



# Iteratieve methoden voor de golfvergelijking

Nicolas Daans

Begeleiders: Wim Vanroose en Siegfried Cools

20 mei 2016

# Inhoudsopgave

<b>Inleiding</b>	<b>1</b>
<b>1 Eendimensionale Helmholtzvergelijking</b>	<b>3</b>
1.1 Homogene Helmholtzvergelijking . . . . .	3
1.2 Trillende snaren en rechts- en linkslopende golven . . . . .	4
1.3 Sommerfeld randvoorwaarden . . . . .	6
1.4 Numerieke oplossing: eindige differentiemethode . . . . .	6
1.5 Transmissie- en reflectiecoëfficiënten . . . . .	9
<b>2 Veralgemeningen van de golfvergelijking</b>	<b>17</b>
2.1 Eindige differentiemethode in twee dimensies . . . . .	17
2.2 Tijdsafhankelijke Helmholtzvergelijking . . . . .	20
<b>3 Iteratieve methodes</b>	<b>24</b>
3.1 Eigenwaarden van coëfficiëntenmatrix . . . . .	24
3.2 Stationaire methoden . . . . .	25
3.3 GMRES . . . . .	30
<b>Besluit en outlook</b>	<b>33</b>

# Inleiding

De  $n$ -dimensionale golfvergelijking (of *Helmholtzvergelijking*) is een partiële differentiaalvergelijking in de  $n$  veranderlijken  $z = (z_1, \dots, z_n)$  gegeven door

$$(-\nabla^2 - k(z)^2)u(z) = f(z), \quad (1)$$

waar  $u(z)$  de onbekende functie is,  $k(z)$  en  $f(z)$  functies van  $z$  zijn en  $\nabla^2 u(z)$  gedefinieerd wordt door

$$\nabla^2 u(z) = \sum_{i=1}^n \frac{\partial^2 u(z)}{\partial z_i^2}. \quad (2)$$

Deze vergelijking duikt op in de beschrijving van allerlei soorten golven, zoals geluidsgolven, lichtgolven en elektromagnetische golven. Vanuit de fysica halen we de volgende terminologie: de functie  $k(x)$  zullen we altijd het *golfgetal* noemen, de functie  $f(x)$  de *bron*. We noemen de vergelijking *homogeen* als  $f(x) = 0$  op het hele definitiegebied en anders noemen we ze *inhomogeen*. Het definitiegebied voor de functie  $f(x)$ ,  $k(x)$  en  $u(x)$  zullen we het *domein* noemen. Hierop zullen we later de nodige restricties leggen.

Voor een fysische afleiding van de eendimensionale golfvergelijking bij het beschrijven van een trillende snaar, verwijzen we naar [4, p. 3-5]. In dit artikel zullen we ons niet toespitsen op de fysische oorsprong van de golfvergelijking, maar op haar numerieke oplossing. Wel zullen we trachten de plaats van de vergelijking en de optredende veranderlijken in de fysica te duiden. Dit zal ook relevant blijken bij het kiezen van goede randvoorwaarden voor de differentiaalvergelijking.

In het eerste hoofdstuk van dit artikel bekijken we het eendimensionale probleem. We leiden een numerieke methode af die de vergelijking discretiseert en omzet naar een groot, ijl lineair stelsel. We bespreken kort een verband met de fysica en gebruiken dit om gepaste randvoorwaarden af te leiden voor de vergelijking. Tenslotte bestuderen we hoe we inkomende en uitgaande golven kunnen kwantificeren op basis van een numerieke oplossing.

In het tweede hoofdstuk worden de numerieke methoden uit het eerste hoofdstuk op twee manieren veralgemeend. Ten eerste wordt gekeken naar hoe de tweedimensionale golfvergelijking kan worden aangepakt. Geheel analoog aan het eendimensionale geval zal dit leiden tot een groot en ijl lineair stelsel. Ten tweede beschouwen we een numerieke methode om een variant van de golfvergelijking - de tijdsafhankelijke Helmholtzvergelijking (1.10) - numeriek op te lossen.

In beide gevallen zullen we niet ingaan op de optimale randvoorwaarden. Voor het eerste probleem worden goede randvoorwaarden voorgesteld in [4, p. 17 e.v.] en in de bachelorproef van mijn collega Teppo Mertens. [2]

Het laatste hoofdstuk gaat over hoe iteratieve methodes gebruikt kunnen worden om de optredende ijle lineaire stelsels op te lossen. Verschillende methoden worden vergeleken op basis van verschillende criteria. In het bijzonder bestuderen we de toepassing van enkele stationaire methoden en GMRES.

Ik wil graag mijn promotors prof. dr. Wim Vanroose en dr. Siegfried Cools bedanken. Professor Vanroose zorgde voor de inspiratie om dit artikel te schrijven en begeleidde mij doorheen het proces door onderzoeksvragen voor te stellen en me te helpen wanneer ik niet verder kon. Siegfried hielp me met een aantal berekeningen en met het programmeren en heeft ook meerdere tussentijdse versies van dit artikel nagelezen; met zijn opmerkingen heb ik veel fouten en onnauwkeurigheden uit dit artikel kunnen halen.

# Hoofdstuk 1

## Eendimensionale Helmholtzvergelijking

### 1.1 Homogene Helmholtzvergelijking

In het eendimensionale geval wordt de vergelijking gereduceerd tot de volgende gewone differentiaalvergelijking

$$\left(-\frac{d^2}{dx^2} - k(x)^2\right)u(x) = f(x), \quad (1.1)$$

waarin we  $z_1$  door  $x$  vervangen, zoals in de rest van dit hoofdstuk. We zullen ook in het hele hoofdstuk onderstellen dat het domein een compact interval is, dat we zonder verlies van algemeenheid van de vorm  $[0, L]$  mogen onderstellen voor  $L > 0$ .

We kijken eerst naar de simpelste variant: de homogene golfvergelijking met constant golfgetal  $k(x) = k$ . Uitwerken en beide leden met  $-1$  vermenigvuldigen levert

$$u''(x) + k^2 u(x) = 0, \quad (1.2)$$

waar we  $u''(x)$  voor  $\frac{\partial^2 u(x)}{\partial x^2}$  noteerden. Dit is een lineaire, homogene differentiaalvergelijking van de tweede orde, zodat haar oplossingen een tweedimensionale vectorruimte opspannen (zie [1, p. 18-29] voor een overzicht van de theorie rond lineaire differentiaalvergelijkingen). In essentie worden haar oplossingen dus bepaald door de keuze van twee lineair onafhankelijke oplossingen, *fundamentele oplossingen* genaamd. In dit artikel zullen we kiezen voor <sup>1</sup>

$$u_1(x) = e^{\iota k x} \quad \text{en} \quad u_2(x) = e^{-\iota k x}. \quad (1.3)$$

Deze keuze is niet uniek en de motivatie voor onze keuze wordt in het volgende deelhoofdstuk gegeven. Aangezien een willekeurige oplossing van vergelijking

---

<sup>1</sup>We gebruiken  $\iota$  voor de complexe eenheid, om verwarring met de index  $i$  later te vermijden.

