplossing van (9.92) geven. Aangezien de beschouwde vergelijking konstante coëfficiënten heeft kunnen we door middel van de substitutie

$$(9.94) \quad V(x,y) = e^{-\omega y} \phi_{\epsilon}(x,y)$$

met $\omega = \frac{1}{2\varepsilon}$, de eerste orde afgeleide laten verdwijnen (zie hoofdstuk 1 van [1]).

De funktie $V(x,y)$ moet dan voldoen aan

$$\Delta V(x,y) - \omega^2 V(x,y) = 0, \quad x > 0, y > 0, \quad (9.95)$$

$$V(x,0) = 0, \quad V(0,y) = e^{-\omega y} \phi(y), \\ \omega = \frac{1}{2\varepsilon}.$$

Laten wij proberen de variabelen te separeren, zodat wij veronderstellen dat

$$V(x,y) = f(x) g(y).$$

Voor f en g vinden wij dan

$$\frac{d^2 f(x)}{dx^2} - (\lambda^2 + \omega^2) f(x) = 0, \quad \text{dus } f(x) = \exp(\pm \sqrt{\lambda^2 + \omega^2} x), \\ \frac{d^2 g(y)}{dy^2} + \lambda^2 g(y) = 0, \quad \text{dus } g(y) = \exp(\pm i\lambda y),$$

waarin λ de separatie variabele is. Voor $g(y)$ kiezen wij $\sin \lambda y$ aangezien voor $y = 0$ dan aan de randwaarde voor $V(x,y)$ is voldaan en voor $f(x)$ de exponentiële functie met het negatieve argument om het gedrag voor $x \rightarrow \infty$ in de hand te houden. Superpositie van deze f en g t.o.v. de parameter λ levert

$$V(x,y) = \int_0^\infty \sin \lambda y \exp(-x \sqrt{\lambda^2 + \omega^2}) c(\lambda) d\lambda,$$

waarin de willekeurige functie $c(\lambda)$ nader bepaald wordt uit de randwaarde voor $x = 0$, namelijk

$$e^{-\omega y} \phi(y) = \int_0^{\infty} \sin \lambda y \quad c(\lambda) \, d\lambda,$$

zodat

$$\begin{aligned} c(\lambda) &= \frac{2}{\pi} \int_0^{\infty} e^{-\omega y} \phi(y) \sin \lambda y \, dy \\ &= \frac{1}{\pi i} \{ \bar{\phi}(\omega - i\lambda) - \bar{\phi}(\omega + i\lambda) \}. \end{aligned}$$

Hierin is

$$\bar{\phi}(s) = \int_0^{\infty} e^{-st} \phi(t) dt.$$

Aangezien $c(\lambda)$ oneven is in λ kunnen wij V ook schrijven als

$$V(x, y) = \frac{1}{2i} \int_{-\infty}^{\infty} \exp(i\lambda y - x \sqrt{\lambda^2 + \omega^2}) \quad c(\lambda) \, d\lambda$$

en als wij nu $\lambda = \omega \operatorname{sht}$ stellen en overgaan op poolcoördinaten ($x = r \cos \theta$, $y = r \sin \theta$) dan levert dit

$$V(x, y) = \frac{\omega}{2i} \int_{-\infty}^{\infty} e^{-\omega r \operatorname{ch}(t-i\theta)} \quad c(\omega \operatorname{sht}) \operatorname{cht} \, dt.$$

Door nu nog t te vervangen door $t + i\theta$ krijgen wij, indien de verschuiving van de nieuwe integratieweg naar de reële as is toegestaan, tenslotte

$$\begin{aligned} (9.96) \quad V(x, y) &= \frac{\omega}{2\pi} \int_{-\infty}^{\infty} e^{-\omega r \operatorname{cht}} \operatorname{ch}(t + i\theta) \{ \bar{\phi}(\omega(1 + i \operatorname{sh}(t+i\theta))) \\ &\quad - \bar{\phi}(\omega(1 - i \operatorname{sh}(t+i\theta))) \} dt. \end{aligned}$$

De opgave die wij ons stellen is het asymptotisch gedrag van de functie $V(x,y)$ te bepalen voor $\omega \rightarrow \infty$, speciaal in de buurt van de lijn $\theta = \frac{\pi}{2}$ (de grenslaag). Wij zullen dit probleem in drie fasen proberen op te lossen, allereerst voor de randfunctie $\phi(y) = 1$, om enig inzicht te verwerven voor het moeilijker geval, daarna voor $\phi(y) = y^{\nu-1}$ met $\text{Re } \nu > 0$ om tenslotte met de gevonden resultaten het probleem voor een algemenere randfunctie $\phi(y)$ te behandelen.

Speciaal geval $\phi(y) \equiv 1$

De oplossing van (9.92) kan met (9.96) als uitgangspunt na enige algebraïsche manipulaties voor het geval $\phi(y) = 1$ worden geschreven als

$$(9.97) \quad V(x,y) = \frac{\sin 2\theta}{\pi} \int_{-\infty}^{\infty} \frac{e^{-\omega r \cosh 2t} dt}{\cosh 2t + \cos 2\theta}.$$

De standaardmethode om een integraal van dit type voor grote ω te ontwikkelen is de zadelpuntmethode. Men bepaalt de zadelpunten van de functie in de exponent, gaat dan na of de integratieweg vervangen kan worden door een weg door het zadelpunt die optimale informatie kan verschaffen over de waarde van de integraal en men voert enkele substituties uit om die informatie te verkrijgen. Bij (9.97) zijn de zadelpunten $t = 0, \pm i k \pi$, $k = 1, 2, 3, \dots$ waarvan alleen $t = 0$ bruikbaar blijkt te zijn en de meest geschikte integratieweg voor dit zadelpunt is de reële t -as. Door de substitutie $u = \sinh \frac{1}{2} t$ wordt dan (9.97) na enig rekenwerk

$$(9.98) \quad V(x,y) = \frac{\sin 2\theta}{4\pi} e^{-\omega r} \int_0^{\infty} t^{-\frac{1}{2}} e^{-2\omega r t} f(t) dt,$$

met
$$f(t) = \frac{1}{(t-t_0)(t-t_1)\sqrt{1+t}},$$

waarin
$$t_0 = -\frac{1 - \sin \theta}{2}, \quad t_1 = -\frac{1 + \sin \theta}{2}.$$

De asymptotische ontwikkeling van deze integraal kan men nu verkrijgen door $f(t)$ in machten van t te ontwikkelen en de reeks dan in (9.98) te substitueren. Dit is echter alleen mogelijk als $\theta \neq \frac{\pi}{2}$ aangezien $f(t)$ een pool van de eerste orde heeft voor $t = t_0$. De coëfficiënten a_n van

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

zijn voor $\theta = \frac{\pi}{2}$ niet gedefinieerd en voor $0 \leq \theta < \frac{\pi}{2}$ niet begrensd, zodat de substitutie van de reeks in (9.98) ons niet voor alle waarden van θ die ons probleem bepalen een bruikbare asymptotische ontwikkeling verschaft. Dit is niet zo vreemd want wij weten dat $V(x,y)$ zich juist bij de grenslaag ($\theta \sim \frac{\pi}{2}$) grillig gedraagt. De pool in $t = t_0$ veroorzaakt de moeilijkheden, de andere (voor $t = t_1$) blijft daarvoor ver genoeg van de oorsprong weg. Als wij b_{-1} het residu van $f(t)$ voor de pool in $t = t_0$ noemen dan kunnen we $f(t)$ opsplitsen in

$$(9.99) \quad f(t) = \frac{b_{-1}}{t - t_0} + g(t),$$

waarin

$$b_{-1} = \frac{\sqrt{2}}{\sin \theta \sqrt{1 + \sin \theta}}$$

en

$$g(t) = \sum_{k=0}^{\infty} b_n t^n ;$$

de reeks convergeert voor $|t| < |t_1|$ en de coëfficiënten b_n zijn begrensd voor $0 \leq \theta \leq \frac{\pi}{2}$. Substitutie van (9.99) in (9.98) levert dan

$$\begin{aligned} V(x,y) = & b_{-1} \frac{\sin 2\theta}{4\pi} e^{-\omega r} \int_0^{\infty} t^{-\frac{1}{2}} e^{-2\omega r t} \frac{dt}{t - t_0} \\ & + \frac{\sin 2\theta}{4\pi} e^{-\omega r} \int_0^{\infty} t^{-\frac{1}{2}} e^{-2\omega r t} g(t) dt . \end{aligned}$$

Substitutie van (9.94) geeft voor $\phi_\epsilon(x,y)$ tenslotte de asymptotische ontwikkeling

$$(9.100) \quad \phi_\epsilon(x,y) \sim \operatorname{erfc} \sqrt{\omega r(1-\sin\theta)} + \frac{\sin 2\theta}{4\pi} e^{-\omega r(1-\sin\theta)} \sum_{n=0}^{\infty} b_n \frac{\Gamma(n+\frac{1}{2})}{(2\omega r)^{n+\frac{1}{2}}}$$

voor $\omega r \rightarrow \infty$, uniform in $0 \leq \theta \leq \pi/2$.

Het asymptotisch gedrag van $\phi_\epsilon(x,y)$ wordt voor $\theta \rightarrow \pi/2$ dus volledig bepaald door de errorfunctie, aangezien de tweede term rechts in (9.100) vanwege de factor $\sin 2\theta$ verdwijnt als $\theta \rightarrow \pi/2$. De errorfunctie is de grenslaagfunctie; in de omgeving van de rand $x = 0$ ($\theta = \frac{\pi}{2}$) speelt deze term een belangrijke rol, daarbuiten is de term asymptotisch te verwaarlozen t.o.v. de andere termen van de reeks.

De methode de pool af te splitsen als het zadelpunt en de pool dicht bij elkaar liggen ontleen we aan van der Waerden [10].

Opmerking 9.7. De asymptotische ontwikkeling (9.100) geldt voor $\omega r \rightarrow \infty$, dus niet voor het gebied in het x - y vlak waarvoor $r = O(\epsilon)$; voor deze omgeving kan echter een andere methode gevolgd worden.

Het geval $\phi(y) = y^{v-1}$

Als $2v - 1$ een natuurlijk getal is kan de vorige methode toegepast worden. Er ontstaat dan een analoge representatie als (9.98) voor $V(x,y)$, maar nu heeft $f(t)$ een pool van de orde $2v - 1$ die als $\theta \rightarrow \frac{\pi}{2}$ de oorsprong nadert. Afsplitsing is mogelijk in de vorm

$$(9.101) \quad f(t) = \sum_{n=1}^{2v-1} \frac{b_{-n}}{(t-t_0)^n} + g(t)$$

en de eindige reeks geeft na substitutie in de integraal $2v-1$ termen, die uitgedrukt kunnen worden in Whittaker functies (zie [7]).

Voor $\theta \rightarrow \frac{\pi}{2}$ zijn echter alle termen singulier, terwijl de som daarvan eindig is en de randfunctie oplevert. Bovendien zijn de coëfficiënten b_{-n} in het algemeen niet in konkrete vorm te geven. Door deze twee aspecten is het probleem niet erg overzichtelijk meer en toepassing van de resultaten op een algemenere randfunctie $\phi(y)$ is niet meer te realiseren. We zullen daarom een andere, hoewel analoge, methode volgen, waarbij slechts verondersteld wordt dat $\operatorname{Re} v > 0$.

Uitgangspunt is ook nu weer representatie (9.96).

De integrand bestaat uit twee termen en we noemen de integraal met de eerste term $V_1(r, \theta)$, zodat

$$(9.102) \quad V_1(r, \theta) = \frac{\omega}{2\pi} \int_{-\infty}^{\infty} e^{-\omega r \cosh t} \operatorname{ch}(t+i\theta) \overline{\phi(\omega(1+\operatorname{ish}(t+i\theta)))} dt.$$

Dan geldt

$$(9.103) \quad V(x, y) = V_1(r, \theta) - V_1(r, -\theta)$$

en alleen de eerste term in het rechterlid van (9.104) levert voor $\theta \rightarrow \pi/2$ de randfunctie op, de andere term verdwijnt in dat geval. We beschouwen daarom eerst (9.102) waarin

$$\phi(s) = \frac{\Gamma(v)}{s^v}$$

zodat

$$(9.104) \quad V_1(r, \theta) = \frac{\Gamma(v)}{2\pi \omega^{v-1}} \int_{-\infty}^{\infty} e^{-\omega r \cosh t} \operatorname{ch}(t+i\theta) \frac{dt}{\{1+\operatorname{ish}(t+i\theta)\}^v}.$$

Stel nu $u = i \operatorname{sh} \frac{1}{2} t$, dan gaat (9.104) over in

$$(9.105) \quad V_1(r, \theta) = \frac{\Gamma(v)}{\pi i \omega^{v-1}} e^{-\omega r} \int_{-i\infty}^{i\infty} e^{2\omega r u^2} f(u) du$$

met

$$f(u) = \left(\frac{1-2u^2}{\sqrt{1-u^2}} \cos\theta + 2u\sin\theta \right) \{1 + 2u\sqrt{1-u^2} \cos\theta - (1-2u^2)\sin\theta\}^{-v}.$$

Aangezien niet verondersteld wordt dat v een geheel getal is kan de singulariteit van $f(u)$ dicht bij de oorsprong (voor $u = u_0 = \sqrt{(1-\sin\theta)}/2$) niet afgesplitst worden. Deze singulariteit is van zodanige aard dat $f(u) (u-u_0)^{2v-1}$ analytisch is in een omgeving van u_0 , zodat voor u in die omgeving

$$(9.106) \quad f(u) (u-u_0)^{2v-1} = \sum_{k=0}^{\infty} a_k (u-u_0)^k.$$

De coëfficiënten zijn voor $0 \leq \theta \leq \frac{\pi}{2}$ begrensde functies van θ , de eerste zijn

$$(9.107) \quad \begin{aligned} a_0 &= \left(\frac{1+\sin\theta}{4} \right)^{v-1}, \\ a_1 &= \frac{1}{4} (2v-3) \sqrt{\frac{1-\sin\theta}{2}} \left(\frac{1+\sin\theta}{4} \right)^{v-2}. \end{aligned}$$

Wij krijgen nu een asymptotische ontwikkeling voor $V_1(r, \theta)$ als we de ontwikkeling voor $f(u)$ uit (9.106) in (9.105) invullen en integratie en sommatie verwisselen, met als gevolg

$$V_1(r, \theta) \sim \frac{\Gamma(v)}{\pi i \omega^{v-1}} e^{-\omega r} \sum_{k=0}^{\infty} a_k \int_{-i\infty}^{i\infty} e^{2\omega r u^2} \left(u + \sqrt{\frac{1-\sin\theta}{2}} \right)^{k+1-2v} du$$

en na evaluatie van de integralen

$$(9.108) \quad \begin{aligned} V_1(r, \theta) &\sim \sqrt{\frac{2}{\pi}} \Gamma(v) (4r)^{v-1} e^{-\omega r \left(\frac{1+\sin\theta}{2} \right)} \\ &\cdot \sum_{k=0}^{\infty} a_k \frac{D_{k+1-2v}(\sqrt{2\omega r(1-\sin\theta)})}{(4\omega r)^{k/2}} \end{aligned}$$

voor $\omega r \rightarrow \infty$; hierin is $D_\nu(z)$ een parabolische cylinderfunctie.

Deze ontwikkeling voor $V_1(r, \theta)$ geldt alleen in de grenslaag aangezien de reeks zijn asymptotische eigenschappen daarbuiten verliest. Als namelijk het argument van de parabolische cylinderfunctie groot wordt geldt voor vaste k

$$a_k \frac{D_{k+1-2\nu}(\sqrt{2\omega r}(1-\sin\theta))}{(4\omega r)^{k/2}} = O((\omega r)^{\frac{1}{2}-\nu} \exp(-\frac{\omega r}{2}(1-\sin\theta))).$$

De termen in de reeks zijn in dit geval dus van gelijke orde, waardoor de reeks zijn betekenis verliest. Een voordeel is echter dat in (9.108) de eerste term voor $\theta \rightarrow \frac{\pi}{2}$ exact de randfunctie oplevert, namelijk voor $\theta = \frac{\pi}{2}$ geldt

$$a_0 = 2^{1-\nu},$$

$$a_k = 0 \quad \text{voor } k = 1, 2, 3, \dots,$$

zodat

$$\begin{aligned} \lim_{\theta \rightarrow \frac{\pi}{2}} V_1(r, \theta) &= \sqrt{\frac{2}{\pi}} \Gamma(\nu) (4y)^{\nu-1} e^{-\omega y} 2^{1-\nu} D_{1-2\nu}(0) \\ &= e^{-\omega y} y^{\nu-1}. \end{aligned}$$

De ontwikkeling buiten de grenslaag kan gevonden worden op de klassieke manier met de zadelpuntsmethode evenals de ontwikkeling voor de tweede term $V_1(r, -\theta)$ in (9.103). Deze ontwikkeling volgt niet uit het vorige door θ te vervangen door $-\theta$ aangezien de coëfficiënten a_k in (9.108) niet begrensd zijn voor $-\pi/2 \leq \theta \leq 0$. De integrand van $V_1(r, -\theta)$ heeft geen polen dicht bij de oorsprong, zodat de zadelpuntsmethode kan worden toegepast.

Het algemene geval

Wij geven alleen de asymptotische oplossing in de grenslaag en beschouwen eerst weer representatie (9.102) voor $V_1(r, \theta)$ waarbij wij veronderstellen dat

$$(9.109) \quad v_1(r, \theta) = \sum_{n=0}^{\infty} b_n v_1^{(n)}(r, \theta)$$

waarin de b_n via $\phi(y)$ gedefinieerd zijn,

$$(9.110) \quad \phi(y) = \sum_{n=0}^{\infty} b_n y^n$$

en

$$v_1^{(n)}(r, \theta) = \frac{n!}{2\pi \omega^n} \int_{-\infty}^{\infty} e^{-\omega r \cosh t} \cosh(t+i\theta) \frac{dt}{\{1+\sinh(t+i\theta)\}^{n+1}}$$

$$\sim n! \sqrt{\frac{2}{\pi}} (4r)^n e^{-\omega r \left(\frac{1+\sin\theta}{2}\right)} \sum_{k=0}^{\infty} a_k \frac{D_{k-1-2n}(\sqrt{2\omega r(1-\sin\theta)})}{(4r\omega)^{k/2}}.$$

Substitutie van deze ontwikkeling voor $v_1^{(n)}$ in (9.109) geeft dan

$$v_1(r, \theta) \sim \sqrt{\frac{2}{\pi}} e^{-\omega r \left(\frac{1+\sin\theta}{2}\right)} \sum_{n=0}^{\infty} b_n n! (4r)^n \sum_{k=0}^{\infty} a_k(n) \frac{D_{k-1-2n}(\sqrt{2\omega r(1-\sin\theta)})}{(4r\omega)^{k/2}}$$

en het verwisselen van de sommaties levert

$$v_1(r, \theta) \sim \sum_{k=0}^{\infty} \frac{A_k}{(4r\omega)^{k/2}}$$

waarin

$$A_k \sim \sqrt{\frac{2}{\pi}} e^{-\omega r \left(\frac{1+\sin\theta}{2}\right)} \sum_{n=0}^{\infty} b_n n! a_k(n) (4r)^n D_{k-1-2n}(\sqrt{2\omega r(1-\sin\theta)}).$$

Wij zullen de eerste coëfficiënten A_0 en A_1 bepalen, daarvoor hebben wij nodig $a_0(n)$ en $a_1(n)$ die in (9.107) zijn aangegeven met $\nu-1 = n$ en voor de parabolische cylinderfunctie kiezen wij een geschikte integraalrepresentatie, voor $k = 0$

$$D_{-1-2n}(\sqrt{2\omega r(1-\sin\theta)}) = \frac{2^{-n-1}}{n!} \sqrt{2\omega r(1-\sin\theta)} \cdot e^{-\frac{1}{2}\omega r(1-\sin\theta)} \int_0^\infty e^{-\omega r(1-\sin\theta)t} \left(\frac{t}{1+t}\right)^n \frac{dt}{\sqrt{1+t}}.$$

Hieruit volgt met gebruikmaking van (9.110)

$$A_0 = \sqrt{\frac{2}{\pi}} e^{-\omega y} \int_{x\sqrt{\frac{\omega}{v}}}^\infty \phi\left(v - \frac{\omega x^2}{t^2}\right) e^{-\frac{1}{2}t^2} dt$$

met $v = \frac{r}{2} (1+\sin\theta)$.

A_0 is een grenslaagfunctie en levert voor $x = 0$ de randfunctie voor het V-probleem: $e^{-\omega y} \phi(y)$. Deze term lijkt veel op de eerste orde term die Eckhaus en de Jager [3] geven, daar is $v = y = r\sin\theta$. Het voordeel is dan dat de grenslaagterm nul is voor $\theta = 0$ en deze voldoet daarmee aan de randvoorwaarde langs de X-as.

Veel winst levert dit echter niet op aangezien deze term alleen voor $\theta \sim \frac{\pi}{2}$ een rol speelt en niet voor $\theta \sim 0$.

Voor A_1 vinden wij op overeenkomstige wijze

$$A_1 = \sqrt{\frac{2}{\pi}} \frac{e^{-\omega r}}{\omega r} \frac{\sqrt{1-\sin\theta}}{\sqrt{1+\sin\theta}} \int_0^\infty e^{-\omega r(1-\sin\theta)t} \cdot \left\{ \frac{vt}{1+t} \phi' \left(\frac{vt}{1+t} \right) + \left(\frac{vt}{1+t} \right)^2 \phi'' \left(\frac{vt}{1+t} \right) \right\} dt,$$

terwijl $A_2, A_3, \dots$ kunnen worden bepaald indien $a_2, a_3, \dots$ bekend zijn.

Opmerking 9.8. Zoals eerder is opgemerkt is de ontwikkeling (9.108) niet geldig voor het gehele kwartvlak maar slechts in de grensmaan. Nader onderzoek heeft uitgewezen dat een analoge ontwikkeling als in (9.108) kan worden afgeleverd, die wel voor het gehele kwartvlak geldig is. Hiervoor wordt verwezen naar [8].

Literatuur

- [1] Coolen, T.M.T.,
Förch, G.J.R.,
Jager, E.M. de,
Pijls, H.G.J. Colloquium elliptische differentiaal-
vergelijkingen, deel 1, M.C. Syllabus 9.1.
Mathematisch Centrum, Amsterdam, 1969.
- [2] Courant, R.,
Hilbert, D. Methods of Mathematical Physics.
Vol. II, Intersc. Publ., 1962.
- [3] Eckhaus, W.,
Jager, E.M. de Asymptotic solutions of singular
perturbation problems for linear differential
equations of elliptic type.
Archive for Rat. Mech. and Analysis 23.
no. 1, (1966).
- [4] Grasman, J. On singular perturbations and parabolic
boundary layers.
Journ. Engin. Math. 2(1968), 163-172.
- [5] Groen, P.P. de An elliptic singular perturbation problem
for non-differential parameters.
MC rapport ZW-1970-001,
Mathematisch Centrum, Amsterdam, 1970.
- [6] Levinson, N. The first boundary value problem.
Annals of Mathematics 51(1950).
- [7] Oberhettinger, F. On a modification of Watson's lemma.
J. Res. Nat. Bur. Stand 63B(1959), 15-17.
- [8] Temme, N.M. Saddle point methods for a singular
perturbation problem.
Rapport TW 121 (verschijnt in 1970)
Mathematisch Centrum, Amsterdam.
- [9] Višik, M.I.,
Lyusternik, L.A. Regular degeneration and boundary layer
for linear differential equations with
small parameter.
Uspehi Mat. Nauk 12(1957).
(American Mathematical Society Translation
Series 2, 20(1962)).

- [10] Waerden, B.L. van der On the method of saddle points.
Appl. Sci. Res. 82(1951), 33-45.
- [11] Wasow, W. Asymptotic solution of boundary value
problems for the differential equation
$$\Delta u - \lambda \frac{\partial u}{\partial x} = \lambda f(x,y).$$

Duke Math. J. 11(1944).

10. OPTIMAL CYCLIC ALTERNATING DIRECTION ITERATION

10.1. Introduction

Beside the methods discussed in section 6 there is another class of methods under the general title of alternating direction methods (see chapter 7 of reference [4] and chapter 6 of reference [5] and the associated bibliographies and discussions).

I shall consider here only two-dimensional problems. For extensions to n dimensions, see for instance [3].

First I shall make a synthesis of the methods of Peaceman-Rachford [1], Douglas-Rachford [2], Douglas [3] and their generalizations by Varga and Wachspress.

10.2. Alternating direction methods

As in section 6 we shall start by writing the discretization of the problem in the form

$$(10.1) \quad Ax = b$$

and the matrix A in the form

$$(10.2) \quad A = A_1 + A_2$$

and in the form

$$(10.3) \quad A = B_1 + B_2,$$

where in the applications A_1 and B_1 are block diagonal matrices, each block of A_1 corresponding to a row of sample points of the discretized problem, and each block of B_1 corresponding to a column of sample points.

A mapping T to solve (10.1) by iteration can be defined implicitly by

$$(10.4) \quad \begin{cases} A_1 x^{(i+\frac{1}{2})} = b - A_2 x^{(i)} \\ B_1 x^{(i+1)} = b - B_2 x^{(i+\frac{1}{2})}, \end{cases}$$

giving

$$(10.5) \quad x^{(i+1)} = T x^{(i)}.$$

Alternately, we can also define a mapping T by

$$(10.6) \quad \begin{cases} C_1 x^{(i+\frac{1}{2})} = b - C_2 x^{(i)} \\ D_1 x^{(i+1)} = D_2 x^{(i+\frac{1}{2})} + (D_1 - D_2) x^{(i)}, \end{cases}$$

where again

$$(10.7) \quad A = C_1 + C_2.$$

It follows from the form of (10.4) and (10.6) and from (10.2), (10.3) and (10.7) that if the iteration associated to T converges, it converges to a solution of (10.1).

Theorem 10.1.

If A_1 , B_1 , C_1 and D_1 have an inverse, if A_2 , B_2 and C_2 satisfy (10.2), (10.3) and (10.7), the mappings associated to (10.4) and (10.6) are identical if and only if

$$(10.8) \quad B_1 D_1^{-1} D_2 = C_1 - B_2 A_1^{-1} C_1,$$

and in particular if

$$(10.9) \quad D_1 = B_1 \text{ and } D_2 = C_1 - B_2 A_1^{-1} C_1.$$

Indeed, if we eliminate $x^{(i+\frac{1}{2})}$ we get

$$\begin{cases} B_1 x^{(i+1)} = (I - B_2 A_1^{-1}) b + B_2 A_1^{-1} A_2 x^{(i)}, \\ D_1 x^{(i+1)} = D_2 C_1^{-1} b + (D_1 - D_2 - D_2 C_1^{-1} C_2) x^{(i)}. \end{cases}$$

The identity of the coefficients of $x^{(i)}$ follows from the fact that if

$x^{(i)} = x^{(i+1)}$ both equations are equivalent to $Ax = b$.

Let $A = H + V + R$. We have the schemes of table 10.1.

Table 10.1

	A_1	$B_1=D_1$	C_1	D_2
Peaceman-Rachford [1]	$H+\frac{1}{2}R+wI$	$V+\frac{1}{2}R+wI$	$H+\frac{1}{2}R+wI$	$2wI$
Douglas-Rachford [2]		$V+\frac{1}{2}R+wI$	$H+\frac{1}{2}R+wI$	wI
Varga [4, (7.71)]		$V+\frac{1}{2}R+wI$	$H+\frac{1}{2}R+wI$	$(2-\beta)wI$
E.L. Wachspress [5, (6-3)]	$H+R+w_h I$	$V+R+w_v I$		
E.L. Wachspress [5, (6-4)]		$V+R+w_v I$	$H+R+w_h I$	$\frac{R}{2} + \frac{w_v+w_h}{2} I$
E.L. Wachspress [5, (6-9)]	$H+R+w_h I$	$V+R+w_v I$	$\frac{1}{2}(H+R+w_h I)$	$\frac{R}{2} + \frac{w_v+w_h}{2} I$

If the partial differential equation is

$$\frac{\partial}{\partial x} (p(x,y) \frac{\partial}{\partial x} u(x,y)) +$$

$$\frac{\partial}{\partial y} (q(x,y) \frac{\partial}{\partial y} u(x,y)) +$$

$$r(x,y) u(x,y) = s(x,y),$$

and we discretize using sample points in a rectangular mesh and basic functions of the form $G_1(x) G_2(y)$, H corresponds to the discretization of the first line, V to that of the second and R to that of the third.

10.3. Cyclic alternating direction method

If we observe that the mapping T_w depends on one parameter w (or on two parameters w_h and w_v), we can deduce another mapping by choosing

$$(10.10) \quad T = T_{w_r} T_{w_{r-1}} \dots T_{w_1},$$

and similarly for two parameters.

10.4. Optimal choice for the Peaceman-Rachford method in the simplest case

We shall discuss in some detail the optimal choice for the Peaceman-Rachford method, at least under special conditions. To get into the problem we shall first assume that $r = 1$.

The rate of convergence is related to the spectral norm ρ of the matrix M associated to T , namely such that

$$(10.11) \quad Tx = Mx + Nb,$$

and we therefore wish to minimize $\rho(M)$.

To enable us to make any headway we have to assume that the matrices

$$(10.12) \quad H_1 = H + \frac{1}{2}R \text{ and } V_1 = V + \frac{1}{2}R$$

have the same eigenvectors, which can also be phrased that H_1 and V_1 commute, because of

Theorem 10.2 (Frobenius). If H_1 and V_1 are Hermitian matrices, then there exists an orthogonal basis of eigenvectors u_i of H_1 and V_1 if and only if $H_1V_1 = V_1H_1$.

For the proof we refer for instance to [4, p. 220].

If u is an eigenvector and the corresponding eigenvalues are λ for V_1 , μ for H_1 and σ for M , we have for the Peaceman-Rachford method

$$(10.13) \quad \begin{cases} (\mu+w)u^{(\frac{1}{2})} = (-\lambda+w)u, \\ (\lambda+w)\sigma u = (-\mu+w)u^{(\frac{1}{2})}, \end{cases}$$

and therefore

$$(10.14) \quad \sigma = \frac{\mu-w}{\mu+w} \cdot \frac{\lambda-w}{\lambda+w}.$$

This defines a function S :

$$S(\lambda, \mu, w) = \sqrt{\sigma}$$

of three variables. Let us first assume that $\lambda = \mu$.

The optimization consists in solving the min-max problem

$$(10.15) \quad \min_w \max_{\lambda_i} |S(\lambda_i, \lambda_i, w)|.$$

The function s , with

$$s(\lambda) = \frac{\lambda - w}{\lambda + w}$$

is monotonic, therefore its extrema occur for the largest or smallest admissible value of λ .

We shall now assume that H_1 and V_1 are positive definite, and that

$$(10.16) \quad (0 <) \min \{\lambda_i\} = \min \{\mu_i\} = b (<) \max \{\lambda_i\} = \max \{\mu_i\} = a.$$

The min-max problem is then reduced to

$$(10.17) \quad \min_w \max \left(\left| \frac{a-w}{a+w} \right|, \left| \frac{b-w}{b+w} \right| \right).$$

w_0 , the optimal w , follows from geometric considerations (see figure 10.1)

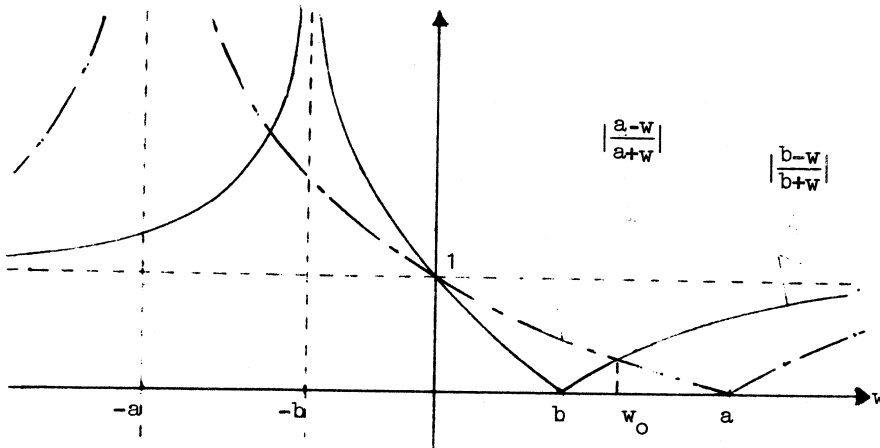


Figure 10.1

and satisfies

$$\begin{cases} \text{a)} & b < w_0 < a, \\ \text{b)} & \left| \frac{a-w_0}{a+w_0} \right| = \left| \frac{b-w_0}{b+w_0} \right|, \end{cases}$$

therefore w_0 is the positive solution $\sqrt{ab}$ of

$$\frac{w_0 - b}{w_0 + b} = \frac{a - w_0}{a + w_0}$$

and the extremum is $(1 - \sqrt{\frac{b}{a}}) / (1 + \sqrt{\frac{b}{a}})$.