(1.2) nu geschreven kan worden als

$$u(x) = Ae^{\iota kx} + Be^{-\iota kx}, \quad (1.4)$$

waar  $A$  en  $B$  constanten zijn, moeten we nog gepaste randvoorwaarden opleggen om tot een unieke oplossing te komen. Deze randvoorwaarden kunnen verschillende vormen aannemen. Enkele daarvan zijn:

- *Dirichlet randvoorwaarden*: Dit is het opleggen van vaste waarden voor de functie  $u(x)$  aan de randen van het domein. Doorgaans kiest men voor *homogene Dirichlet randvoorwaarden*, wat neerkomt op

$$u(0) = u(L) = 0. \quad (1.5)$$

- *Neumann randvoorwaarden*: Dit is het opleggen van waarden voor de eerste afgeleide van  $u(x)$  in de randpunten. Opnieuw kiest men dikwijls voor *homogene Neumann randvoorwaarden*:

$$u'(0) = u'(L) = 0. \quad (1.6)$$

Laten we bij wijze van illustratie uitrekenen wat homogene Dirichlet randvoorwaarden in ons geval opleveren. We bekomen een stelsel in  $A$  en  $B$ :

$$u(0) = Ae^{\iota k0} + Be^{-\iota k0} = A + B = 0 \quad (1.7)$$

$$u(L) = Ae^{\iota kL} + Be^{-\iota kL} = 0 \quad (1.8)$$

Dit stelsel heeft  $A = B = 0$  als triviale oplossing. Er gebeurt echter iets opmerkelijks wanneer

$$L = \frac{n\pi}{k} \quad \text{voor } n \in \mathbb{N}. \quad (1.9)$$

In dit geval is  $A = -B = c$  een oplossing voor *elk* getal  $c$ ! Niet alleen is dit een heel tegeninuitieve uitkomst - er zouden enkel niet-triviale oplossingen zijn voor zeer specifieke waarden voor  $L$  - het toont ook aan dat de homogene Dirichlet randvoorwaarden niet altijd tot een unieke oplossing hoeven te leiden. Expliciet nog kan men inzien dat, als  $u(x)$  een oplossing is van de vergelijking met Dirichlet randvoorwaarden en  $L$  zoals hierboven, dan  $u(x) + \alpha \sin(kx)$  ook voor elke waarde van  $\alpha$ .

De homogene Neumann randvoorwaarden leiden tot een haast identiek probleem. De vraag is dan ook welke randvoorwaarden wél geschikt zijn. Men zou kunnen trachten op zoek te gaan naar inhomogene Dirichlet/Neumann randvoorwaarden of een combinatie van de twee; wij zullen onze inspiratie halen bij een toepassing uit de fysica.

## 1.2 Trillende snaren en rechts- en linkslopende golven

Beschouw een snaar van lengte  $L$ , opgespannen tussen  $x$ -waarden 0 en  $L$ . Uit [4, p3 - 4] weten we dat de verticale uitwijking  $u(x, t)$  van een punt op de snaar

met  $x$ -waarde  $x$  op tijdstip  $t$  voldoet aan de (partiële) differentiaalvergelijking

$$\left( \frac{1}{c(x)^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} \right) u(x, t) = F(x, t), \quad (1.10)$$

waar  $c(x)$  een fysische (materiaal)constante is en  $F(x, t)$  een trillingsbron voorstelt.

Onderstellen we dat de trillingsbron periodiek is, zeg maar een oscillatie met constante hoekfrequentie  $\omega$

$$F(x, t) = e^{-i\omega t} f(x), \quad (1.11)$$

dan kunnen we onderstellen dat de uitwijking in functie van de tijd gegeven wordt door

$$u(x, t) = e^{-i\omega t} u(x). \quad (1.12)$$

Met andere woorden: we onderstellen dat de oplossing geschreven kan worden als het product van een tijdsafhankelijke en een plaatsafhankelijke component. De vorige twee vergelijkingen invullen in (1.10) en uitwerken levert

$$\left( \frac{1}{c(x)^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} \right) e^{-i\omega t} u(x) = e^{-i\omega t} f(x) \quad (1.13)$$

$$\frac{-\omega^2}{c(x)^2} e^{-i\omega t} u(x) - e^{-i\omega t} u''(x) = e^{-i\omega t} f(x) \quad (1.14)$$

$$-u''(x) - k(x)^2 u(x) = f(x) \quad (1.15)$$

waar we in de laatste stap beide leden delen door  $e^{-i\omega t}$  en  $k(x) = \omega/c(x)$  stelden. Als we eisen dat er geen bron is in (een omgeving van) de randpunten van het domein, dat wil zeggen  $f(x) = 0$  voor  $0 \leq x < \epsilon$  en  $L - \epsilon < x \leq L$  ( $\epsilon$  voldoende klein), en dat  $k = k(x)$  daar constant is, dan vinden we de homogene vergelijking (1.2) terug. Vullen we de oplossing (1.4) in (1.12) in, dan vinden we

$$u(x, t) = A e^{-i\omega t + ikx} + B e^{-i\omega t - ikx} \quad (1.16)$$

We observeren dat de eerste term constant wordt wanneer

$$kx - \omega t = 0 \quad (1.17)$$

zodat, wanneer de tijd met  $k$  toeneemt, de golf  $A e^{ikx}$  met  $\omega$  naar rechts verschoven is. We spreken van een *rechtslopende golf*. Analoog is de tweede term constant wanneer

$$-kx - \omega t = 0 \quad (1.18)$$

zodat we van een *linkslopende golf* spreken. Dit motiveert nu ook de keuze voor  $e^{\pm ikx}$  als fundamentele oplossingen in het vorige deelhoofdstuk.

### 1.3 Sommerfeld randvoorwaarden

De afleiding in het vorige hoofdstuk suggereert een specifiek soort randvoorwaarden. Als de trillingsbron zich volledig binnen het domein bevindt, dan mogen er aan de randen van het domein enkel uitgaande golven zijn, dat wil zeggen linkslopende golven aan de linkerkant en rechtslopende aan de rechterkant. Deze conditie blijkt precies gerealiseerd te worden door het opleggen van volgende randvoorwaarden:

$$\lim_{x \rightarrow 0^+} \left( \frac{du(x)}{dx} + \iota k u(x) \right) = 0 \quad (1.19)$$

$$\lim_{x \rightarrow L^-} \left( \frac{du(x)}{dx} - \iota k u(x) \right) = 0. \quad (1.20)$$

Inderdaad, schrijven we de oplossing op de rechterkant van het domein in de vorm  $u(x) = Ae^{\iota kx} + Be^{-\iota kx}$ , dan vinden we uit (1.20) de eis

$$\iota k(Ae^{\iota kx} - Be^{-\iota kx}) = \iota k(Ae^{\iota kx} + Be^{-\iota kx}) \quad (1.21)$$

waaruit we  $B = 0$  halen, zodat enkel de rechtslopende golf  $Ae^{\iota kx}$  overblijft. Op een analoge manier vinden we dat  $A = 0$  moet gelden aan de linkerkant van het domein. We krijgen dus juist wat we wilden opleggen.

De voorwaarden (1.19) en (1.20) noemt men *Sommerfeld randvoorwaarden*. Men spreekt algemener ook over *uitgaande golf randvoorwaarden*. Sommerfeld bewees dat deze randvoorwaarden altijd tot een unieke oplossing leiden. [3]

### 1.4 Numerieke oplossing: eindige differentiemethode

We beschouwen nu terug de meest algemene vorm van de eendimensionale golfvergelijking uit het begin van dit hoofdstuk

$$\left( -\frac{d^2}{dx^2} - k(x)^2 \right) u(x) = f(x). \quad (1.22)$$

Als  $k(x)$  en  $f(x)$  plaatsafhankelijk zijn, bestaat er geen analytische oplossingsmethode meer voor deze vergelijking. We doen dan beroep op een numerieke oplossingsmethode, die we de *eindige differentiemethode* noemen.