To generalize we have

- a) to consider $r > 1$,
- b) to relax the condition on the range of the eigenvalues for H_1 and V_1 ,
- c) to consider the possibility of $w_v \neq w_h$.

10.5. The simple cyclic case for Peaceman-Rachford

Let us first examine a), we shall consider b) and c) in section 10.12.

The min-max problem is

$$(10.18) \quad \min_{w_j} \max_{\lambda_i, \mu_i} \prod_{j=1}^r \sqrt{\left| \frac{\mu_i - w_j}{\mu_i + w_j} \right| \cdot \left| \frac{\lambda_i - w_j}{\lambda_i + w_j} \right|}.$$

Again we assume (10.16). We shall replace (10.18) by the continuous analogue

$$(10.19) \quad \min_{w_j} \max_{\substack{x \in [b, a] \\ y \in [b, a]}} \prod_{j=1}^r \sqrt{\left| \frac{x - w_j}{x + w_j} \right| \cdot \left| \frac{y - w_j}{y + w_j} \right|}.$$

The error thus made is discussed in section 10.6 after b⁴. Moreover, if w_j is outside (b, a) the corresponding factor is extremal when $w_j = a$ or b ; we can therefore restrict ourselves to $w_j \in [b, a]$.

10.6. A classical min-max problem

Consider the following problem

$$(10.20) \quad \kappa_0 = \min_{w_j \in [b, a]} \max_{x \in [b, a]} \frac{\prod_{j=1}^r (x - w_j)}{\prod_{j=1}^r (x + w_j)}, \quad 0 < b < a.$$

We have, see for instance Achieser [6, chapter II],

Theorem 10.3

There exists an unique solution of problem (10.20) with $r + 1$ extrema all equal in modulus and alternating in sign.

Proof

$$a) \quad f(w_1, \dots, w_r) = \max_{x \in [b, a]} \left| \frac{\prod_{j=1}^r (x - w_j)}{\prod_{j=1}^r (x + w_j)} \right|$$

is continuous and bounded from above by $(\frac{a-b}{2b})^r$ in $[b, a]^r$. It therefore achieves its minimal value for some $(w_1, \dots, w_r) \in [b, a]^r$.

b) Let $P(x) = \prod_{j=1}^r (x - w_j)$, $Q(x) = P(x) / P(-x)$. Let $P_{\text{opt.}}$ (or $Q_{\text{opt.}}$) be an optimal P (or Q), and let $\kappa(Q) = \max_{x \in [b, a]} |Q(x)|$. We wish to show that

$b_1)$ if Q has $r + 1$ local extrema $m_i = Q(x_i)$ with alternating sign, then

$$\kappa_0 \geq \min |m_i|,$$

$b_2)$ if Q_1 has $r + 1$ local extrema m_i with alternating sign such that $|m_i| = \kappa'$, for all i , then $\kappa_0 = \kappa'$ and two of these extrema occur necessarily at b and a ,

$b_3)$ if the number t of points x_i for which $\kappa_2 = |Q_2(x_i)| = \kappa(Q_2)$ is less than $r + 1$, then we can find a polynomial such that for the corresponding Q , $\kappa(Q) < \kappa(Q_2)$,

$b_4)$ $P_{\text{opt.}}$ is unique.

Because the optimal solution achieves its extremal value on the boundary, if the number of eigenvalues μ_i is large and evenly distributed, the improvement that can be achieved when $r > 1$ for the discrete problem is slight, because of the existence of some μ_i near a value v_j for which the

ational function Q has an interior extremum and the error $Q(u_i) - Q(v_j)$ is of the order of $(u_i - v_j)^2$.

We now give a proof of b_1 to b_4 :

) let $\kappa_0 < \min |m_i|$ and

$$\begin{aligned} 10.21) \quad V(x) &= (Q(x) - Q_{\text{opt.}}(x)) \cdot P(-x) \cdot P_{\text{opt.}}(-x) \\ &= P(x) \cdot P_{\text{opt.}}(-x) - P(-x) \cdot P_{\text{opt.}}(x). \end{aligned}$$

$V(x_i)$ has the same sign as $Q(x_i)$, therefore the values $V(x_i)$ alternate in sign and therefore V has r roots ρ_j in $[b, a]$. But V is an odd polynomial and has a degree of at most $2r-1$. V having $2r$ distinct roots $\pm \rho_j$ we must have $V = 0$, which is in contradiction with the hypothesis.

2) Q_1 has at least $r - 1$ interior extrema y_i for which $V(x) = P(-x) \cdot DP(x) + P(x) \cdot DP(-x)$ is zero (D is the derivative operator); but V is an even polynomial which has at most $2r - 2$ roots $\pm y_i$, therefore Q_1 cannot have more than $r - 1$ interior extrema and therefore not more than $r + 1$ local extrema; from b_1) $\kappa_0 \geq \kappa'$, therefore Q_1 is optimal.

3) Lemma 10.1

If P, Q and S are polynomials of degree p, q and $s \leq p + q - 1$, and P and Q are relatively prime, then there exist polynomials R and T , having a degree of at most $q - 1$ and $p - 1$, such that

$$10.22) \quad PR + QT = S.$$

By the Euclidian algorithm applied to P and Q we can construct polynomials V and W such that $PV + QW = 1$, therefore

$$PVS + QWS = S.$$

If we divide VS by Q , and if Q_1 is the quotient and R is the remainder we have

$$Q(PQ_1 + WS) + PR = S.$$

Because the degree of $S - PR$ is at most $p + q - 1$, the degree of

$$T = PQ_1 + WS$$

is at most $p - 1$.

Let κ' be the maximum of the absolute values of the local extrema different from κ_2 . Let us extract from the roots of P_2 those ζ_j which separate points x_i for which the extrema are of opposite sign; by hypothesis and from 2) there are at most $r - 1$ of them.

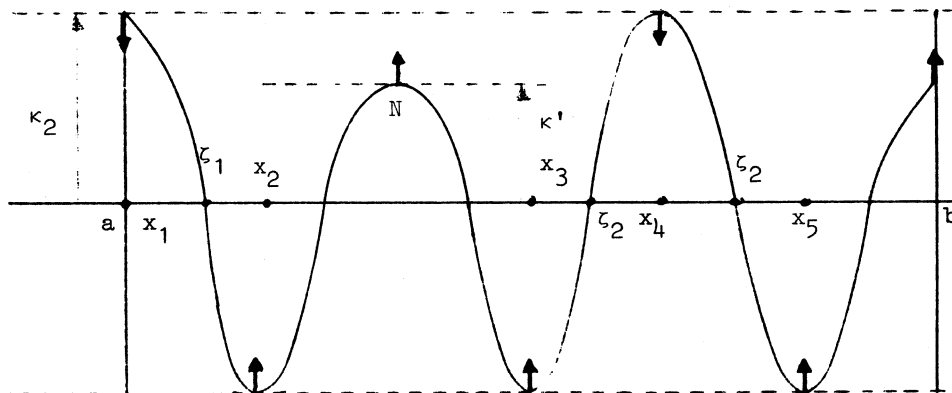


Figure 10.2

Let $S(x) = \prod_{j=1}^s (x - \zeta_j)$, from the lemma follows the existence of polynomials R and T of degree at most $r - 1$ such that

$$R(-x) \cdot P_2(x) - T(x) \cdot P_2(-x) = x \cdot S(x) \cdot S(-x).$$

If we replace x by $-x$ and add we get

$$(R(-x) - T(-x)) P_2(x) = (T(x) - R(x)) P_2(-x),$$

therefore $P_2(x)$ divides $R(x) - T(x)$ which is of less degree and therefore $R(x) = T(x)$.

Consider

$$P = P_2 - \rho R$$

and the corresponding Q :

$$Q(x) = P(x) / P(-x).$$

$$Q(x) - Q_2(x) = \frac{\rho x S(x) S(-x)}{(P_2(-x) - \rho R(-x)) P_2(-x)}.$$

If ρ is sufficiently small the denominator is positive, and it is sufficient to choose the sign of ρ to be that of $-Q_2(x_1)$ to have $Q(x_1) - Q_2(x_1)$ to be of the opposite sign of $Q_2(x_1)$, and ρ also small enough so that the height of the point N does not increase by more than $(\kappa_2 - \kappa')/2$, say.

-) If there are two polynomials P and P_{opt} which are optimal, let x_i be the positions of the extrema of P_{opt} ; we know that they are $r + 1$ in number.

$$\pm Q(x_i) \leq |Q(x_i)| \leq |Q_{\text{opt.}}(x_i)|,$$

therefore $(Q_{\text{opt.}} - Q)(x_i)$ has the sign of $Q_{\text{opt.}}(x_i)$ or is zero. Let y_i be those t points x_i for which it is zero, and let

$$S(x) = \prod_{i=1}^t (x - y_i).$$

If V is defined as in (10.21), V is divisible by $S(x) \cdot S(-x)$ and the quotient of degree $2(r-t) - 1$ has in $[b, a]$ $r + 1 - t$ values alternating in sign and therefore at least $2(r-t)$ roots, therefore $V = 0$.

10.7. The case $r = 2^p$ and heuristics

Unicity enables us to solve the problem entirely when $r = 2^p$. Indeed, if Q is the optimal solution of (10.20) then

$$(10.23) \quad Q(x, w_j) = \frac{\prod_{j=1}^r (x - w_j)}{\prod_{j=1}^r (x + w_j)} = \pm \frac{\prod_{j=1}^r \left(\frac{ab}{x} - \frac{ab}{w_j} \right)}{\prod_{j=1}^r \left(\frac{ab}{x} + \frac{ab}{w_j} \right)} = Q(y, \frac{ab}{w_j}),$$

where

$$y = \frac{ab}{x}.$$

The last function has the same domain $[b, a]$ and has the same form; its roots must therefore be the same and

$$(10.24) \quad \frac{ab}{w_j} = w_{r-j+1} \quad (j=1 \dots r).$$

In particular, if $r = 1$ we have at once the result of section 10.4, and if r is odd, one of the roots is $\sqrt{ab}$.

If r is even, we can combine the corresponding factors as follows

$$(10.25) \quad \frac{(x - w_j)(x - \frac{ab}{w_j})}{(x + w_j)(x + \frac{ab}{w_j})} = \frac{\frac{1}{2}(x + \frac{ab}{x}) - \frac{1}{2}(w_j + \frac{ab}{w_j})}{\frac{1}{2}(x + \frac{ab}{x}) + \frac{1}{2}(w_j + \frac{ab}{w_j})}.$$

If we set

$$(10.26) \quad y = \frac{1}{2}(x + \frac{ab}{x})$$

and

$$(10.27) \quad v_j = \frac{1}{2}(w_j + \frac{ab}{w_j}),$$

we see that the problem is reduced to that of $r/2$ and the interval

$$(10.28) \quad [\sqrt{ab}, \frac{a+b}{2}];$$

from v_j we derive w_j using (10.27).

If r is a multiple of 2^p we reduce the problem to the case $2^{-p}r$, and we can solve the problem entirely if $r = 2^p$. Then the successive intervals are the successive a_i, b_i of the arithmetic-geometric mean iteration [10, 17.6.1]*)

$$(10.29) \quad \begin{cases} a_{i+1} = \frac{a_i + b_i}{2}, & b_{i+1} = \sqrt{a_i b_i}, & c_{i+1} = \frac{1}{2}(a_i - b_i) \\ a_0 = a, & b_0 = b \end{cases}$$

studied by Gauss; (10.26) is the rational form of the Gauss transformation [7, p. 197], and the extremum is

$$(10.30) \quad \kappa_{2^p} = \frac{1 - \sqrt{k'}}{1 + \sqrt{k'}}^p, \text{ where } k'_i = b_i/a_i, k_i = c_i/a_i.$$

This suggests at once the connection with elliptic functions. Moreover, in view of the properties of the dn function of Jacobi [10, 16.8.3]:

$$(10.31) \quad \text{dn}(u+K, k) \cdot \text{dn}(u, k) = k' = b/a$$

and [10, 16.14.4]

$$(10.32) \quad a_1 \text{dn}\left(\frac{2a_1 u}{a}, k_1\right) = \frac{a}{2}(\text{dn}(u, k) + \text{dn}(u+K, k)) \\ = \frac{1}{2}(a \cdot \text{dn}(u, k) + \frac{b}{\text{dn}(u, k)})$$

where $k = \sqrt{1-k'^2}$, we see that

*) We have given, whenever possible, a precise reference to the Handbook of Mathematical Functions [10]; the proof can be traced to the associated references, where we recommend Whittaker and Watson [15].

$$(10.33) \quad w_j = a \cdot \operatorname{dn}\left(\frac{2j-1}{2r} K, k\right) \text{ when } r = 2^p,$$

and we can conjecture that the same formula is true in general.

10.8. The solution for arbitrary r

Theorem 10.4. (W.B. Jordan)

$$(10.34) \quad \frac{\operatorname{dn}-\sqrt{k'_p}}{\operatorname{dn}+\sqrt{k'_p}} (Mu, k_p) = \prod_{j=1}^r \frac{\operatorname{dn}-d_j}{\operatorname{dn}+d_j} (u, k),$$

where

$$(10.35) \quad d_j = \operatorname{dn}\left(\frac{2j-1}{2r} K, k\right);$$

k_p is the modulus and $k'_p = \sqrt{1-k_p^2}$ the complementary modulus of the Jacobi transformation of degree [9]

$$(10.36) \quad r = \frac{K_k / K'_k}{K_{k_p} / K'_{k_p}},$$

where K_k, K'_k are the complete elliptic integrals of the first kind associated to k and k' , and

$$(10.37) \quad M = K'_{k_p} / K'_k.$$

The Landen transformation corresponds to the special case $r = 2$.

Theorem 10.5. (W.B. Jordan)

The solution of

$$(10.38) \quad \min_{w_j} \max_{x \in [k', 1]} \left| \prod_{j=1}^r \left(\frac{x-w_j}{x+w_j} \right) \right|, \quad 0 < k' < 1,$$

is given by

$$(10.39) \quad w_j = d_j,$$

and the extrema occur at

$$(10.40) \quad m_j = \operatorname{dn} \left(\frac{jK}{r}, k \right), \quad j = 0, \dots, r,$$

and are equal in modulus to

$$(10.41) \quad \rho_r = \frac{1 - \sqrt{k'}}{1 + \sqrt{k'}} \frac{p}{p}.$$

The proof of theorem 10.5 follows from theorem 10.4 if we observe that the function of the first member of (10.34) is periodic with a real period $2K_k / M = 2K_k / r$ and has the extremal values (10.41); if we set $x = \operatorname{dn}(w)$ in (10.38), we have a rational function of the right form. The proof of theorem 10.4 depends partly on

Theorem 10.6

If two doubly periodic functions with the same periods have the same zeros and the same poles (with the same multiplicity) their ratio is constant [11, application of theorem 1, p. 74].