We onderstellen dat we een oplossing  $u(x)$  hebben en kiezen  $N + 1 \geq 3$  equidistante punten in het interval  $[0, L]$ :

$$x_j = jh \quad j \in \{0, 1, \dots, N\}, \quad (1.23)$$

waar  $h = L/N$ . Verder stellen we voor iedere  $i \in \{0, 1, \dots, N\}$

$$k_i = k(x_i) \quad f_i = f(x_i) \quad (1.24)$$



Door het uitvoeren van een Taylor-expansie rond  $x_i$  vinden we dan voor  $i \in \{0, 1, \dots, N-1\}$

$$u(x_{i+1}) = u(x_i) + hu'(x_i) + \frac{h^2}{2}u''(x_i) + \frac{h^3}{6}u^{(3)}(x_i) + \frac{h^4}{24}u^{(4)}(x_i + \epsilon_{i,1}h) \quad (1.25)$$

en voor  $i \in \{1, 2, \dots, N\}$

$$u(x_{i-1}) = u(x_i) - hu'(x_i) + \frac{h^2}{2}u''(x_i) - \frac{h^3}{6}u^{(3)}(x_i) + \frac{h^4}{24}u^{(4)}(x_i + \epsilon_{i,2}h). \quad (1.26)$$

waar  $\epsilon_{i,1}, \epsilon_{i,2} \in [0, 1]$ . In de onderstelling dat  $u^{(4)}(x)$  lokaal begrensd is, zal de laatste term van de twee bovenstaande expansies naar nul naderen voor voldoende kleine  $h$  (voldoende grote  $N$ ). Optellen van de uitdrukkingen voor  $u(x_{i+1})$  en  $u(x_{i-1})$  levert zelfs

$$u(x_{i+1}) + u(x_{i-1}) = 2u(x_i) + h^2u''(x_i) + \mathcal{O}(h^4). \quad (1.27)$$

Nu gebruiken we dat  $u(x)$  een oplossing is van (1.22). Dan moet immers gelden dat

$$-u''(x_i) - k_i^2u(x_i) = f_i. \quad (1.28)$$

De vorige twee vergelijkingen combineren levert na herschikking

$$\frac{1}{h^2}u(x_{i-1}) + \left(k_i^2 - \frac{2}{h^2}\right)u(x_i) + \frac{1}{h^2}u(x_{i+1}) = -f_i + \mathcal{O}(h^2) \quad (1.29)$$

en dit voor  $i \in \{1, 2, \dots, N-1\}$ . We bekommen een lineair stelsel van  $N-1$  vergelijkingen in de  $N+1$  onbekenden  $u_i$  met  $i \in \{0, 1, 2, \dots, N\}$  door voor  $i \in \{1, \dots, N-1\}$

$$\frac{1}{h^2}u_{i-1} + \left(k_i^2 - \frac{2}{h^2}\right)u_i + \frac{1}{h^2}u_{i+1} = -f_i \quad (1.30)$$

te stellen. We vatten de  $u_i$  dus op als benaderingen voor  $u(x_i)$ . De overgebleven twee vergelijkingen, benodigd om tot een unieke oplossing te komen, zullen we uit de Sommerfeld randvoorwaarden halen.

Sommerfeld (1.19) toegepast op de linkerrand levert

$$u'(x_0) = -\iota k_0 u(x_0). \quad (1.31)$$

We zullen de resultaten (1.25) en (1.28) met de Sommerfeld randvoorwaarde combineren:

$$u(x_1) = u(x_0) + hu'(x_0) + \frac{h^2}{2}u''(x_0) + \mathcal{O}(h^3) \quad (1.32)$$

$$= u(x_0) - \iota k_0 hu(x_0) + \frac{h^2}{2}(-f_0 - k_0^2u(x_0)) + \mathcal{O}(h^3) \quad (1.33)$$

$$= u(x_0) \left(1 - \iota k_0 h - \frac{k_0^2 h^2}{2}\right) + \mathcal{O}(h^3), \quad (1.34)$$

waarbij we in de laatste gelijkheid ook nog  $f_0 = 0$  gebruikten (de bron ligt volledig in het inwendige van het domein). Analoog krijgen we voor de rechterrاند (door (1.20) en (1.26) te gebruiken)

$$u(x_{N-1}) = u(x_N) - hu'(x_N) + \frac{h^2}{2}u''(x_N) + \mathcal{O}(h^3) \quad (1.35)$$

$$= u(x_N) - \iota k_N hu(x_N) + \frac{h^2}{2}(-f_N - k_N^2 u(x_N)) + \mathcal{O}(h^3) \quad (1.36)$$

$$= u(x_N) \left( 1 - \iota k_N h - \frac{k_N^2 h^2}{2} \right) + \mathcal{O}(h^3). \quad (1.37)$$

De twee vergelijkingen in  $u_1$  en  $u_0$  (respectievelijk  $u_N$  en  $u_{N-1}$ )

$$u_1 = u_0 \left( 1 - \iota k_0 h - \frac{k_0^2 h^2}{2} \right) \quad (1.38)$$

$$u_{N-1} = u_N \left( 1 - \iota k_N h - \frac{k_N^2 h^2}{2} \right) \quad (1.39)$$

brengen ons tot de gewenste  $N + 1$  vergelijkingen. In matrixvorm heeft het stelsel dat we oplossen dus als coëfficiëntenmatrix  $A$ :

$$\begin{bmatrix} \frac{\iota k_0}{h} + \frac{k_0^2}{2} - \frac{1}{h^2} & \frac{1}{h^2} & 0 & 0 & \cdots & 0 \\ \frac{1}{h^2} & \frac{-2}{h^2} + k_1^2 & \frac{1}{h^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{h^2} & \frac{-2}{h^2} + k_2^2 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & \frac{1}{h^2} & 0 \\ \vdots & \vdots & \ddots & \frac{1}{h^2} & \frac{-2}{h^2} + k_{N-1}^2 & \frac{1}{h^2} \\ 0 & 0 & \cdots & 0 & \frac{1}{h^2} & \frac{\iota k_N}{h} + \frac{k_N^2}{2} - \frac{1}{h^2} \end{bmatrix}$$

en we berekenen dus  $u_0, u_1, \dots, u_N$  door

$$A \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_N \end{bmatrix} = \begin{bmatrix} -f_0 \\ -f_1 \\ \vdots \\ -f_N \end{bmatrix} \quad (1.40)$$

op te lossen, waarbij  $f_0 = f_N = 0$  moet gelden. Merk op dat we in de vergelijkingen (1.38) en (1.39) nog door een factor  $h^2$  gedeeld hebben met als doel de matrix  $A$  symmetrisch te maken. Dit zal nuttig zijn wanneer we in hoofdstuk 3 iteratieve methoden willen gebruiken om deze en gelijkaardige problemen aan te pakken.

In één dimensie is het echter nog mogelijk om met klassieke methoden dit stelsel op te lossen. Enkele voorbeelden van numerieke oplossingen met Sommerfeld randvoorwaarden worden geschetst in fig. 1.1 en fig. 1.2. De rechterhand wordt steeds beschreven door de puntbron

$$f(x) = \begin{cases} 1 & x = 10 \\ 0 & x \neq -10 \end{cases} \quad (1.41)$$

Merk op dat de kromme buiten de puntbron een sinusoidale golf beschrijft, zoals voorspeld door (1.4). In deze plots hebben we enkel het reële deel van de oplossing getekend. Denken we terug aan de interpretatie van zich voortplantende golven, dan correspondeert dit soort plot met een 'momentopname' van zo'n golf. Willen we weten hoe de golf er over een tijd  $t$  uitziet, dan moeten we deze functie vermenigvuldigen met  $e^{-i\omega t}$ , zoals beschreven in (1.12).