The proof of theorem 10.4 also depends on the following properties of the function dn [10, 16.2]: dn has periods $2K$ and $4iK'$; exactly twice it has each complex value in the rectangle of the periods [11, p. 76], its poles are $\pm iK'$ and [10, 16.8.3 and 16.5.2]

$$(10.42) \quad \operatorname{dn}(u + 2iK', k) = -\operatorname{dn}(u, k).$$

Finally

$$(10.43) \quad \operatorname{dn}\left(\pm \frac{K}{2}, k\right) = \sqrt{k'}.$$

The functions occurring in (10.34) have the same imaginary period if (10.37) holds. The left hand function of (10.34) is of the real period $2K_{k_p} / M$ and therefore also of the period $2rK_{k_p} / M$, and this is equal to $2K_k$ if (10.36) holds. The zeros of the left hand function of (10.34) occur at

$$\frac{1}{M} \left(\pm K_{k_p} / 2 + j \cdot K_{k_p} \right)$$

and its poles at

Start with an arbitrary point $P_1 = M_1$, determine M_{2r+1} on the circle such that M_1LM_{2r+1} is a straight line; draw the circle c with centre on AB and horizontal tangents passing through M_1 and M_{2r+1} ; draw from M_1 the successive tangents to c , meeting C at $P_3, P_5, \dots$, the r th tangent giving the point P_{2r+1} ; if $P_{2r+1} = M_{2r+1}$ then $P_{2j-1} = M_{2j-1}$; because of the monotonic and continuous behaviour on the circumference $AP_1P_3 \dots B$ it is easy to see what to do if P_{2r+1} is further on the circumference than M_{2r+1} , namely P_1 must be chosen closer to A etc. When we have succeeded, we continue the construction and obtain by symmetry a closed polygon of $2r$ sides inscribed in C and circumscribed to c . By a theorem of Poncelet, the same will be true if we start the construction from any point of C ; in particular if we start from $A = M_0$, the successive tangents give the points $M_2, M_4, \dots, M_{2r} = B$. The circles C, c and the point circle L have the same (horizontal) radical axis at a distance $b^2 / (a-b)$ above B .

10.10. Mechanical interpretation

We consider a circular motion (diameter $1+k'$) of a mass in the vertical gravitational field, such that when the mass is at A it has velocity V_A and when it is at B it has velocity $k' \cdot V_A$, where A and B are the same points as in figure 10.3. If half the period is ρK , the mass at time $\frac{\rho K}{2r} j$ is at the point M_j given in the geometrical interpretation. Moreover, if $2r$ masses, having the same circular motion, are at some time in M_{2j} , $j = 0(1)2r-1$, their centre of gravity also carries out a circular motion and the ratio of the velocity at the highest and lowest position of the centre of gravity is k'_1 . The proof will be sketched, some of the details may be found in [8].

The motion in a vertical gravitational field of a mass m constrained on the circle

$$x^2 + (z-R)^2 = R^2$$

is given by

$$\left\{ \begin{array}{l} m \frac{d^2 x}{dt^2} = (mg \frac{x}{r}) \cdot \frac{z-R}{R} \\ m \frac{d^2 z}{dt^2} = - (mg \frac{x}{R}) \cdot \frac{x}{R} , \end{array} \right.$$

because the force is the projection of the weight mg on the tangent to the circle, i.e. $mg \frac{x}{R}$, and because the unit vector along the tangent has horizontal and vertical components $\frac{z-R}{R}$ and $-\frac{x}{R}$.

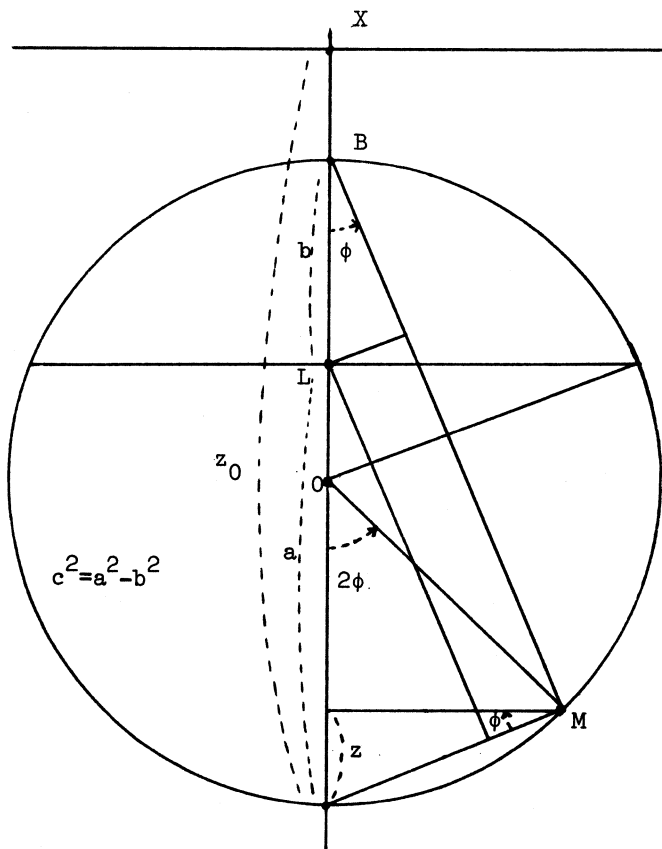
If we multiply respectively by $\frac{dx}{dt}$, $\frac{dy}{dt}$, and if we add, integrate and multiply by $2/m$ we get

$$v^2 = \frac{d\vec{x}}{dt} \cdot \frac{d\vec{x}}{dt} = 2g (z_0 - z),$$

where z_0 is a constant of integration, whose physical interpretation is the height at which the mass has zero velocity.

Rewriting in polar coordinates we get

$$(2R)^2 \left(\frac{d\phi}{dt}\right)^2 = 2g(z_0 - 2R \sin^2 \phi).$$



A(t=0)

Figure 10.4

On the other hand, if $XL^2 = XB \cdot XA$ then

$$z_0 - 2R = XB = \frac{b^2}{a-b}, \quad z_0 = XA = \frac{a^2}{a-b}, \quad 2R = \frac{a^2 - b^2}{a-b} = \frac{c^2}{a-b}.$$

Therefore

$$\frac{(2R)^2}{2g} (a-b) \left(\frac{d\phi}{dt}\right)^2 = a^2 - c^2 \sin^2 \phi,$$

and with an appropriate choice of units

$$(10.44) \quad \left(\frac{d\phi}{dt}\right)^2 = a^2 - c^2 \sin^2 \phi = a^2 \cos^2 \phi + b^2 \sin^2 \phi,$$

or

$$(10.45) \quad t = \int_0^\phi \frac{d\phi}{\sqrt{a^2 \cos^2 \phi + b^2 \sin^2 \phi}} = \int_0^\phi \frac{d\phi}{\sqrt{a^2 - c^2 \sin^2 \phi}},$$

which defines a function $(\phi, at=u)$.

The inverse function (u, ϕ) is noted am (am for amplitude) [10, 16.1.4] and

$$(10.46) \quad \phi = \text{am}(u).$$

The Jacobi elliptic functions are [10, 16.1.5]

$$(10.47) \quad \text{sn} = \sin \bullet \text{am}, \quad \text{cn} = \cos \bullet \text{am}, \quad \text{dn} = \sqrt{\text{cn}^2 + \frac{b^2}{a^2} \text{sn}^2},$$

with parameter

$$k = c/a,$$

and

$$(10.48) \quad LM^2 = LP^2 + LQ^2 = b^2 \sin^2 \phi + a^2 \cos^2 \phi = a^2 \text{dn}^2 u.$$

Half the pendular period is [10, 16.6.1]

$$(10.49) \quad \frac{1}{a} \int_0^{\pi/2} \frac{d\phi}{\sqrt{1-k^2 \sin^2 \phi}} = \frac{1}{a} K_k.$$

The property the centre of gravity of the $2r$ masses has, is an interpretation of formulas analogous to those of Jacobi [9, p. 48, 23 (16)-(17) and remark top of p. 43]:

$$(10.50) \quad \left\{ \begin{array}{l} \frac{k_p M}{k} \operatorname{sn}(Mu, k_p) = \sum_{j=0}^{n-1} \operatorname{sn}(u+u_j, k), \\ \frac{k_p M}{k} \operatorname{cn}(Mu, k_p) = \sum_{j=0}^{n-1} \operatorname{cn}(u+u_j, k), \\ u_j = \frac{4K}{n} j, \quad n = 4r, \quad k_p \text{ and } M \text{ as above.} \end{array} \right.$$

Namely

$$(10.51) \quad \frac{k_p^2 M^2}{k^2} S(Mu, k_p) = \sum_{j=0}^{n-1} S(u+t_j, k),$$

$$(10.52) \quad \frac{k_p^2 M^2}{k^2} C(Mu, k_p) + \rho = \sum_{j=0}^{n-1} C(u+t_j, k),$$

where $S = \operatorname{sn} \cdot \operatorname{cn}$, $C = \operatorname{cn}^2 - \operatorname{sn}^2$, ρ is some constant, $t_j = \frac{2K}{n} j$, $n = 2r$. The proof depends on

Theorem 10.7

Two doubly periodic functions with the same poles and the same principal parts at all poles differ by a constant [11, p. 76].

S and C have real periods $2K$, and imaginary periods $4iK'$ and $2iK'$ respectively and double poles at $\pm iK'$ with principal part respectively

$$\pm \frac{1}{ik^2 u^2} \quad \text{and} \quad -\frac{2}{k^2 u^2}.$$

For (10.51) the constant is zero by symmetry, for (10.52) it has no simple expression, but can be determined by setting for instance $u = 0$ in (10.52).

The proof required for the geometric interpretation and for the Poncelet theorem is as follows: let M and N execute the same pendular movement. Let

$$\frac{UM}{UN} = \frac{MM_1}{NN_1}.$$

But

$$\lim_{M_1 \rightarrow M} \frac{MM_1}{NN_1} = \frac{V_M}{V_N},$$

therefore

$$\lim_{M_1 \rightarrow M} \frac{UM}{UN} = \frac{TM}{TN},$$

and the curve to which MN remains tangent touches MN at T, such that we have (10.53). Such a curve is a circle which has with the given circle the radical axis XS in common.

The geometrical interpretation provides a method for the determination of the optimal parameters. First, we can as in Wachspress [5, p.199], when $r = 2^p \cdot r'$, r' odd, determine from the optimal parameters for the case r' those for r using $(2^p - 1) \cdot r'$ square roots (see (10.27)). For any value of r' we can obtain the largest parameter $w_1 = \text{dn}(\frac{K}{2r'})$ using a Landen transformation [10, 17.5.12]. The other parameters $w_j = \text{dn}(\frac{(2j-1)K}{2r'})$ can be obtained by rational operations. From the geometric construction it follows that the coordinates of M_{2i+1} are rational functions of those of M_{2i-1} and M_{2i-3} ; this led us to the relations linking the sn and cn functions and by analogy to (10.57) we have

$$(10.54) \quad w_{j+1} \cdot w_{j-1} = \frac{k^2(w_j^2 + w^2 - 1) + (1 - w^2)(1 - w_j^2)}{k^2 - (1 - w^2)(1 - w_j^2)},$$

or

$$(10.55) \quad w_{j+1} \cdot w_{j-1} = \frac{w_j^2(w^2 - k'^2) + (1 - w^2)k'^2}{w_j^2(1 - w^2) + (w^2 - k'^2)},$$

where

$$(10.56) \quad w = \text{dn}(\frac{K}{r'}) = \frac{w_1^4 - 2k'^2 w_1^2 + k'^2}{-w_1^4 + 2w_1^2 - k'^2} \text{ and } w_0 = w_1.$$

The proof depends on the addition formulas for the dn function of Jacobi [10, 16.17.3] and on

$$w_{j+1} = \operatorname{dn}\left(\left(\frac{2j+1}{2}\right)u_1\right), w = \operatorname{dn}(u_1), u_1 = \frac{K}{r'}.$$

Indeed,

$$\begin{aligned} \operatorname{dn}(u+u_1) \cdot \operatorname{dn}(u-u_1) &= \frac{\operatorname{dn}^2 u \cdot \operatorname{dn}^2 u_1 - k^4 \operatorname{sn}^2 u \cdot \operatorname{sn}^2 u_1 \cdot \operatorname{cn}^2 u \cdot \operatorname{cn}^2 u_1}{(1-k^2 \operatorname{sn}^2 u \cdot \operatorname{sn}^2 u_1)^2}, \\ \operatorname{dn}(u+u_1) \cdot \operatorname{dn}(u-u_1) &= \frac{1-k^2(\operatorname{sn}^2 u + \operatorname{sn}^2 u_1) + k^2 \operatorname{sn}^2 u \cdot \operatorname{sn}^2 u_1}{1-k^2 \operatorname{sn}^2 u \cdot \operatorname{sn}^2 u_1}, \\ (10.57) \quad \operatorname{dn}(u+u_1) \cdot \operatorname{dn}(u-u_1) &= \frac{k^2(\operatorname{dn}^2 u + \operatorname{dn}^2 u_1 - 1) + (1-\operatorname{dn}^2 u)(1-\operatorname{dn}^2 u_1)}{k^2 - (1-\operatorname{dn}^2 u)(1-\operatorname{dn}^2 u_1)} \end{aligned}$$

(see also [9, p. 88-51(7)]).

10.11. Rate of convergence

The rate of convergence for the mapping T is

$$-\ln(\rho_r^2).$$

The average rate of convergence for the mapping T_{w_i} is therefore

$$(10.58) \quad \gamma = -\ln(\rho_r^{2/r}).$$

We have (see (10.29) and (10.30)):

$$\sqrt{k_2} = \sqrt{\frac{c_2}{a_2}} = \sqrt{\frac{a_1 - b_1}{a_1 + b_1}} = \frac{c_1}{a_1 + b_1} = \frac{a_0 - b_0}{a_0 + b_0 + 2\sqrt{a_0 b_0}} = \frac{\sqrt{a_0} - \sqrt{b_0}}{\sqrt{a_0} + \sqrt{b_0}} = \frac{1 - \sqrt{k'_0}}{1 + \sqrt{k'_0}},$$

therefore, if k'_q is the value derived from k'_p after two iterations (10.29), we have

$$(10.59) \quad \rho_r = \frac{1 - \sqrt{k'_p}}{1 + \sqrt{k'_p}} = \sqrt{k'_q},$$

and k_q is related to the nome q_q by [10, 16.38.5 and 16.38.7]

$$(10.60) \quad \sqrt{k_q} = \frac{2q_q^{\frac{1}{4}}(1+q_q^2+q_q^6+\dots+q_q^{i(i+1)}+\dots)}{1+2q_q+2q_q^4+\dots+2q_q^{i^2}+\dots}.$$

Moreover, the relation between the nomes q , q_p and q_q is

$$(10.61) \quad q_q = q_p^4, \quad q_p = q^r,$$

because [10, 16.27]

$$(10.62) \quad q = e^{-\pi K'/K}.$$

Therefore we have

$$(10.63) \quad \gamma = -\ln q^2 - \frac{2}{r} \ln 2 - \frac{2}{r} \ln \frac{1+q_q^2+\dots}{1+2q_q+\dots},$$

from which follows the rate of convergence for large values of r

$$(10.64) \quad \gamma_\infty = -\ln q^2,$$

where r should be determined such that the second term is a fraction (.1 say) of the first.

If k is small we can use [10, 17.3.21]

$$(10.65) \quad q = \frac{k^2}{16} + 8\left(\frac{k^2}{16}\right)^2 + 84\left(\frac{k^2}{16}\right)^3 + 992\left(\frac{k^2}{16}\right)^4 + \dots,$$

therefore we have from (10.63)

$$(10.66) \quad \gamma = -4\ln k + \left(8 - \frac{2}{r}\right) \ln 2 - \frac{2}{r} \ln \frac{1+q_q^2+\dots}{1+2q_q+\dots} - 2\ln\left(1+8\frac{k^2}{16}+84\left(\frac{k^2}{16}\right)^2+\dots\right).$$

If k' is small, we can use the relation for the complementary nome

$$(10.67) \quad q' = e^{-\pi K/K'},$$

namely (10.65) with q and k replaced by q' and k' . We have from (10.63)

$$(10.68) \quad \gamma = \frac{\pi^2}{\ln(\frac{4}{k'}) - \frac{k'^2}{4} - \frac{13}{128} k'^4 \dots} - \frac{2}{r} \ln 2 - \frac{2}{r} \ln \frac{1+q_q^2+\dots}{1+2q_q+\dots},$$

because

$$\ln(1 + 8(\frac{k'^2}{16}) + 84(\frac{k'^2}{16})^2 + \dots) = .8 \frac{k'^2}{16} + 52(\frac{k'^2}{16})^2 + \dots$$

Compare with [5, 6-57].

10.12. Optimization in the general case.

In the general case we have to replace (10.18) by

$$(10.69) \quad \min_{w_{h_j}, w_{v_j}} \max_{\lambda_i, \mu_i} \prod_{j=1}^r \left| \frac{\mu_i - w_{v_j}}{\mu_i + w_{h_j}} \cdot \frac{\lambda_i - w_{h_j}}{\lambda_i + w_{v_j}} \right|,$$

where $0 < a = \min \{\mu_i\} (<) b = \max \{\mu_i\}$,

$0 < c = \min \{\lambda_i\} (<) d = \max \{\lambda_i\}$.