## Eindige differentiemethode met andere randvoorwaarden

We kunnen de eindige differentiemethode nog steeds toepassen met andere randvoorwaarden, zoals bijvoorbeeld de homogene Dirichlet randvoorwaarden. De afleiding in het vorige deel blijft dan grotendeels gelden, maar de vergelijkingen (1.38) en (1.39) vervangen we simpelweg door

$$u_0 = u_N = 0 \quad (1.42)$$

De coëfficiëntenmatrix  $A$  wordt dan ook

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ \frac{1}{h^2} & \frac{-2}{h^2} + k_1^2 & \frac{1}{h^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{h^2} & \frac{-2}{h^2} + k_2^2 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & \frac{1}{h^2} & 0 \\ \vdots & \vdots & \ddots & \frac{1}{h^2} & \frac{-2}{h^2} + k_{N-1}^2 & \frac{1}{h^2} \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix}$$

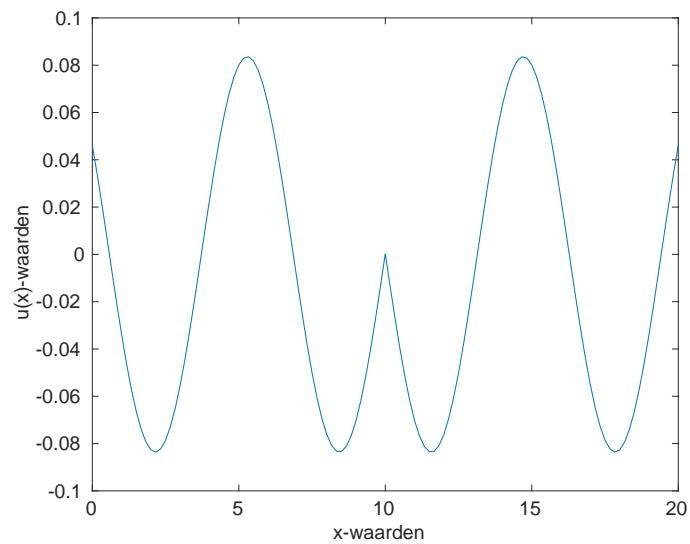
Figuur 1.3 geeft een voorbeeld van een oplossing met Dirichlet randvoorwaarden en de puntbron uit vorige sectie. Merk op dat de functie inderdaad nul wordt op de rand van het domein. Dit soort alternatieve randvoorwaarden kunnen eventueel gebruikt worden bij het simuleren van 'reflecterende randen' of andere fenomenen waarbij de randen een actieve rol spelen. Meestal willen we echter dat de randen niet interfereren met de golven in het domein, zodat we in principe dezelfde oplossing moeten bekomen na uitbreiding of beperking van het domein. De Sommerfeld randvoorwaarden willigen deze eis (in één dimensie) in.

## 1.5 Transmissie- en reflectiecoëfficiënten

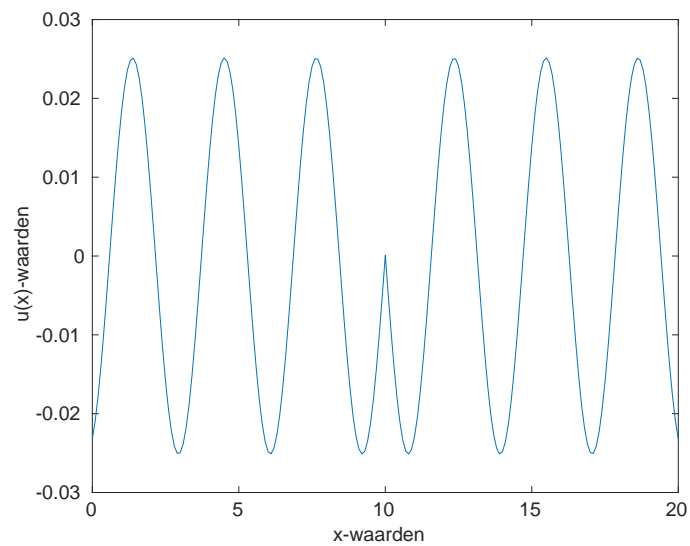
We werken voor de rest van dit hoofdstuk met de algemene eendimensionale Helmholtzvergelijking (1.22) en onderstellen dat op elke oplossing Sommerfeld randvoorwaarden opgelegd zijn.

We zagen dat, door het opleggen van de Sommerfeld randvoorwaarden, er aan de rechterrands van het domein enkel een rechtslopende golf kan zijn, dat wil zeggen van de vorm

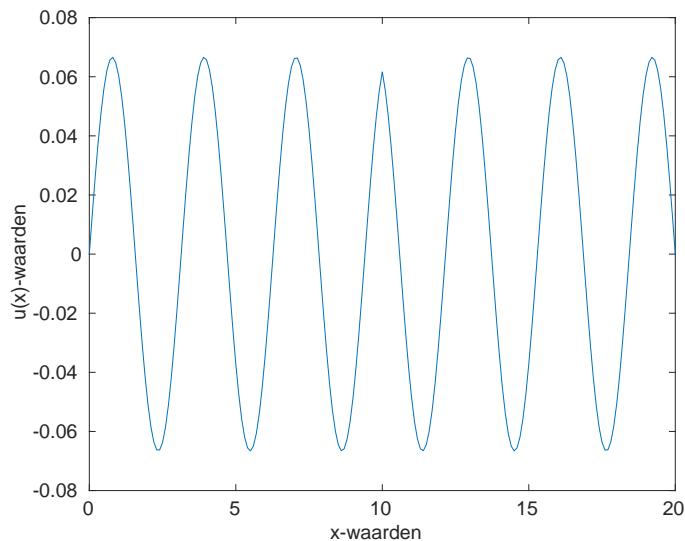
$$u(x) = Ae^{\iota kx} \quad (1.43)$$



Figuur 1.1: Oplossing met Sommerfeld randvoorwaarden,  $N = 120$ ,  $L = 20$  en  $k(x) = 1$ .



Figuur 1.2: Oplossing met Sommerfeld randvoorwaarden,  $N = 200$ ,  $L = 20$  en  $k(x) = 2$ .



Figuur 1.3: Oplossing met Dirichlet randvoorwaarden,  $L = 20$ ,  $k = 2$  en  $N = 200$ .

voor een zekere constante  $A \in \mathbb{C}$  en analoog kan er aan de linkerrand van het domein enkel een linkslopende golf zijn van de vorm

$$u(x) = Be^{-\iota kx} \quad (1.44)$$

voor een constante  $B \in \mathbb{C}$ . We zullen  $A$  de *transmissiecoëfficiënt* en  $B$  de *reflectiecoëfficiënt* noemen. Het doel van dit deelhoofdstuk is om een methode te vinden om deze waarden terug te vinden, gegeven een numerieke oplossing van de Helmholtzvergelijking.

De *Wronskiaanse determinant* (zie [1, p27]) van een paar functies van  $D \subseteq \mathbb{R}$  naar  $\mathbb{C}$ ,  $(u(x), v(x))$ , is de functie

$$W(u(\cdot), v(\cdot)) : D \rightarrow \mathbb{C} : x \mapsto \begin{vmatrix} u(x) & v(x) \\ u'(x) & v'(x) \end{vmatrix} \quad (1.45)$$

( $D$  een open deel van  $\mathbb{R}$  waarop  $u(x)$  en  $v(x)$  afleidbaar zijn). Laten we  $u(x)$  van de vorm  $Ae^{\iota kx} + Be^{-\iota kx}$  zijn, dan is:

$$W(u(x), e^{\iota kx}) = W(Ae^{\iota kx} + Be^{-\iota kx}, e^{\iota kx}) \quad (1.46)$$

$$= \begin{vmatrix} Ae^{\iota kx} + Be^{-\iota kx} & e^{\iota kx} \\ A\iota ke^{\iota kx} - B\iota ke^{-\iota kx} & \iota ke^{\iota kx} \end{vmatrix} \quad (1.47)$$

$$= B \begin{vmatrix} e^{-\iota kx} & e^{\iota kx} \\ -\iota ke^{-\iota kx} & \iota ke^{\iota kx} \end{vmatrix} = B(2\iota k) \quad (1.48)$$

en we besluiten dat

$$B = \frac{W(u(x), e^{\iota k x})}{2\iota k}. \quad (1.49)$$

Op analoge wijze bekomen we

$$A = \frac{W(u(x), e^{-\iota k x})}{-2\iota k}. \quad (1.50)$$

Nu zouden we deze waarde graag numeriek willen berekenen. Hiervoor is het nodig dat we uit de waarden  $u_i$  benaderingen voor  $u'(x_i)$  kunnen vinden. Voor een  $i \in \{1, 2, \dots, N-1\}$ , trek de vergelijking (1.26) van vergelijking (1.25) af om te bekomen

$$u(x_{i+1}) - u(x_{i-1}) = 2hu'(x_i) + \mathcal{O}(h^3) \quad (1.51)$$

$$u'(x_i) = \frac{u(x_{i+1}) - u(x_{i-1})}{2h} + \mathcal{O}(h^2). \quad (1.52)$$

Met de benadering

$$u'(x_i) \approx \frac{u_{i+1} - u_{i-1}}{2h} \quad (1.53)$$

kunnen we (1.49) verder uitrekenen in de punten  $u_i^2$ :

$$B = \frac{W(u(x_i), e^{\iota k_i x_i})}{2\iota k_i} = \frac{u(x_i)\iota k_i e^{\iota k_i x_i} - u'(x_i)e^{\iota k_i x}}{2\iota k_i} \quad (1.54)$$

$$\approx \frac{u_i \iota k_i e^{\iota k_i x_i} - \frac{u_{i+1} - u_{i-1}}{2h} e^{\iota k_i x_i}}{2\iota k_i} \quad (1.55)$$

$$= \frac{e^{\iota k_i x_i}}{2} \left( u_i + \frac{u_{i+1} - u_{i-1}}{2h k_i} \iota \right) \quad (1.56)$$

en analoog

$$A \approx \frac{e^{-\iota k_i x_i}}{2} \left( u_i - \frac{u_{i+1} - u_{i-1}}{2h k_i} \iota \right). \quad (1.57)$$

Laten we als besluit van dit hoofdstuk een uitgewerkt voorbeeld beschouwen. We stellen  $L = 10$ ,  $N = 300$  en

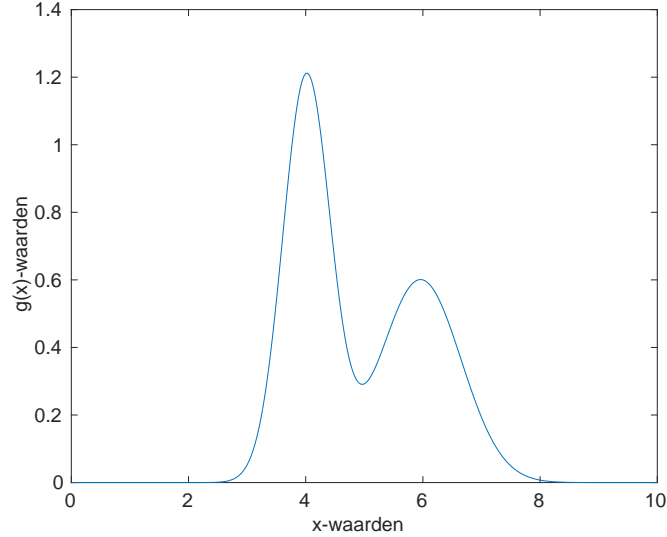
$$k(x) = k_0 + \iota g(x) \quad (1.58)$$

waarbij  $k_0 \in [0, +\infty[$  constant is en

$$g(x) = (10 - x)x \left( e^{-(x-3)^2} + 0.5e^{-2(x-6)^2} \right). \quad (1.59)$$

---

<sup>2</sup>Deze berekening houdt natuurlijk enkel steek wanneer  $u(x)$  in een omgeving van  $x_i$  inderdaad van de vorm  $Ae^{\iota k x} + Be^{-\iota k x}$  is voor constanten  $A$  en  $B$ .



Figuur 1.4: Plot van de functie  $g(x)$  uit (1.59)

De idee hierachter is de volgende: een golf  $e^{-\iota k_0 x}$  komt links in het domein aan. Er is geen bron in het domein, maar wel een obstakel gevormd door  $g(x)$ , dat de golf zal dempen. Immers,

$$e^{-\iota k(x)x} = e^{-\iota(k_0 + \iota g(x))x} = e^{-\iota k_0 x} e^{-g(x)x}. \quad (1.60)$$

We willen weten hoe de golf  $u(x)$  er uit ziet die ontstaat als gevolg van die inkomende golf. Omdat er geen bron in het domein aanwezig is, zal de totale golf  $u(x) + e^{-\iota k_0 x}$  moeten voldoen aan de homogene golfvergelijking. Dit betekent dat

$$-(u(x) + e^{-\iota k_0 x})'' - k(x)^2(u(x) + e^{-\iota k_0 x}) = 0 \quad (1.61)$$

hetgeen na uitwerking levert

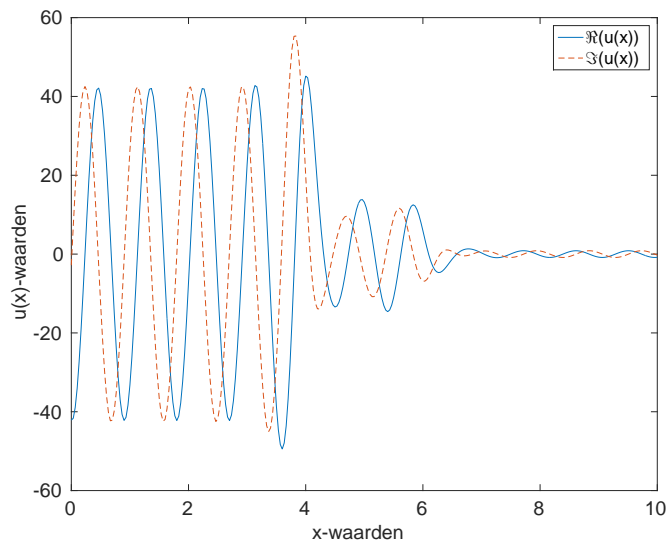
$$-u''(x) - k(x)^2 u(x) = (k(x)^2 - k_0^2) e^{-\iota k_0 x} \quad (1.62)$$

zodat we een inhomogene golfvergelijking in  $u(x)$  bekommen met als bron

$$f(x) = (k(x)^2 - k_0^2) e^{-\iota k_0 x}. \quad (1.63)$$

Merk op dat deze rechterhand in de randen van het domein nul wordt, omdat  $g(x)$  in de randen van het domein nul wordt.

We lossen dit numeriek op met Sommerfeld randvoorwaarden. In de plot van de oplossing (fig. 1.5) zien we dat er een eerste demping gebeurt ter hoogte van de eerste 'bult' in  $g(x)$  en nog een demping bij de tweede bult. Merk ook op dat de oplossing sinusoidaal is in de buurt van de randen van het domein, maar niet ertussen. Bovendien verschillen de inkomende en uitgaande amplitudes.



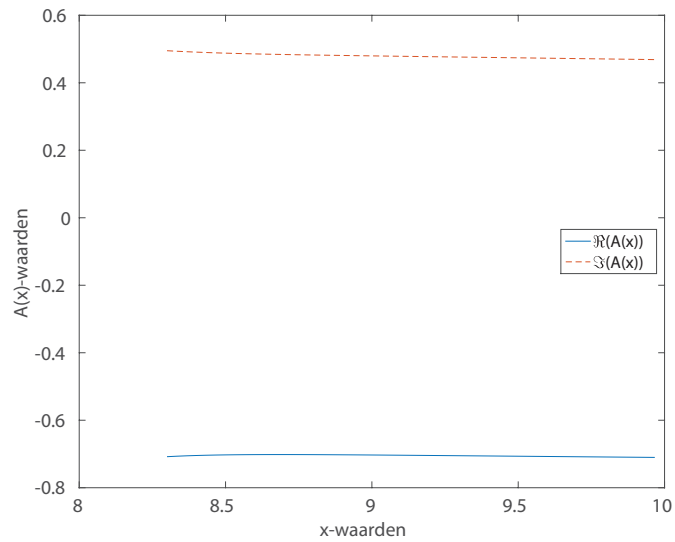
Figuur 1.5: Oplossing van het voorbeeldprobleem met  $k_0 = 7$ .