The μ_i and λ_i are related in that they are the eigenvalues of H_1 and V_1 corresponding to the same eigenvector. Let us first make a linear fractional transformation (of the form $z = f(x) = \frac{\alpha x + \beta}{\gamma x + \delta}$) such that we have the correspondence

$$(10.70) \quad \begin{cases} x & -d & -c & a & b \\ z & -1 & -k' & k' & 1. \end{cases}$$

In implicit form the transformation is

$$(10.71) \quad \frac{1+k'}{2k'} \cdot \frac{z-k'}{z-1} = \frac{x-a}{x-b} \cdot \frac{c+b}{c+a},$$

which ensures the last three correspondences, the first one is ensured by an appropriate choice of k' , namely such that

$$(10.72) \quad \frac{1+k'}{2k'} \cdot \frac{1+k'}{2} = \frac{a+d}{b+d} \cdot \frac{b+c}{a+c} = \tau.$$

The product of the roots for k' is 1, and the sum is

$$2(2\tau-1) > 2,$$

because $\tau > 1$, as follows from

$$b(d-c) > a(d-c);$$

therefore the roots are real and positive, and one is less than 1 (the root we shall choose in the following), the other is larger than 1. Let

$$\begin{cases} w'_{h_j} = -f(-w_{h_j}), & w'_{v_j} = f(w_{v_j}), \\ \lambda'_i = -f(-\lambda_i), & \mu'_i = f(\mu_i). \end{cases}$$

(10.69) can be written in the form

$$\min_{w'_{h_j}, w'_{v_j}} \max_{\lambda'_i, \mu'_i} \prod_{j=1}^r \left| \frac{\mu'_i - w'_{v_j}}{\mu'_i + w'_{h_j}} \cdot \frac{\lambda'_i - w'_{h_j}}{\lambda'_i + w'_{v_j}} \right|.$$

This may be checked directly by using

$$\frac{z-w'}{z-w'_1} = \frac{x-w}{x-w_1} \cdot \frac{\gamma w_1 + \delta}{\gamma w + \delta},$$

but it also follows from an important property of linear fractional transformations, namely the preservation of the anharmonic ratio. As in section 10.6 we replace (10.69) by

$$(10.73) \quad \min_{w'_{h_j}, w'_{v_j}} \max_{\substack{x \in [k', 1] \\ y \in [k', 1]}} \prod_{j=1}^r \left| \frac{x-w'_{v_j}}{x+w'_{h_j}} \cdot \frac{y-w'_{h_j}}{y+w'_{v_j}} \right|.$$

We have to find w'_{v_j}, w'_{h_j} such that

$$(10.74) \quad \min_{w'_{h_j}, w'_{v_j}} \left(\max_{x \in [k', 1]} \left| \frac{P_1(x)}{P_2(-x)} \right| \cdot \max_{x \in [k', 1]} \left| \frac{P_2(x)}{P_1(-x)} \right| \right),$$

where P_1 and P_2 have the coefficient of x^r in common, and P_1 and P_2 have all the roots in $[k', 1]$. The results of section 10.6 extend directly and in a straightforward fashion to the above case, if we rewrite (10.74) in the form

$$(10.75) \quad \min_Q \max_{\substack{x \in [-1, -k'] \\ x \in [k', 1]}} |Q(x)|,$$

where, when x is positive, Q is a rational function which is a ratio of two polynomials of degree r with the same coefficient of x^r , and where $Q(x) = 1/Q(-x)$ for negative x .

We leave this as an exercise. Uniqueness implies that the roots of P_1 are the same as those of P_2 , therefore $w'_{h_j} = w'_{v_j}$. Using (10.71) and (10.72) the problem is reduced to the simple cyclic case treated in section 10.8.

10.13. Other related min-max problems

Wachspress has indicated how the solution of the problem (10.20) can be derived from the solution of the problem solved by Cauey:

$$(10.76) \quad \min_{a_j} \max_{0 \leq z \leq h < 1} \left| \prod_{j=1}^r \frac{a_j^2 - z^2}{1 - a_j^2 z^2} \right|.$$

Because this relationship is important in that it shows the way to the solution of a class of min-max problems, we shall give in detail a derivation of the solution of (10.76), using the known solution of (10.20).

The 1-2 transformation

$$x = f(z) = \frac{1-z^2}{1+z^2} \quad \text{or} \quad z^2 = \frac{1-x}{1+x}$$

transforms the interval $[k', 1]$ into the two intervals $[0, h]$ and $[-h, 0]$, where $h = \frac{1-k'}{1+k'}$. If $w_j = f(a_j)$, then

$$\frac{x-w_j}{x+w_j} = \frac{a_j^2 - z^2}{1 - a_j^2 z^2}.$$

The solution of (10.76) is therefore

$$a_j^2 = \frac{1 - \operatorname{dn}(u_j, k)}{1 + \operatorname{dn}(u_j, k)}, \quad u_j = \frac{2j-1}{2r} K.$$

Rewriting the expression for a_j in terms of Jacobi functions with argument $u_j/2$ we obtain [10, 16.18.5]

$$a_j = k \frac{\text{sn} \cdot \text{cn}}{\text{dn}} \left(\frac{u_j}{2}, k \right)$$

and by means of the Landen transformation [10, 16.14.2] we further have

$$a_j = \frac{k}{1+k'} \text{sn} \left(\frac{1+k'}{2} u_j, k_1 \right) = \sqrt{k_1} \text{sn} \left(\frac{1+k'}{2} u_j, k_1 \right),$$

where k_1 follows from k according to (10.29) and (10.30). But as [10, 17.3.2, 17.5.2 and 17.5.3]:

$$\frac{1+k'}{2} K = K_1,$$

there holds

$$a_j = \sqrt{k_1} \text{sn} \left(\frac{2j-1}{2r} K_1, k_1 \right).$$

10.14. Remark

After this exposition was completed, professor Richard Askey was kind enough to bring to my attention a related exposition by Todd [12], [13], which points to another related application by Solotareff [14, p. 280] and gives [12] the relation between the mesh size and the rate of convergence.

This exposition is best understood in terms of the very first paper of Jacobi on elliptic functions [9, p. 39] and the details given at the beginning of his Fundamenta Nova [9, p. 49 f]. See also [9, p. 415].

The theorem 10.3 of Achieser characterizes the optimal rational function Q by the properties

$$\left\{ \begin{array}{l} Q(v_j) = \frac{P(v_j)}{P(-v_j)} = (-1)^j \kappa_0, DQ(v_j) = 0, j = 1, \dots, r-1 \\ Q(1) = (-1)^r Q(k') = \kappa_0 < 1. \end{array} \right.$$

Moreover

$$Q(-x) = 1/Q(x).$$

Therefore, replacing κ_0 by κ ,

$$T(x) = (-P(x) + \kappa P(-x))(P(x) + \kappa P(-x))(-\kappa P(x) + P(-x))(\kappa P(x) + P(-x))$$

has the roots ± 1 , $\pm k'$ and the double roots $\pm v_j$, or with

$$R(x) = \prod_{j=1}^{r-1} (x^2 - v_j^2)$$

we have

$$(10.77) \quad T(x) = N(1-x^2)(x^2-k'^2) R^2(x),$$

where N is a constant, because the degree of T is $4r$ and that of R is $2(r-1)$.

If $\bar{P}(x) = P(-x)$ and D denotes the derivative operator, we have Jacobi's identity [16, p. 40]

$$(P - \alpha \bar{P}) D\bar{P} - \bar{P} D(P - \alpha \bar{P}) = P D\bar{P} - \bar{P} D P.$$

It follows that any double root of $P - \alpha \bar{P}$ is also a root of

$$S = \bar{P} D P - P D \bar{P}.$$

Therefore R divides S , but S has degree $2r - 2$; hence we have the identity

$$(10.78) \quad T(x) = M^2(1-x^2)(x^2-k'^2) S^2(x).$$

The sign of the constant M^2 follows from $T(x) < 0$ when $x > 1$.

If we divide by $(P(-x))^4$, take the square root and integrate, we obtain [12, p. 59]

$$(10.79) \quad \int_{y_0}^{\kappa} \frac{DQ}{\sqrt{(\kappa^2 - Q^2)(1 - \kappa^2 Q^2)}} = M \int_{x_0}^1 \frac{dx}{\sqrt{(1-x^2)(x^2-k'^2)}} = Mu$$

with $y_0 = P(x_0)/P(-x_0)$.

If we invert both integrals of (10.79) we obtain a Jacobi type formula [9, p.51]

$$(10.80) \quad \kappa \operatorname{cd}(\operatorname{Mu}, \kappa^2) = \frac{P(\operatorname{dn}(u, \kappa))}{P(-\operatorname{dn}(u, \kappa))},$$

which gives at once the roots

$$(10.81) \quad d_j = \frac{K(\kappa^2)(2j-1)}{M}, \quad j = 1, \dots, r,$$

and the relation between the periods, also given in section 10.8

$$(10.82) \quad \frac{K(k)}{r} = \frac{K(\kappa^2)}{M}.$$

The identity (10.77) can be generalized for other min-max problems and furnishes an easy way to obtain an optimal solution when r is small.

Literature

- [1] Peaceman, D.W.,
Rachford, H.H., Jr. The numerical solution of parabolic and
 elliptic differential equations.
J. Soc. Industr. Appl. Math., 3, 28-41, 1955.
- [2] Douglas, Jim Jr.,
Rachford, H.H., Jr. On the numerical solution of heat conduction
 problems in two or three space variables.
Trans. Amer. Math. Soc., 82, 421-439, 1956.
- [3] Douglas, Jim Jr.,
Kellog, R.B.,
Varga, R.S. Alternate-direction iteration methods for
 n space variables.
Math. of Comp., 17, 1963.
- [4] Varga, R.S. Matrix iterative analysis.
Prentice-Hall, Inc., Englewood Cliffs,
New Jersey, 1962.
- [5] Wachspress, E.L. Iterative solution of elliptic systems.
Prentice-Hall, Inc., Englewood Cliffs,
New Jersey, 1966.
- [6] Achieser, N.I. Theory of approximation (translation).
Frederick Ungar Publ. Co., New York, 1956.
- [7] Gauss, C.F. Anziehung eines elliptischen Ringes.
Ostwalt's Klassiker der exakten Wissen-
schaften, 255, 1927.
- [8] Lemaître, G. Cours de mécanique.
Louvain Nauwelaerts, 1956.
- [9] Jacobi, C.G.J. Gesammelte Werke I.
Reimer, Berlin, 1881.
- [10] -- Handbook of mathematical functions.
Applied Math. Series 55, Washington National
Bureau of Standards, 1965.
- [11] Knopp, K. Theory of functions II.
Dover, New York, 1947.

- [12] Todd, J. Optimal ADI parameters, Functional analysis, Approximation theory, Numerische Mathematik. Tagungen in Obervolfad, 22 bis 25 Juni, 1965.
- [13] Gaier, D., On the rate of convergence of optimal ADI processes.
 Todd, J. Numer. Math., 9, 452-459, 1966.
- [14] Achieser, N.I. Vorlesungen über Approximationstheorie. Akad. Verlag, Berlin, 1953.
- [15] Whittaker, E.T., A course of modern analysis, chs. 20 to 22.
 Watson, G.N. Cambridge University Press, 4th edition, Cambridge, England, 1952.

11. I.N. VEKUA'S THEORIE DER ELLIPTISCHE DIFFERENTIAALVERGELIJKINGEN MET ANALYTISCHE COËFFICIËNTEN

11.1. Inleiding

De Russische wiskundige I.N. Vekua heeft een nieuwe methode ontwikkeld om de oplossingen van elliptische p.d.v. met twee onafhankelijke variabelen en analytische coëfficiënten te representeren, gebaseerd op de Riemannfunctie en een Volterra-integraalvergelijking met complexe variabelen [1]. We beschouwen de lineaire elliptische p.d.v. van de tweede orde met twee onafhankelijke variabelen voor de onbekende functie $u(x,y)$ in zijn normaalvorm:

$$(E) \quad E(u) = \Delta u + a(x,y) \frac{\partial u}{\partial x} + b(x,y) \frac{\partial u}{\partial y} + c(x,y) u = f(x,y).$$

Als $f(x,y) \equiv 0$ noemen we vergelijking (E) homogeen en noteren hem als (E_0) . We nemen aan dat de coëfficiënten a , b , c en het rechterlid f van (E) complexwaardige analytische functies zijn van de twee reële variabelen x , y tegelijk in een enkelvoudig samenhangend gebied T van het $z = x + iy$ -vlak, d.w.z. dat ze ontwikkeld kunnen worden in een dubbelreeks

$$\sum_{m,n=0}^{\infty} \alpha_{mn} (x-x_0)^m (y-y_0)^n,$$

die absoluut convergent is in een omgeving van het willekeurig punt $(x_0, y_0) \in T$. We noemen T het definitiegebied van de vergelijking (E). Volgens de theorie der functies van meerdere complexe variabelen kunnen de bekende functies a , b , c en f dan op éénduidige wijze analytisch voortgezet worden in het gebied van de complexe variabelen $x = x_1 + ix_2$, $y = y_1 + iy_2$, zodanig dat de voortgezette functies samenvallen met de oorspronkelijke functies als $x_2 = y_2 = 0$.

We krijgen dan functies $a(x,y)$, $b(x,y)$, $c(x,y)$ en $f(x,y)$ van twee complexe variabelen, die analytisch zijn in een zeker gebied (D_1, D'_1) , met $x \in D_1$, $y \in D'_1$, van de ruimte van twee complexe variabelen. Er geldt: $T \subset (D_1, D'_1)$. We voeren nu twee nieuwe onafhankelijke variabelen in

$$(11.1) \quad z = x + iy \qquad \zeta = x - iy$$

(merk op: $\zeta = \bar{z} \Leftrightarrow x$ en y reëel).

Door deze transformatie gaat een functie $f(x,y)$ over in een daarmee samenhangende functie

$$F(z, \zeta) = f(x, y) = f\left(\frac{z+\zeta}{2}, \frac{z-\zeta}{2i}\right)$$

(we hebben f als representant voor de functies a , b , c en f gekozen).

Aangezien z en ζ analytisch zijn en $\begin{vmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \\ \frac{\partial \zeta}{\partial x} & \frac{\partial \zeta}{\partial y} \end{vmatrix} \neq 0$, zal $F(z, \zeta)$ met $f(x, y)$

analytisch zijn en wel in een zeker gebied (D, D') , $z \in D$, $\zeta \in D'$; D , D' enkelvoudig samenhangende gebieden in resp. het complexe z -vlak en het complexe ζ -vlak.

Als $\zeta = \bar{z}$ dan is $D' = \bar{D}$, de gespiegelde van D t.o.v. $\text{Im } z = 0$ in het z -vlak. Er geldt $D \supset T$.

Voorbeeld 11.1.

Zij T een enkelvoudig samenhangend gebied en $u(x, y)$ een reële harmonische functie in T . Er bestaat dan een functie $f(z)$ van een complexe variabele analytisch in T , zodanig dat

$$u(x, y) = \text{Re } f(z).$$

We definiëren

$$\bar{f}(z) = \overline{f(\bar{z})}.$$

Het is duidelijk dat $\bar{f}(z)$ analytisch is voor $z \in \bar{T}$. Beschouw nu de functie

$$U(z, \zeta) = \frac{1}{2} [f(z) + \bar{f}(\zeta)].$$

Deze functie is analytisch voor $(z, \zeta) \in (T, \bar{T})$ en er geldt, voor reële x en y :

$$U(z, \bar{z}) = \frac{1}{2} [f(z) + \overline{f(z)}] = u(x, y).$$

Uit deze constructie volgt dat $U(z, \zeta)$ de analytische voortzetting is van $u(x, y)$ in het gebied $(T, \bar{T})$.

Uit (11.1) volgt

$$(11.2) \quad x = \frac{1}{2} (z + \zeta) \quad y = \frac{1}{2i} (z - \zeta).$$

Beschouw nu de differentiaaloperatoren

$$(11.3) \quad \frac{\partial}{\partial z} = \frac{1}{2} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) \text{ en } \frac{\partial}{\partial \zeta} = \frac{1}{2} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right).$$

Deze kunnen toegepast worden op iedere differentieerbare functie van twee complexe variabelen x en y . We moeten echter bedenken dat, als x en y reëel zijn, $\frac{\partial U}{\partial z}$ en $\frac{\partial U}{\partial \zeta}$ in het algemeen niet beschouwd kunnen worden als de partiële afgeleiden van U met betrekking tot z en ζ . Dat kan alleen als U een analytische functie is van z en ζ .