Met behulp van de formules uit dit hoofdstuk kunnen we de waarden  $A$  en  $B$  numeriek uitrekenen in alle punten tussen 0 en 10. Voor de transmissiecoëfficiënt  $A$  bijvoorbeeld zien we dat deze voor  $k_0 = 7$  inderdaad een constante waarde  $-0.7101 - 0.4687i$  aanneemt aan de rechterraand van het domein (fig. 1.6).

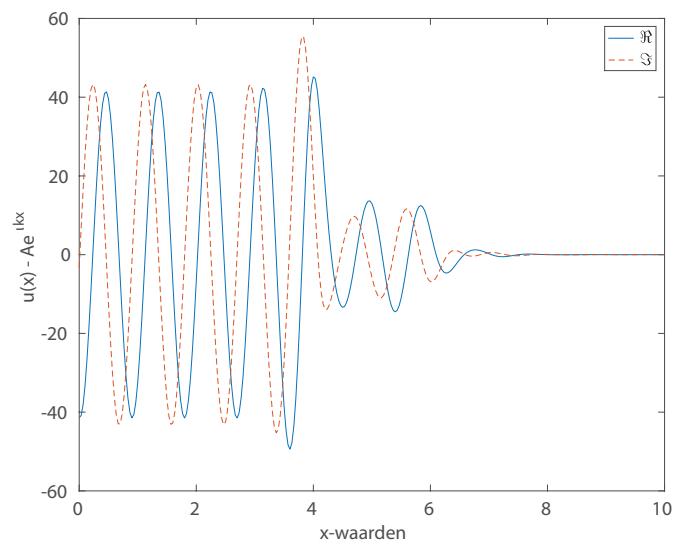
Om verder nog eens te illustreren dat de oplossing van de vorm  $Ae^{ikx}$  is aan de rechterraand, kunnen we voor  $k_0 = 7$  de plot bekijken van  $u - Ae^{ikx}$  met  $A = -0.7101 - 0.4687i$ , die inderdaad constant nul wordt rechts (fig. 1.7).

Tenslotte kan het interessant zijn om, gegeven de functie  $f(x)$ ,  $L$  en  $N$ , de evolutie van transmissie- en reflectiecoëfficiënten te plotten in functie van het (constant) golfgetal  $k_0$  (fig. 1.8 en fig. 1.9). Bij de eerste ziet men duidelijk dat de transmissie afneemt naarmate het golfgetal toeneemt; kortere golven worden harder gedempt. Dit fenomeen verklaart ook waarom lage tonen minder hard gedempt worden dan hoge tonen. Opvallend zijn echter ook de lokale pieken in de grafiek.

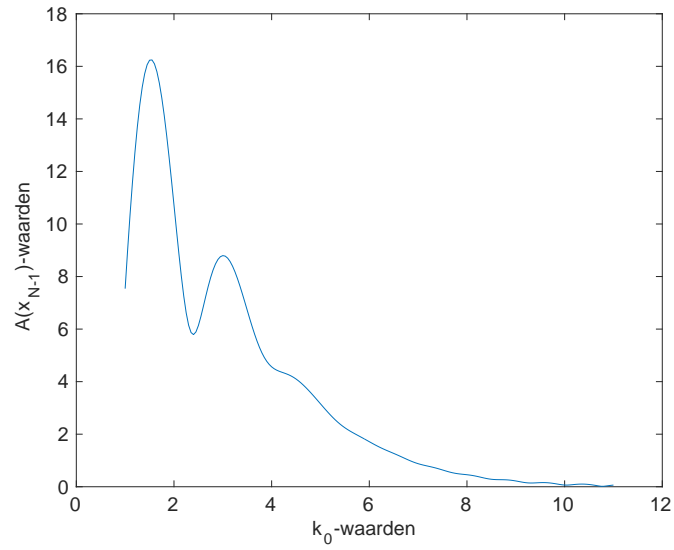




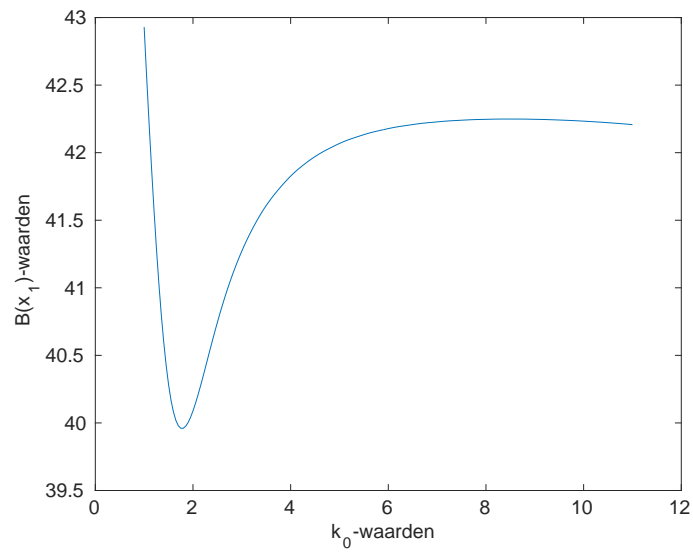
Figuur 1.6: Transmissiecoëfficiënt aan de rechterraand van het domein bij  $k_0 = 7$ .



Figuur 1.7: Plot van  $u(x) - Ae^{ikx}$ : merk op hoe de term  $Ae^{ikx}$  helemaal met  $u(x)$  samenvalt aan de rechterkant van het domein.



Figuur 1.8: Evolutie van de transmissiecoëfficiënt  $A$  in het punt  $x_{N-1}$ .



Figuur 1.9: Evolutie van de reflectiecoëfficiënt  $B$  in het punt  $x_1$ .

## Hoofdstuk 2

# Veralgemeningen van de golfvergelijking

### 2.1 Eindige differentiemethode in twee dimensies

We vertrekken vanuit de algemene 2-dimensionale Helmholtzvergelijking, zoals beschreven in (1) voor  $n = 2$ . We werken op een begrensd, vierkant domein<sup>1</sup>  $[0, L]^2$  en willen opnieuw een *grid* van equidistante punten kiezen, dat ons toelaat de differentiaalvergelijking om te vormen tot een stelsel van lineaire vergelijkingen.

Stel  $u(x, y)$  een oplossing, neem  $N \geq 2$  en definieer  $(N + 1)^2$  equidistante punten

$$x_{j,l} = (jh, lh) \quad j, l \in \{0, 1, \dots, N\}. \quad (2.1)$$

Analoog aan het eindimensionale geval, stellen we voor  $i, j \in \{0, 1, \dots, N\}$

$$k_{i,j} = k(x_{i,j}) \quad f_{i,j} = f(x_{i,j}). \quad (2.2)$$

We noteren verder voor  $l, m \in \{0, 1, \dots, \}$

$$u^{(l,m)}(x) = \frac{\partial^{l+m} u}{\partial x^l \partial y^m}(x). \quad (2.3)$$

We kunnen (1.25) en (1.26) uit het eindimensionale geval herschrijven met partiële afgeleiden. Neem  $j \in \{0, 1, \dots, N\}$  vast. Voor  $i \in \{0, 1, \dots, N - 1\}$

---

<sup>1</sup>Het is niet moeilijk om te veralgemenen naar een hoger aantal dimensies en naar balkvormige domeinen van de vorm  $\Omega = [0, L_1] \times [0, L_2] \times \dots \times [0, L_n]$ . Om de notatie in dit deelhoofdstuk niet te zwaar te maken, maken we de afleiding voor  $n = 2$

hebben we

$$u(x_{i+1,j}) = u(x_{i,j} + (h, 0)) \quad (2.4)$$

$$= u(x_{i,j}) + hu^{(1,0)}(x_{i,j}) + \frac{h^2}{2}u^{(2,0)}(x_{i,j}) \quad (2.5)$$

$$+ \frac{h^3}{6}u^{(3,0)}(x_{i,j}) + \frac{h^4}{24}u^{(4,0)}(x_{i,j} + \epsilon_{i,j,1}h(1, 0)) \quad (2.6)$$

en voor  $i \in \{1, 2, \dots, N\}$  geldt:

$$u(x_{i-1,j}) = u(x_{i,j} - (h, 0)) \quad (2.7)$$

$$= u(x_{i,j}) - hu^{(1,0)}(x_{i,j}) + \frac{h^2}{2}u^{(2,0)}(x_{i,j}) \quad (2.8)$$

$$- \frac{h^3}{6}u^{(3,0)}(x_{i,j}) + \frac{h^4}{24}u^{(4,0)}(x_{i,j} - \epsilon_{i,j,2}h(1, 0)). \quad (2.9)$$

Beide vergelijkingen optellen (voor  $i \in \{1, \dots, N-1\}$ ) geeft

$$u(x_{i+1,j}) + u(x_{i-1,j}) = 2u(x_{i,j}) + h^2u^{(2,0)}(x_{i,j}) + \mathcal{O}(h^4) \quad (2.10)$$

of door te herschikken

$$u^{(2,0)}(x_{i,j}) = \frac{u(x_{i+1,j}) - 2u(x_{i,j}) + u(x_{i-1,j}))}{h^2} + \mathcal{O}(h^2). \quad (2.11)$$

Op een analoge manier bekommen we voor  $i \in \{0, 1, \dots, N\}$ ,  $j \in \{1, \dots, N-1\}$  tot op eerste orde

$$u^{(0,2)}(x_{i,j}) = \frac{u(x_{i,j+1}) - 2u(x_{i,j}) + u(x_{i,j-1}))}{h^2} + \mathcal{O}(h^2) \quad (2.12)$$

We willen nu uitdrukken dat  $u(x, y)$  een oplossing is van de Helmholtzvergelijking, dat wil zeggen

$$-\frac{\partial^2 u}{\partial x^2}(x, y) - \frac{\partial^2 u}{\partial y^2}(x, y) - k(x, y)^2 u(x, y) = f(x, y) \quad (2.13)$$

in alle punten  $(x, y) \in ]0, L[^2$ . Door deze vergelijking te bekijken in de punten  $x_{i,j}$  voor  $i, j \in \{1, \dots, N-1\}$  en de voorgaande formules voor  $u^{(2,0)}(x_{i,j})$  en  $u^{(0,2)}(x_{i,j})$  te beschouwen, vinden we

$$f_{i,j} = -k_{i,j}^2 u(x_{i,j}) - \frac{u(x_{i+1,j}) - 2u(x_{i,j}) + u(x_{i-1,j}))}{h^2} \quad (2.14)$$

$$- \frac{u(x_{i,j+1}) - 2u(x_{i,j}) + u(x_{i,j-1}))}{h^2} + \mathcal{O}(h^2) \quad (2.15)$$

Benaderen we nu weer  $u(x_{i,j})$  door  $u_{i,j}$  voor  $i, j \in \{0, \dots, N\}$ , dan kunnen we stellen voor  $i, j \in \{1, \dots, N-1\}$

$$f_{i,j} = -k_{i,j}^2 u_{i,j} - \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} - \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{h^2}. \quad (2.16)$$

Herschikken levert ons

$$\frac{u_{i-1,j}}{h^2} + \frac{u_{i,j-1}}{h^2} + \frac{u_{i+1,j}}{h^2} + \frac{u_{i,j+1}}{h^2} + u_{i,j} \left( k_{i,j}^2 - \frac{4}{h^2} \right) = -f_{i,j} \quad (2.17)$$

en dit voor  $i, j \in \{1, \dots, N-1\}$ . We vinden aldus  $(N-1)^2$  vergelijkingen in de  $(N+1)^2$  veranderlijken  $(u_{i,j})_{i,j=0}^N$ . Er zijn nog randvoorwaarden nodig om de overige  $(N+1)^2 - (N-1)^2 = 4N$  lineaire vergelijkingen te bekommen, benodigd om tot een unieke oplossing te komen.

### Sommerfeld randvoorwaarden

Het is mogelijk om de Sommerfeld randvoorwaarden toe te passen in ieder punt op de rand van het domein. We bespreken dit slechts kort, aangezien zal blijken dat dit niet de ideale randvoorwaarden meer zijn in het tweedimensionale geval.

Noemen we  $A$  de coëfficiëntenmatrix voor Sommerfeld randvoorwaarden uit sectie 1.4, dan kan men aantonen dat de matrix voor de eindige differentiemethode in twee dimensies (en voor Sommerfeld randvoorwaarden) gegeven wordt door

$$I_{N+1} \otimes A + A \otimes I_{N+1}, \quad (2.18)$$

waar  $I_{N+1}$  de eenheidsmatrix met  $N+1$  rijen en kolommen voorstelt en  $\otimes$  het Kroneckerproduct is. Hierbij moet worden opgemerkt dat we de  $u_{i,j}$  (net als de  $f_{i,j}$ ) in één enkele vector hebben moeten steken, willen we het stelsel kunnen oplossen. We gebruiken hierbij de lexifcografische ordening:

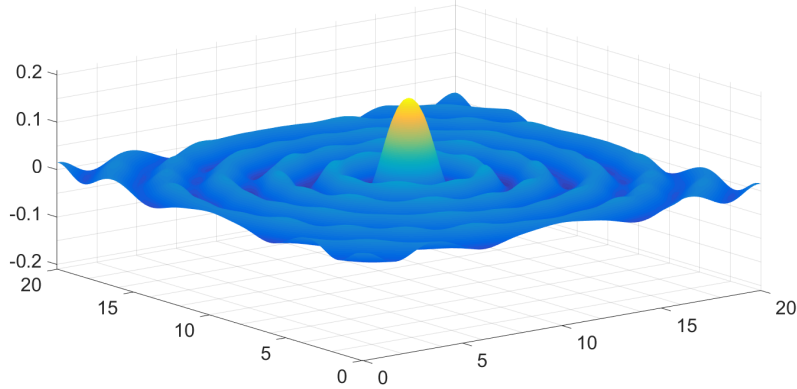
$$u_{0,0} \ u_{0,1} \ \dots \ u_{0,N+1} \ u_{1,0} \ \dots \ u_{N+1,N} \ u_{N+1,N+1}. \quad (2.19)$$

Ter illustratie maken we een 2D-plot van een oplossing van de golfvergelijking met Sommerfeld randvoorwaarden. De gekozen bron is

$$f(x, y) = e^{-(x-10)^2 - (y-10)^2}, \quad (2.20)$$

die radiaal symmetrisch is rond het middelpunt van het domein,  $(10, 10)$ . We verwachten dus een oplossing die deze radiale symmetrie erft. Op de plot zien we echter kleine afwijkingen van deze symmetrie: de golven hebben 'bulten'. De intuïtieve verklaring hiervoor is dat de Sommerfeld randvoorwaarden enkel de uitgaande golfvoorwaarden opleggen in de richting loodrecht op de rand van het domein, terwijl dit niet per se de richting is in dewelke de golven het domein verlaten. Voor een technischere verklaring over waarom deze randvoorwaarden niet het gewenste resultaat opleveren, zie bijvoorbeeld [4, p. 16].