Door middel van de operatoren (11.3) gaat de Laplaciaan van de analytische functie $u(x, y)$ over in:

$$(11.4) \quad \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 4 \frac{\partial^2 U}{\partial z \partial \zeta}.$$

Met behulp van (11.2), (11.3) en (11.4) gaat de vergelijking (E) over in

$$(F) \quad F(U) = \frac{\partial^2 U}{\partial z \partial \zeta} + A(z, \zeta) \frac{\partial U}{\partial z} + B(z, \zeta) \frac{\partial U}{\partial \zeta} + C(z, \zeta) U = F(z, \zeta)$$

met

$$(11.5) \quad \left\{ \begin{array}{l} A(z, \zeta) = \frac{1}{4} \left[a\left(\frac{z+\zeta}{2}, \frac{z-\zeta}{2i}\right) + ib\left(\frac{z+\zeta}{2}, \frac{z-\zeta}{2i}\right) \right], \\ B(z, \zeta) = \frac{1}{4} \left[a\left(\frac{z+\zeta}{2}, \frac{z-\zeta}{2i}\right) - ib\left(\frac{z+\zeta}{2}, \frac{z-\zeta}{2i}\right) \right], \\ C(z, \zeta) = \frac{1}{4} c\left(\frac{z+\zeta}{2}, \frac{z-\zeta}{2i}\right), \\ F(z, \zeta) = \frac{1}{4} f\left(\frac{z+\zeta}{2}, \frac{z-\zeta}{2i}\right), \\ U(z, \zeta) = u\left(\frac{z+\zeta}{2}, \frac{z-\zeta}{2i}\right), \end{array} \right.$$

waarbij A , B , C en F analytisch zijn in z , ζ in een gebied (D, D') .

Definitie 11.1.

Een enkelvoudig samenhangend deelgebied D van het definitiegebied T , waarvoor geldt dat $A(z, \zeta)$, $B(z, \zeta)$ en $C(z, \zeta)$ analytische functies zijn van de twee complexe variabelen z, ζ in het gebied (D, D') , $z \in D$, $\zeta \in D'$, noemen we een fundamenteaalgebied van de vergelijking (E).

Er bestaat altijd een fundamenteaalgebied voor de vergelijking (E) met analytische coëfficiënten, nl. iedere voldoende kleine ρ -omgeving

$$|x - x_0| < \rho, \quad |y - y_0| < \rho$$

van een willekeurig punt (x_0, y_0) uit het definitiegebied T . Als, in het bijzonder, de coëfficiënten a, b, c van (E) gehele functies zijn van x en y dan zijn ook A, B, C gehele functies van z en ζ . In dit geval kan het gehele complexe z -vlak als fundamenteaalgebied genomen worden.

Als D een fundamenteaalgebied is voor (E), dan is ieder enkelvoudig samenhangend deelgebied van D ook fundamenteaalgebied voor (E).

Als $F(z, \zeta) \equiv 0$ noteren we de vergelijking (F) als (F_0) .

Naast (E) zullen we ook de vergelijking

$$(E^*) \quad E^*(v) = \Delta v - \frac{\partial a v}{\partial x} - \frac{\partial b v}{\partial y} + c v = f^*(x, y)$$

beschouwen, waarin f^* een analytische functie is.

We noemen (E^*) de geadjungeerde vergelijking van (E). Het is duidelijk dat de geadjungeerde vergelijking van (E^*) weer E is en dat E en E^* een zelfde definitiegebied T en een zelfde fundamenteaalgebied D hebben.

Door middel van (11.2), (11.3) en (11.4) gaat $(E)^*$ over in

$$(F^*) \quad F^*(V) = \frac{\partial^2 V}{\partial z \partial \zeta} - \frac{\partial A V}{\partial z} - \frac{\partial B V}{\partial \zeta} + C V = F^*(z, \zeta),$$

hetgeen juist de geadjungeerde vergelijking van (F) is.

Als $F^*(z, \zeta) \equiv 0$ noteren we (F^*) als (F_0^*) .

Als $f^*(x, y) \equiv 0$ noteren we (E^*) als (E_0^*) .

Onder $E(u)$ zullen we steeds verstaan de operator E toegepast op de functie u met betrekking tot de variabelen x en y :

$$E(u) = \Delta u + a(x,y) \frac{\partial u}{\partial x} + b(x,y) \frac{\partial u}{\partial y} + c(x,y) u.$$

$F(U)$ is dezelfde operatie maar nu uitgedrukt in de variabelen z en ζ . In het bovenstaande hebben we de twee reële variabelen x, y voorkomend in de vergelijking (E) ingebed in de 4-dimensionale ruimte van de complexe variabelen z, ζ . Door de substitutie (11.1) is de elliptische p.v.d. (E) overgevoerd in de vergelijking (F). Deze heeft formeel de vorm van een hyperbolische p.d.v. We zijn nu in staat de theorie der hyperbolische p.d.v. toe te passen op de elliptische p.d.v. (E) met analytische coëfficiënten en twee onafhankelijke variabelen. Aldus zullen we komen tot een expliciete integraalvoorstelling van alle reguliere oplossingen van (E). We noemen $u(x,y)$ een reguliere oplossing als hij continue partiële afgeleiden van de eerste en tweede orde bezit in T en voldoet aan (E). Uit deze expliciete integraalvoorstelling zijn diverse eigenschappen van de oplossingen van (E) af te leiden, zoals analyticiteit der oplossingen (stelling van Picard) en existentie en eenduidigheid van randwaardeproblemen. Vekua [1] heeft ook representaties gegeven voor oplossingen gedefinieerd in meervoudig samenhangende gebieden en voor oplossingen van vergelijkingen van hogere orde en systemen van elliptische vergelijkingen; hierop zullen we echter niet ingaan.

11.2. De Riemannfunctie van de vergelijking (E_0)

Beschouw de homogene vergelijking

$$(E_0) \quad E(u) = \Delta u + a(x,y) \frac{\partial u}{\partial x} + b(x,y) \frac{\partial u}{\partial y} + c(x,y) u = 0.$$

We kunnen deze vergelijking ook als volgt schrijven

$$(F_0) \quad F(U) = \frac{\partial^2 U}{\partial z \partial \bar{\zeta}} + A(z, \bar{\zeta}) \frac{\partial U}{\partial z} + B(z, \bar{\zeta}) \frac{\partial U}{\partial \bar{\zeta}} + C(z, \bar{\zeta}) U = 0.$$

D zij een fundamenteaalgebied van (E_0) (en dus ook van (E_0^*)).

De bijbehorende homogene geadjungeerde vergelijking is

$$(F_0^*) \quad F^*(V) = \frac{\partial^2 V}{\partial z \partial \bar{\zeta}} - \frac{\partial AV}{\partial z} - \frac{\partial BV}{\partial \bar{\zeta}} + CV = 0.$$

De operator F en de geadjungeerde operator F^* bezitten de volgende essentiële eigenschap:

$$(11.6) \quad VF(U) - UF^*(V) = \frac{\partial}{\partial z} \left(V \frac{\partial U}{\partial \zeta} + AUV \right) + \frac{\partial}{\partial \zeta} \left(-U \frac{\partial V}{\partial z} + BUV \right)$$

voor twee willekeurige analytische functies $U(z, \zeta)$ en $V(z, \zeta)$ in het gebied (D, D') .

We zullen nu de functie $G(z, \zeta, t, \tau)$ gaan beschouwen die voldoet aan de geadjungeerde homogene vergelijking

$$(11.7) \quad F_{z, \zeta}^* [G(z, \zeta, t, \tau)] = 0$$

en de karakteristieke randvoorwaarden

$$(11.8) \quad \left\{ \begin{array}{l} \frac{\partial G(z, \zeta, t, \tau)}{\partial \zeta} = A(z, \zeta) G(z, \zeta, t, \tau) \text{ voor } z = t \in D; \zeta, \tau \in D', \\ \frac{\partial G(z, \zeta, t, \tau)}{\partial z} = B(z, \zeta) G(z, \zeta, t, \tau) \text{ voor } \zeta = \tau \in D'; z, t \in D, \\ G(t, \tau, t, \tau) = 1, \end{array} \right.$$

waarbij t en τ vaste punten zijn in resp. D en D' .

Uit (11.8) volgt

$$(11.9) \quad \left\{ \begin{array}{l} G(z, \zeta, t, \tau)|_{z=t} = \exp \int_{\tau}^{\zeta} A(t, \eta) d\eta, \\ G(z, \zeta, t, \tau)|_{\zeta=\tau} = \exp \int_t^z B(\xi, \tau) d\xi. \end{array} \right.$$

De functie $G(z, \zeta, t, \tau)$ noemen we de Riemann functie van de vergelijking (E_0) . Hij hangt uitsluitend af van de differentiaalvergelijking en niet van het beschouwde gebied. Een en ander is volkomen analoog aan de invoering van de Riemannfunctie voor een hyperbolische differentiaalvergelijking met reële variabelen.

We zullen nu een bewijsschets geven van de existentie van de functie $G(z, \zeta, t, \tau)$ evenals van de ondubbelzinnigheid en analyticiteit in het gebied (D, D') .

Uit (11.7) volgt

$$\begin{aligned} & \int_t^z d\xi \int_\tau^\zeta \frac{\partial^2 G(\xi, \eta, t, \tau)}{\partial \xi \partial \eta} d\eta - \int_t^z d\xi \int_\tau^\zeta \frac{\partial A(\xi, \eta) G(\xi, \eta, t, \tau)}{\partial \xi} d\eta + \\ & + \int_t^z d\xi \int_\tau^\zeta \frac{\partial B(\xi, \eta) G(\xi, \eta, t, \tau)}{\partial \eta} d\eta + \int_t^z d\xi \int_\tau^\zeta C(\xi, \eta) G(\xi, \eta, t, \tau) d\eta = 0. \end{aligned}$$

Integratie levert

$$\begin{aligned} (11.10) \quad & G(z, \zeta, t, \tau) - G(t, \zeta, t, \tau) - G(z, \tau, t, \tau) + G(t, \tau, t, \tau) + \\ & - \int_\tau^\zeta A(z, \eta) G(z, \eta, t, \tau) d\eta + \int_\tau^\zeta A(t, \eta) G(t, \eta, t, \tau) d\eta + \\ & - \int_t^z B(\xi, \zeta) G(\xi, \zeta, t, \tau) d\xi + \int_t^z B(\xi, \tau) G(\xi, \tau, t, \tau) d\xi + \\ & + \int_t^z d\xi \int_\tau^\zeta C(\xi, \eta) G(\xi, \eta, t, \tau) d\eta = 0. \end{aligned}$$

Uit (11.8) volgt

$$\int_\tau^\zeta \frac{\partial G(t, \eta, t, \tau)}{\partial \eta} d\eta = \int_\tau^\zeta A(t, \eta) G(t, \eta, t, \tau) d\eta,$$

zodat

$$G(t, \zeta, t, \tau) - \int_\tau^\zeta A(t, \eta) G(t, \eta, t, \tau) d\eta = G(t, \tau, t, \tau)$$

en analoog

$$G(z, \tau, t, \tau) - \int_t^z B(\xi, \tau) G(\xi, \tau, t, \tau) d\xi = G(t, \tau, t, \tau).$$

Wanneer we deze beide laatste betrekkingen invullen in (11.10) en gebruik maken van het feit dat $G(t, \tau, t, \tau) = 1$ krijgen we

$$(11.11) \quad G(z, \zeta, t, \tau) - \int_{\tau}^{\zeta} A(z, \eta) G(z, \eta, t, \tau) d\eta + \\ - \int_t^z B(\xi, \zeta) G(\xi, \zeta, t, \tau) d\xi + \int_t^z d\xi \int_{\tau}^{\zeta} C(\xi, \eta) G(\xi, \eta, t, \tau) d\eta = 1.$$

(11.11) is een integraalvergelijking van het Volterra type voor de functie $G(z, \zeta, t, \tau)$. Een dergelijke integraalvergelijking heeft altijd een ondubbelzinnig bepaalde oplossing $G(z, \zeta, t, \tau)$. $G(z, \zeta, t, \tau)$ is een analytische functie van de vier variabelen z, ζ, t, τ in het gebied $z, t \in D, \zeta, \tau \in D'$. Voor het bewijs van de existentie, eenduidigheid en analyticiteit zij verwezen naar [1], pag. 11-16.

In bepaalde gevallen kan $G(z, \zeta, t, \tau)$ expliciet geconstrueerd worden met behulp van de methode van de successieve approximatie [1], pag. 20.

De voorwaarden (11.9) zijn equivalent met

$$(11.12) \quad \begin{cases} \frac{\partial G(t, \zeta, t, \tau)}{\partial \tau} + A(t, \tau) G(t, \zeta, t, \tau) = 0 \text{ voor } t = z \\ \frac{\partial G(z, \tau, t, \tau)}{\partial t} + B(t, \tau) G(z, \tau, t, \tau) = 0 \text{ voor } \tau = \zeta \\ G(z, \zeta, z, \zeta) = 1. \end{cases}$$

Wanneer we nu in (11.6) voor $V(z, \zeta)$ de Riemannfunctie $G(z, \zeta, t, \tau)$ invullen en gebruik maken van het feit dat $F_{z, \zeta}^*(G) = 0$, gaat deze relatie over in

$$(11.13) \quad GF(U) = \frac{\partial}{\partial z} [G(\frac{\partial U}{\partial \zeta} + AU)] + \frac{\partial}{\partial \zeta} [U(BG - \frac{\partial G}{\partial z})]$$

geldig voor elke functie $U(z, \zeta)$ analytisch in het gebied (D, D') .

We zullen nu een andere uitdrukking afleiden voor de eerste term in het rechterlid van (11.13):

$$\frac{\partial^2 UG}{\partial z \partial \zeta} = \frac{\partial}{\partial z} (G \frac{\partial U}{\partial \zeta} + U \frac{\partial G}{\partial \zeta}) = \frac{\partial}{\partial z} [G(\frac{\partial U}{\partial \zeta} + AU)] + \frac{\partial}{\partial z} [U(\frac{\partial G}{\partial \zeta} - AG)].$$

Dit ingevuld in (11.13) levert

$$(11.14) \quad \frac{\partial^2 UG}{\partial z \partial \zeta} = \frac{\partial}{\partial z} [U(\frac{\partial G}{\partial \zeta} - AG)] + \frac{\partial}{\partial \zeta} [U(\frac{\partial G}{\partial z} - BG)] + GF(U).$$

Verwissel in (11.14) de paren (z, ζ) en (t, τ) en integreer beide zijden van de nieuwe betrekking naar t in het interval (z_0, z) en naar τ in het interval (ζ_0, ζ) waarbij z_0, ζ_0 vaste punten zijn ($z_0 \in D, \zeta_0 \in D'$).

$$\begin{aligned} & \int_{z_0}^z dt \int_{\zeta_0}^{\zeta} \frac{\partial^2 U(t, \tau) G(t, \tau, z, \zeta)}{\partial t \partial \tau} d\tau = \\ & \int_{z_0}^z dt \int_{\zeta_0}^{\zeta} \frac{\partial}{\partial t} [U(t, \tau) (\frac{\partial G(t, \tau, z, \zeta)}{\partial \tau} - A(t, \tau) G(t, \tau, z, \zeta))] d\tau + \\ & + \int_{z_0}^z dt \int_{\zeta_0}^{\zeta} \frac{\partial}{\partial \tau} [U(t, \tau) (\frac{\partial G(t, \tau, z, \zeta)}{\partial t} - B(t, \tau) G(t, \tau, z, \zeta))] d\tau + \\ & + \int_{z_0}^z dt \int_{\zeta_0}^{\zeta} G(t, \tau, z, \zeta) F[U(t, \tau)] d\tau. \end{aligned}$$

Integratie levert

$$\begin{aligned} & U(z, \zeta) G(z, \zeta, z, \zeta) + U(z_0, \zeta_0) G(z_0, \zeta_0, z, \zeta) + \\ & - U(z, \zeta_0) G(z, \zeta_0, z, \zeta) - U(z_0, \zeta) G(z_0, \zeta, z, \zeta) = \\ & \int_{\zeta_0}^{\zeta} [U(z, \tau) (\frac{\partial G(z, \tau, z, \zeta)}{\partial \tau} - A(z, \tau) G(z, \tau, z, \zeta))] + \\ & - \{U(z_0, \tau) (\frac{\partial G(z_0, \tau, z, \zeta)}{\partial \tau} - A(z_0, \tau) G(z_0, \tau, z, \zeta))\} d\tau + \\ & + \int_{z_0}^z [U(t, \zeta) (\frac{\partial G(t, \zeta, z, \zeta)}{\partial t} - B(t, \zeta) G(t, \zeta, z, \zeta))] + \\ & - \{U(t, \zeta_0) (\frac{\partial G(t, \zeta_0, z, \zeta)}{\partial t} - B(t, \zeta_0) G(t, \zeta_0, z, \zeta))\} dt + \\ & + \int_{z_0}^z dt \int_{\zeta_0}^{\zeta} GF(U) d\tau. \end{aligned}$$

Maak nu gebruik van (11.8)

$$\left\{ \begin{array}{l} \frac{\partial G(t, \tau, z, \zeta)}{\partial \tau} - A(t, \tau) G(t, \tau, z, \zeta) = 0 \text{ als } t = z \\ \frac{\partial G(t, \tau, z, \zeta)}{\partial t} - B(t, \tau) G(t, \tau, z, \zeta) = 0 \text{ als } \tau = \zeta \\ G(z, \zeta, z, \zeta) = 1 \end{array} \right.$$

en integreer partieël. Dit levert

$$\begin{aligned} (11.15) \quad U(z, \zeta) &= U(z_0, \zeta_0) G(z_0, \zeta_0, z, \zeta) + \\ &+ \int_{z_0}^z G(t, \zeta_0, z, \zeta) \left[\frac{\partial U(t, \zeta_0)}{\partial t} + B(t, \zeta_0) U(t, \zeta_0) \right] dt + \\ &+ \int_{\zeta_0}^{\zeta} G(z_0, \tau, z, \zeta) \left[\frac{\partial U(z_0, \tau)}{\partial \tau} + A(z_0, \tau) U(z_0, \tau) \right] d\tau + \\ &+ \int_{z_0}^z dt \int_{\zeta_0}^{\zeta} G(t, \tau, z, \zeta) F[U(t, \tau)] d\tau. \end{aligned}$$

Deze identiteit geldt evenals (11.6) voor elke functie $U(z, \zeta)$ analytisch in (D, D') . We kunnen diens gevolg voor $U(z, \zeta)$ in het bijzonder $G(z_0, \zeta_0, z, \zeta)$ kiezen. Ingevuld in (11.15) levert dit

$$\begin{aligned} G(z_0, \zeta_0, z, \zeta) &= G(z_0, \zeta_0, z_0, \zeta_0) G(z_0, \zeta_0, z, \zeta) + \\ &+ \int_{z_0}^z G(t, \zeta_0, z, \zeta) \left[\frac{\partial G(z_0, \zeta_0, t, \zeta_0)}{\partial t} + B(t, \zeta_0) G(z_0, \zeta_0, t, \zeta_0) \right] dt + \\ &+ \int_{\zeta_0}^{\zeta} G(z_0, \tau, z, \zeta) \left[\frac{\partial G(z_0, \zeta_0, z_0, \tau)}{\partial \tau} + A(z_0, \tau) G(z_0, \zeta_0, z_0, \tau) \right] d\tau + \\ &+ \int_{z_0}^z dt \int_{\zeta_0}^{\zeta} G(t, \tau, z, \zeta) F[G(z_0, \zeta_0, t, \tau)] d\tau. \end{aligned}$$

Met behulp van (11.12) volgt hieruit

$$(11.16) \quad \int_{z_0}^z dt \int_{\zeta_0}^{\zeta} G(t, \tau, z, \zeta) F[G(z_0, \zeta_0, t, \tau)] d\tau = 0.$$

Noem $F[G(z_0, \zeta_0, t, \tau)] = \alpha(t, \tau)$.

$\alpha(t, \tau)$ is een analytische functie van (t, τ) in (D, D') waarvoor de volgende Volterra integraalvergelijking van de eerste soort geldt

$$\int_{z_0}^z dt \int_{\zeta_0}^{\zeta} G(t, \tau, z, \zeta) \alpha(t, \tau) d\tau = 0.$$

Differentiëren naar z en ζ levert een integraalvergelijking van Volterra van de tweede soort:

$$\begin{aligned} \alpha(z, \zeta) + \int_{z_0}^z \frac{\partial G(t, \tau, z, \zeta)}{\partial z} \alpha(t, \tau) dt + \int_{\zeta_0}^{\zeta} \frac{\partial G(z, \tau, z, \zeta)}{\partial \zeta} \alpha(z, \tau) d\tau + \\ + \int_{z_0}^z dt \int_{\zeta_0}^{\zeta} \frac{\partial^2 G(t, \tau, z, \zeta)}{\partial z \partial \zeta} \alpha(t, \tau) d\tau = 0. \end{aligned}$$

Aangezien uit deze integraalvergelijking volgt dat de afgeleiden van elke orde van $\alpha(z, \zeta)$ gelijk nul zijn in het punt (z_0, ζ_0) , heeft hij slechts de nuloplossing.

Er geldt dus

$$F_{t, \tau} [G(z_0, \zeta_0, t, \tau)] = 0.$$

Dit leidt tot de volgende stelling:

Stelling 11.1.

$G(z, \zeta, t, \tau)$ voldoet als functie van zijn laatste twee argumenten t, τ aan de vergelijking (F_0) en aan de voorwaarden (11.12). Hieruit volgt, dat $G(z, \zeta, t, \tau)$ als functie van zijn laatste twee argumenten de Riemannfunctie is van de vergelijking E_0^* , die immers (E_0) als geadjungeerde vergelijking heeft.

Stelling 11.2.

De functie $U_0(z, \zeta)$ gedefinieerd door

$$(11.17) \quad U_0(z, \zeta) = \int_{z_0}^z dt \int_{\zeta_0}^{\zeta} G(t, \tau, z, \zeta) F(t, \tau) d\tau,$$

waarbij F een willekeurige analytische functie is van zijn argumenten in (D, D') , voldoet aan de vergelijking

$$(F) \quad F(U) = \frac{\partial^2 U}{\partial z \partial \zeta} + A(z, \zeta) \frac{\partial U}{\partial z} + B(z, \zeta) \frac{\partial U}{\partial \zeta} + C(z, \zeta) U = F(z, \zeta).$$

Dientengevolge is $U_0(z, \zeta)$ een particuliere oplossing van vergelijking (F).

Bewijs

Uit (11.17) volgt dat $U_0(z, \zeta)$ een analytische functie is van (z, ζ) in (D, D') , die voldoet aan

$$\begin{aligned} \text{a)} \quad & U_0(z_0, \zeta_0) = 0 \\ \text{b)} \quad & U_0(z_0, \zeta) = U_0(z, \zeta_0) = 0 \\ \text{c)} \quad & \frac{\partial U_0(z_0, \zeta)}{\partial \zeta} = \frac{\partial U_0(z, \zeta_0)}{\partial z} = 0. \end{aligned}$$

Ingevuld in de voor elke analytische functie geldende relatie (11.15) levert dit

$$U_0(z, \zeta) = \int_{z_0}^z dt \int_{\zeta_0}^{\zeta} G(t, \tau, z, \zeta) F[U_0(t, \tau)] d\tau.$$

Met behulp van (11.17) wordt dit

$$\int_{z_0}^z dt \int_{\zeta_0}^{\zeta} G(t, \tau, z, \zeta) \{F[U_0(t, \tau)] - F(t, \tau)\} d\tau = 0,$$

waaruit op analoge wijze als in het bewijs van stelling 11.1 volgt

$$F[U_0(z, \zeta)] = F(z, \zeta).$$

11.3. Analytische oplossingen van (E_0)

Stelling 11.3.

Iedere in (D, D') analytische oplossing $U(z, \zeta)$ van vergelijking (F_0) wordt gerepresenteerd door de formule

$$(11.18) \quad U(z, \zeta) = \alpha_0 G(z_0, \zeta_0, z, \zeta) + \int_{z_0}^z \phi(t) G(t, \zeta_0, z, \zeta) dt + \\ + \int_{\zeta_0}^{\zeta} \phi^*(\tau) G(z_0, \tau, z, \zeta) d\tau$$

met willekeurige α_0 en willekeurige in D resp. D' analytische functies $\phi(z)$ resp. $\phi^*(\zeta)$.

Bewijs

Laat $U(z, \zeta)$ een analytische functie zijn van z, ζ in (D, D') , die voldoet aan $F(U) = 0$. Dan levert (11.15)

$$U(z, \zeta) = U(z_0, \zeta_0) G(z_0, \zeta_0, z, \zeta) + \\ + \int_{z_0}^z G(t, \zeta_0, z, \zeta) \left[\frac{\partial U(t, \zeta_0)}{\partial t} + B(t, \zeta_0) U(t, \zeta_0) \right] dt + \\ + \int_{\zeta_0}^{\zeta} G(z_0, \tau, z, \zeta) \left[\frac{\partial U(z_0, \tau)}{\partial \tau} + A(z_0, \tau) U(z_0, \tau) \right] d\tau,$$

zodat $U(z, \zeta)$ in de vorm (11.18) gebracht is met

$$(11.19) \quad \begin{cases} \alpha_0 = U(z_0, \zeta_0) \\ \phi(z) = \frac{\partial U(z, \zeta_0)}{\partial z} + B(z, \zeta_0) U(z, \zeta_0) \\ \phi^*(\zeta) = \frac{\partial U(z_0, \zeta)}{\partial \zeta} + A(z_0, \zeta) U(z_0, \zeta). \end{cases}$$

Het is duidelijk dat $\phi(z)$ en $\phi^*(\zeta)$ analytische functies zijn in D resp. D' . Omgekeerd geldt voor willekeurige α_0 en willekeurige analytische functies $\phi(z)$ en $\phi^*(\zeta)$ in D resp. D' dat de functie $U(z, \zeta)$ gegeven door betrekking (11.18) analytisch is en voldoet aan (F_0) in het gebied (D, D') . Immers de operator F toegepast op (11.18) levert

$$\begin{aligned}
F(U) = & \alpha_0 F[G(z_0, \zeta_0, z, \zeta)] + \int_{z_0}^z \phi(t) F[G(t, \zeta_0, z, \zeta)] dt + \\
& + \int_{\zeta_0}^{\zeta} \phi^*(\tau) F[G(z_0, \tau, z, \zeta)] d\tau + \\
& + \phi(z) \left[\frac{\partial G(z, \zeta_0, z, \zeta)}{\partial \zeta} + A(z, \zeta) G(z, \zeta_0, z, \zeta) \right] + \\
& + \phi^*(\zeta) \left[\frac{\partial G(z_0, \zeta, z, \zeta)}{\partial z} + B(z, \zeta) G(z_0, \zeta, z, \zeta) \right].
\end{aligned}$$

Met behulp van (11.12) en stelling 11.1 volgt hieruit

$$F(U) = 0.$$

Indien we in (11.18) substitueren $z = x + iy$, $\zeta = x - iy$, waarbij $(x, y) \in D$ (hetgeen impliceert dat x en y reëel zijn, zodat $\zeta = \bar{z}$, $D' = \bar{D}$), krijgen we een functie $u(x, y)$, die analytisch is in D en voldoet aan de vergelijking

$$(E_0) \quad E(u) = \Delta u + a(x, y) \frac{\partial u}{\partial x} + b(x, y) \frac{\partial u}{\partial y} + c(x, y) u = 0.$$

We komen zo tot de volgende stelling:

Stelling 11.4.

D zij een fundamenteaalgebied van (E_0) . Dan representeert de formule

$$\begin{aligned}
(11.20) \quad u(x, y) = & \alpha_0 G(z_0, \zeta_0, z, \bar{z}) + \int_{z_0}^z \phi(t) G(t, \zeta_0, z, \bar{z}) dt + \\
& + \int_{\zeta_0}^{\bar{z}} \phi^*(\tau) G(z_0, \tau, z, \bar{z}) d\tau \quad (z=x+iy, \bar{z}=x-iy, (x, y) \in D),
\end{aligned}$$

waarbij α_0 een willekeurige constante is, $\phi(z)$ en $\phi^*(\bar{z})$ willekeurige analytische functies zijn in resp. D , $\bar{D}$, alle oplossingen van (E_0) analytisch in D .

Opmerking 11.1.

Aangezien z_0, ζ_0 willekeurige vaste punten zijn in resp. D, D' mogen we stellen $\zeta_0 = \bar{z}_0$.

We zullen in §11.7 aantonen dat formule (11.20) ook alle in D reguliere oplossingen van (E_0) representeert.

We zullen nu de representatie (11.18) in een iets andere vorm brengen, die voor het vervolg handiger is. We definiëren hiertoe:

$$(11.21) \quad \begin{cases} \phi_0(z) = \frac{1}{2} \alpha_0 + \int_{z_0}^z \phi(t) dt \\ \phi_0^*(\zeta) = \frac{1}{2} \alpha_0 + \int_{\zeta_0}^{\zeta} \phi^*(\tau) d\tau, \end{cases}$$

waarbij $\phi(z)$ en $\phi^*(\zeta)$ willekeurige in D resp. D' analytische functies zijn. Ook $\phi_0(z)$ en $\phi_0^*(\zeta)$ zijn dan analytisch in D resp. D' . Er geldt:

$$(11.22) \quad \phi_0(z_0) = \phi_0^*(\zeta_0) = \frac{1}{2} \alpha_0.$$

Uit (11.18) en (11.21) volgt

$$\begin{aligned} U(z, \zeta) = & \alpha_0 G(z_0, \zeta_0, z, \zeta) + \int_{z_0}^z G(t, \zeta_0, z, \zeta) \frac{d\phi_0(t)}{dt} dt + \\ & + \int_{\zeta_0}^{\zeta} G(z_0, \tau, z, \zeta) \frac{d\phi_0^*(\tau)}{d\tau} d\tau. \end{aligned}$$

Partiële integratie levert

$$\begin{aligned} (11.23) \quad U(z, \zeta) = & G(z, \zeta_0, z, \zeta) \phi_0(z) - \int_{z_0}^z \phi_0(t) \frac{\partial}{\partial t} G(t, \zeta_0, z, \zeta) dt + \\ & + G(z_0, \zeta, z, \zeta) \phi_0^*(\zeta) - \int_{\zeta_0}^{\zeta} \phi_0^*(\tau) \frac{\partial}{\partial \tau} G(z_0, \tau, z, \zeta) d\tau. \end{aligned}$$

Evenals (11.18) representeert (11.23) alle in (D, D') analytische oplossingen van (F_0) , indien we voor $\phi_0(z), \phi_0^*(\zeta)$ willekeurige in D resp. D' analytische functies nemen.

11.4. Fundamenteaaloplossingen

Formule (11.23) stelt ons in staat een aantal belangrijke particuliere oplossingen van de vergelijking (E_0) in handen te krijgen. In het bijzonder kunnen we uit (11.23) zgn. fundamenteaaloplossingen afleiden.

Definitie 11.2.

Zij D een fundamenteaalgebied van (E_0) . Zij $(x_0, y_0) \in D$. Iedere oplossing van (E_0) die zich in $D \setminus (x_0, y_0)$ gedraagt als

$$(11.24) \quad g(x, y) \log [(x-x_0)^2 + (y-y_0)^2] + g_0(x, y),$$

met g, g_0 in D reguliere functies en met $g(x_0, y_0) \neq 0$, noemen we fundamenteaaloplossing van (E_0) .

Het punt (x_0, y_0) noemen we de pool van de fundamenteaaloplossing. Een fundamenteaaloplossing heet genormaliseerd indien $g(x_0, y_0) = -\frac{1}{4\pi}$.

Uit de definitie volgt dat, indien we (11.24) vermenigvuldigen met een constante ongelijk nul en er een in D reguliere oplossing van (E_0) bij optellen, we opnieuw een fundamenteaaloplossing hebben. We kunnen dus extra eisen aan de fundamenteaaloplossing opleggen.

We zullen nu fundamenteaaloplossingen voor (E_0) afleiden.

Substitueer in (11.23)

$$(11.25) \quad \phi_0(z) = \log(z-z_0) \qquad \phi_0^*(\zeta) = \log(\zeta-\zeta_0),$$

waarbij we in D een coupure aanbrengen vanuit z_0 en in D' een coupure vanuit ζ_0 , opdat de functies $\phi_0(z)$ en $\phi_0^*(\zeta)$ analytisch zijn. Dit levert

$$\begin{aligned} U(z, \zeta) = & G(z, \zeta_0, z, \zeta) \log(z-z_0) - \int_{z_0}^z \log(t-z_0) \frac{\partial}{\partial t} G(t, \zeta_0, z, \zeta) dt + \\ & + G(z_0, \zeta, z, \zeta) \log(\zeta-\zeta_0) - \int_{\zeta_0}^{\zeta} \log(\tau-\zeta_0) \frac{\partial}{\partial \tau} G(z_0, \tau, z, \zeta) d\tau, \end{aligned}$$

een in (D, D') , met uitzondering van de coupures, analytische oplossing van (F_0) .

Voer een nieuwe integratievariabele σ in volgens

$$t - z_0 = \sigma(z - z_0) \quad \tau - \zeta_0 = \sigma(\zeta - \zeta_0)$$

$$\begin{aligned} U(z, \zeta) &= G(z, \zeta_0, z, \zeta) \log(z - z_0) + \\ &- \int_0^1 [\log \sigma + \log(z - z_0)] \frac{\partial}{\partial \sigma} G(z_0 + (z - z_0)\sigma, \zeta_0, z, \zeta) d\sigma + \\ &+ G(z_0, \zeta, z, \zeta) \log(\zeta - \zeta_0) + \\ &- \int_0^1 [\log \sigma + \log(\zeta - \zeta_0)] \frac{\partial}{\partial \sigma} G(z_0, \zeta_0 + (\zeta - \zeta_0)\sigma, z, \zeta) d\sigma \\ &= G(z_0, \zeta_0, z, \zeta) \log(z - z_0) + G(z_0, \zeta_0, z, \zeta) \log(\zeta - \zeta_0) + \\ &- \int_0^1 \log \sigma \frac{\partial}{\partial \sigma} [G(z_0 + (z - z_0)\sigma, \zeta_0, z, \zeta) + G(z_0, \zeta_0 + (\zeta - \zeta_0)\sigma, z, \zeta)] d\sigma. \end{aligned}$$

We hebben aldus de volgende particuliere oplossing van (F_0) verkregen:

$$\begin{aligned} (11.26) \quad \Omega_1(z, \zeta, z_0, \zeta_0) &= G(z_0, \zeta_0, z, \zeta) \log[(z - z_0)(\zeta - \zeta_0)] + \\ &- \Omega_0(z, \zeta, z_0, \zeta_0) \end{aligned}$$

waarin

$$\begin{aligned} (11.27) \quad \Omega_0(z, \zeta, z_0, \zeta_0) &= \\ &\int_0^1 \log \sigma \frac{\partial}{\partial \sigma} [G(z_0 + (z - z_0)\sigma, \zeta_0, z, \zeta) + G(z_0, \zeta_0 + (\zeta - \zeta_0)\sigma, z, \zeta)] d\sigma. \end{aligned}$$

Ω_1 is een meerwaardige functie. Echter, als we in (11.26) z, ζ, z_0, ζ_0 vervangen door $z = x + iy, \zeta = x - iy, z_0 = x_0 + iy_0, \zeta_0 = x_0 - iy_0$ waarbij $(x, y), (x_0, y_0)$ punten uit het gebied D zijn, dan krijgen we de volgende eenwaardige functie

$$(11.28) \quad \omega_1(x, y, x_0, y_0) = g(x, y, x_0, y_0) \log [(x-x_0)^2 + (y-y_0)^2] + \\ - \omega_0(x, y, x_0, y_0).$$

$$\omega_1(x, y, x_0, y_0) = \Omega_1(x+iy, x-iy, x_0+iy_0, x_0-iy_0)$$

$$g(x, y, x_0, y_0) = G(x_0+iy_0, x_0-iy_0, x+iy, x-iy)$$

$$\omega_0(x, y, x_0, y_0) = \Omega_0(x+iy, x-iy, x_0+iy_0, x_0-iy_0).$$

$\omega_1(x, y, x_0, y_0)$ is een analytische oplossing van (E_0) voor $(x, y) \in D \setminus (x_0, y_0)$. Aangezien $g(x_0, y_0, x_0, y_0) = 1$ en uit analyticiteit regulariteit volgt, is $\omega_1(x, y, x_0, y_0)$ een fundamentealoplossing van (E_0) met een pool in (x_0, y_0) .

Stelling 11.5.

Er bestaan fundamentealoplossingen $\omega(x, y, x_0, y_0)$ van (E_0) met een pool in het punt (x_0, y_0) , die, met betrekking tot de variabelen x_0, y_0 , fundamentealoplossing zijn voor de geadjungeerde vergelijking (E_0^*) met een pool in het punt (x, y) . $\omega(x, y, x_0, y_0)$ is een analytische functie, zowel van x, y als van x_0, y_0 , in het gebied D met uitzondering van het punt (x_0, y_0) resp. (x, y) . Dergelijke fundamentealoplossingen noemen we standaardfundamentealoplossingen.

Voor het bewijs zij verwezen naar [1], pag. 29. Met behulp van distributietheorie kunnen we afleiden

$$(11.29) \quad E_{xy} [\omega(x, y, \xi, \eta)] = c \delta(P, P_0) \quad P(x, y), P_0(\xi, \eta),$$

$$(11.30) \quad E_{\xi\eta}^* [\omega(x, y, \xi, \eta)] = c \delta(P, P_0) \quad P(\xi, \eta), P_0(x, y),$$

waarin c een willekeurige constante is en δ de deltafunctie van Dirac.

In het vervolg zullen we voor ω de genormaliseerde standaardfundamentealoplossing kiezen. De constante c in de formules (11.29) en (11.30) reduceert zich dan tot $c = -1$.

Opmerking 11.2.

De fundamenteeloplossing (11.24) van (E_0) is een generalisatie van de oplossing $u = \log r$ voor de Laplace vergelijking $\Delta u = 0$. Bij het oplossen van het Dirichlet probleem voor elliptische differentiaalvergelijkingen gebruiken we speciale fundamenteeloplossingen, de zgn. Greense functies. Deze hebben de eigenschap gelijk aan nul te zijn op de rand van het beschouwde gebied wat bereikt kan worden door een geschikte keuze van het reguliere deel (vgl. spiegelmethode voor de Laplacevergelijking).

11.5. Analyticiteit van de oplossingen van (E_0)

Zij W een gebied waarvan de randkromme L stuksgewijs glad is en een positieve omlooprichting heeft. Zijn $u(x,y)$, $v(x,y)$ eenwaardige functies met continue eerste en tweede orde partiële afgeleiden in W en continue eerste orde afgeleiden in $W + L$. Dan geldt, op grond van de divergentiestelling van Gauss, de formule

$$(11.31) \quad \iint_W [vE_{xy}(u) - uE_{xy}^*(v)] dx dy = \int_L [u \frac{dv}{dn} - v \frac{du}{dn} - uv(a \cos(n,x) + b \cos(n,y))] ds,$$

waarin n de naar binnen gerichte normaal en s de booglengte van L is; a , b , c zijn de coëfficiënten van de differentiaaloperator E . Deze zijn analytisch in $W + L$. Kies nu in (11.31) voor u een reguliere oplossing $u(x,y)$ van (E_0) in het gebied W ($W \subset D$) en voor $v(x,y)$ de genormaliseerde standaardfundamenteeloplossing $\omega(\xi, \eta, x, y)$. Dan geldt op grond van (11.30):

$$(11.32) \quad \iint_W u(x,y) \delta(P, P_0) dx dy = \int_L [u(x,y) \frac{d\omega(\xi, \eta, x, y)}{dn} - \omega(\xi, \eta, x, y) Nu] ds,$$

waarin $P_0(\xi, \eta)$ een willekeurig punt van W is, (x,y) het integratiepunt zowel in het linkerlid (punt van het gebied W) als in het rechterlid (punt van de rand L) en

$$Nu = \frac{du(x,y)}{dn} + [a \cos(n,x) + b \cos(n,y)] u(x,y).$$

Uit (11.32) volgt, op grond van de eigenschappen van de δ functie van Dirac en het feit dat $u(x,y)$ regulier is

$$u(\xi, \eta) = \int_L [u(x,y) \frac{d\omega(\xi, \eta, x, y)}{dn} - \omega(\xi, \eta, x, y) Nu] ds,$$

of ook

$$(11.33) \quad u(x,y) = \int_L [u(\xi, \eta) \frac{d\omega(x, y, \xi, \eta)}{dn} - \omega(x, y, \xi, \eta) Nu] ds$$

waarbij (ξ, η) het integratiepunt is en

$$(11.34) \quad Nu = \frac{du(\xi, \eta)}{dn} + [a \cos(n, \xi) + b \cos(n, \eta)] u(\xi, \eta).$$

Stelling 11.6. (Picard)

Als a , b , c analytische functies zijn in het gebied W , dan is iedere in W reguliere oplossing van (E_0) een analytische functie van de variabelen x en y in W .

Het bewijs volgt onmiddellijk uit de voor de reguliere oplossing $u(x,y)$ geldige formule (11.33), aangezien aangetoond is dat $\omega(x,y,\xi,\eta)$ een analytische functie is van x en y .

11.6. Analytische voortzetting van de oplossingen van (E_0)

Zij D een fundamenteelgebied van (E_0) . Zij W een enkelvoudig samenhangend gebied met stuksgewijs gladde randkromme L , zodanig dat $W + L \subset D$. Iedere in D reguliere oplossing $u(x,y)$ van (E_0) kan dan in W gerepresenteerd worden met behulp van formule (11.33). Aangezien volgens stelling 11.6 $u(x,y)$ een analytische functie is van de beide reële variabelen x en y in W , is $u(x,y)$ analytisch voortzetbaar voor complexe waarden van de argumenten. Formule (11.33) stelt ons in staat een expliciete representatie af te leiden voor een analytische voortzetting van de reguliere oplossing $u(x,y)$.

Vervang in het rechterlid van (11.33) x door $\frac{1}{2} (z+\zeta)$ en y door $\frac{1}{2i} (z-\zeta)$,

waarbij $z \in W$, $\zeta \in \bar{W}$. We krijgen dan de functie

$$(11.35) \quad U(z, \zeta) = \int_L [U(t, \bar{t}) \frac{d}{dn} \Omega(z, \zeta, t, \bar{t}) - \Omega(z, \zeta, t, \bar{t}) NU] ds,$$

waarin

$$\Omega(z, \zeta, t, \bar{t}) = \omega\left(\frac{z+\zeta}{2}, \frac{z-\zeta}{2i}, \frac{t+\bar{t}}{2}, \frac{t-\bar{t}}{2i}\right)$$

en $t = \xi + i\eta \in L$.

Op grond van het feit dat ω de genormaliseerde standaardfundamenteaaloplossing is en op grond van de formules (11.26) en (11.28), geldt

$$(11.36) \quad \Omega(z, \zeta, t, \bar{t}) = -\frac{1}{4\pi} G(t, \bar{t}, z, \zeta) \log [(t-z)(\bar{t}-\zeta)] + \\ + \Omega'_0(z, \zeta, t, \bar{t}).$$

Het reguliere deel $\Omega'_0(z, \zeta, t, \bar{t})$ is een analytische functie van zijn argumenten in $(D, \bar{D}, D, \bar{D})$ evenals $G(t, \bar{t}, z, \zeta)$. Ook $\log (t-z)(\bar{t}-\zeta)$ is een analytische functie van z en ζ in W resp. $\bar{W}$, aangezien t en $\bar{t}$ randpunten zijn en de coupures buiten W resp. $\bar{W}$ gekozen worden. Dientengevolge is $\Omega(z, \zeta, t, \bar{t})$ en dus $U(z, \zeta)$ een analytische functie van z, ζ in $(W, \bar{W})$.

We kunnen W echter zodanig kiezen dat zijn rand de rand van D willekeurig dicht benaderd. Ook geldt, dat $u(x, y)$ volledig onafhankelijk is van de keuze van W , zolang $W \subset D$.

Dus $U(z, \zeta)$ is een analytische functie van z, ζ in $(D, \bar{D})$, die samenvalt met $u(x, y)$ als $z = x + iy$, $\zeta = x - iy$, met $(x, y) \in D$. Hieruit volgt, dat $U(z, \zeta)$ een analytische voortzetting is van $u(x, y)$ in het gebied van complexe waarden van x en y . We kunnen nu de volgende stelling formuleren:

Stelling 11.7. (Vekua)

De analytische voortzetting van de oplossing $u(x, y)$ van (E_0) , regulier in het fundamenteaalgebied D , is een analytische functie van de variabelen $z = x + iy$, $\zeta = x - iy$ in $(D, \bar{D})$ en wordt gegeven door (11.35), waarin de kromme L een willekeurige samenhangende stuksgewijs gladde gesloten kromme in D mag zijn.

Conclusie

Als $u(x,y)$ een in D reguliere oplossing van (E_0) is, dan wordt door de uitdrukking

$$(11.37) \quad U(z, \zeta) = u\left(\frac{z+\zeta}{2}, \frac{z-\zeta}{2i}\right)$$

een analytische functie van z, ζ in $(D, \overline{D})$ gegeven, die $u(x,y)$ analytisch voortzet in het gebied van complexe x en y .

11.7. Algemene representatie van de reguliere oplossingen van (E_0)

D zij een fundamenteelgebied van (E_0) en $u(x,y)$ een in D reguliere oplossing van (E_0) . Opgrond van de stelling van Vekua kunnen we $u(x,y)$ analytisch voortzetten volgens

$$U(z, \zeta) = u\left(\frac{z+\zeta}{2}, \frac{z-\zeta}{2i}\right),$$

een analytische functie van z, ζ in $(D, \overline{D})$, die voldoet aan $F(U) = 0$.

Volgens stelling 11.3 geldt dan

$$(11.18) \quad U(z, \zeta) = \alpha_0 G(z_0, \zeta_0, z, \zeta) + \int_{z_0}^z \phi(t) G(t, \zeta_0, z, \zeta) dt + \\ + \int_{\zeta_0}^{\zeta} \phi^*(\tau) G(z_0, \tau, z, \zeta) d\tau,$$

waarbij z_0, ζ_0 vaste punten zijn in D resp. $\overline{D}$ en $\alpha_0, \phi(z), \phi^*(\zeta)$ voldoen aan (11.19).

Door in (11.18) te substitueren $z = x + iy, \zeta = x - iy$, waarbij $(x,y) \in D$ krijgen we

$$(11.20) \quad u(x,y) = \alpha_0 G(z_0, \zeta_0, z, \overline{z}) + \int_{z_0}^z \phi(t) G(t, \zeta_0, z, \overline{z}) dt + \\ + \int_{\zeta_0}^{\overline{z}} \phi^*(\tau) G(z_0, \tau, z, \overline{z}) d\tau.$$

We hebben hiermee bewezen dat elke in D reguliere oplossing $u(x,y)$ van (E_0)

in de vorm (11.20) geschreven kan worden, waarbij de constante α_0 en de in D resp. $\bar{D}$ analytische functies $\phi(z)$, $\phi^*(\zeta)$ op eenduidige wijze bepaald zijn door $u(x,y)$ via (11.19).

Omgekeerd is het rechterlid van (11.20) volgens stelling 11.4, voor elke constante α_0 en willekeurige in D resp. $\bar{D}$ analytische functies $\phi(z)$, $\phi^*(\zeta)$, een analytische (en dus reguliere) oplossing van (E_0) .

We hebben hiermee bewezen:

Stelling 11.8

Zij D een fundamenteaalgebied van (E_0) . Dan levert formule (11.20), waarin α_0 een willekeurige constante is en $\phi(z)$, $\phi^*(\zeta)$ willekeurige analytische functies zijn in D resp. $\bar{D}$, alle in D reguliere oplossingen van (E_0) , terwijl α_0 , ϕ , ϕ^* eenduidig gedefinieerd zijn in termen van $u(x,y)$ via (11.19).

Literatuur

- [1] Vekua, I.N. New methods for solving elliptic equations.
North-Holland publ. comp., Amsterdam, 1968.
- [2] Henrici, P. ZAMP 8(1957), pag. 169-203.
- [3] Pogorzelski, Integral equations and their applications.
Pergamon press, Warschau, 1966.