De manier waarop dit probleem wordt aangepakt, is door zogenaamde *absorbing boundary conditions*, waarbij het oorspronkelijke domein  $\Omega$  een stukje wordt uitgebreid naar  $\tilde{\Omega}$  en de golven op dit uitgebreide domein geleidelijk gedempt worden. Een voorbeeld van een techniek die dit resultaat realiseert, is *Exterior Complex Scaling (ECS)*. Voor een uitwerking hiervan, zie [4, p. 17 e.v.]



Figuur 2.1: 2D oplossing met Sommerfeld randvoorwaarden

## 2.2 Tijdsafhankelijke Helmholtzvergelijking

Een tweede veralgemening is om te vertrekken van het eindimensionale tijdsafhankelijke probleem (1.10). We willen dan de evolutie van een golf in de tijd modelleren.

De eerste stap is om de tweede-orde differentiaalvergelijking (1.10) om te zetten naar een stelsel eerste-orde differentiaalvergelijkingen. Beschouw daartoe het stelsel

$$\begin{cases} \frac{\partial u}{\partial t} = c \frac{\partial v}{\partial x} + c^2 G(x, t) \\ \frac{\partial v}{\partial t} = c \frac{\partial u}{\partial x} \end{cases} \quad (2.21)$$

waarbij  $G(x, t)$  een primitieve is van  $F(x, t)$  met betrekking tot  $t$ , dat wil zeggen een functie zodanig dat  $\frac{\partial}{\partial t} G(x, t) = F(x, t)$ . We beweren dat het voldoende is dit stelsel op te lossen, om 1.10 op te lossen.

Stel eerst  $(u, v)$  een oplossing van (2.21). Onder enkele gladheidsaannames op  $v$  vinden we dan

$$\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = \frac{1}{c^2} \frac{\partial}{\partial t} \left( c \frac{\partial v}{\partial x} + c^2 G(x, t) \right) - \frac{\partial}{\partial x} \left( \frac{1}{c} \frac{\partial v}{\partial t} \right) \quad (2.22)$$

$$= \frac{1}{c} \frac{\partial^2 v}{\partial t \partial x} + F(x, t) - \frac{1}{c} \frac{\partial^2 v}{\partial x \partial t} \quad (2.23)$$

$$= F(x, t) \quad (2.24)$$

zodat  $u$  een oplossing is van (1.10). Omgekeerd, onderstel dat  $u$  een oplossing is van (1.10). We willen tonen dat er dan een functie  $v$  bestaat zodat  $(u, v)$  een oplossing is van (2.21), en dat deze  $v$  op een constante term na uniek bepaald is.

Beschouw hiertoe eerst een willekeurige primitieve  $H(x, t)$  van  $\frac{\partial u}{\partial x}$  ten opzichte van  $t$ . Definieer

$$v = cH(x, t) + C(x) \quad (2.25)$$

voor een nog later te bepalen functie  $C(x)$ , dan geldt alvast

$$\frac{\partial v}{\partial t} = c \frac{\partial u}{\partial x}. \quad (2.26)$$

Verder hebben we achtereenvolgens

$$\frac{\partial^2 u}{\partial t^2} = c^2 \left( F(x, t) + \frac{\partial^2 u}{\partial x^2} \right) \quad (2.27)$$

$$\frac{\partial u}{\partial t} = c^2 G(x, t) + c^2 \frac{\partial}{\partial x} H(x, t) + D(x) \quad (2.28)$$

voor een zekere functie  $D(x)$ .

$$\frac{\partial u}{\partial t} = c^2 G(x, t) + c^2 \frac{\partial}{\partial x} \left( \frac{v}{c} - \frac{C(x)}{c} \right) + D(x) \quad (2.29)$$

$$= c^2 G(x, t) + c \frac{\partial v}{\partial x} - c \frac{\partial C}{\partial x} + D(x) \quad (2.30)$$

Als we dus  $C(x) = \frac{1}{c} E(x)$  stellen, met  $E(x)$  een willekeurige primitieve van  $D(x)$  met betrekking tot  $x$ , vinden we dat, met de gedefinieerde  $v$ ,  $(u, v)$  een oplossing is van (2.21).

Tenslotte, stel dat  $(u, v)$  en  $(u, \tilde{v})$  beide oplossingen zijn van (2.21) (zelfde  $u$ ). We vinden dan achtereenvolgens uit de eerste vergelijking in (2.21)

$$\frac{\partial v}{\partial x} = \frac{1}{c} \left( \frac{\partial u}{\partial x} - c^2 G(x, t) \right) = \frac{\partial \tilde{v}}{\partial x} \quad (2.31)$$

$$\frac{\partial}{\partial x} (v - \tilde{v}) = \frac{\partial v}{\partial x} - \frac{\partial \tilde{v}}{\partial x} = 0 \quad (2.32)$$

zodat  $v - \tilde{v}$  niet van  $x$  afhangt. Analoog vinden we uit de tweede vergelijking in (2.21) dat  $v - \tilde{v}$  niet van  $t$  afhangt, zodat  $v - \tilde{v}$  constant is, wat we wilden bewijzen.

We houden ons dus vanaf nu met het stelsel (2.21) bezig. We kunnen dit in matrixvorm herschrijven als

$$\begin{bmatrix} \frac{\partial u}{\partial t} \\ \frac{\partial v}{\partial t} \end{bmatrix} = c \begin{bmatrix} 0 & \frac{\partial}{\partial x} \\ \frac{\partial}{\partial x} & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} + c^2 \begin{bmatrix} G(x, t) \\ 0 \end{bmatrix}. \quad (2.33)$$

De volgende vraag is hoe we deze vergelijking kunnen herschrijven met behulp van eindige differenties, zodat we ze benaderend kunnen uitrekenen wanneer we  $u$  en  $v$  in  $x_0, \dots, x_N$  kennen. We gebruiken hiervoor opnieuw de eerder afgeleide centrale differenties, waarbij we in het vervolg  $u'_i$  voor de benadering van  $\frac{\partial u}{\partial x}(x_i)$  en  $\dot{u}_i$  voor de benadering van  $\frac{\partial u}{\partial t}(x_i)$  noteren, beide op een vast moment  $t$ .

$$u'_i \approx \frac{u_{i+1} - u_{i-1}}{2h} \quad v'_i \approx \frac{v_{i+1} - v_{i-1}}{2h} \quad (2.34)$$

Dit geeft dan

$$\begin{bmatrix} \dot{u}_1 \\ \vdots \\ \dot{u}_{N-1} \\ \dot{v}_1 \\ \vdots \\ \dot{v}_{N-1} \end{bmatrix} = \begin{bmatrix} 0 & B \\ B & 0 \end{bmatrix} \begin{bmatrix} u_0 \\ \vdots \\ u_N \\ v_0 \\ \vdots \\ v_N \end{bmatrix} + c^2 \begin{bmatrix} G(x_1, t) \\ \vdots \\ G(x_{N-1}, t) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (2.35)$$

$$\text{met } B = \frac{1}{2h} \begin{bmatrix} 1 & 0 & -1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \ddots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ddots & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 & -1 \end{bmatrix}. \quad (2.36)$$

We bekommen aldus  $\dot{u}_i$  en  $\dot{v}_i$  ( $i \in \{1, \dots, N-1\}$ ) gegeven  $u_i$  en  $v_i$  ( $i \in \{0, \dots, N\}$ ) op een bepaald tijdstip  $t$ . Nu zullen we aantonen dat we deze kunnen gebruiken om de  $u_i$  en  $v_i$  te bepalen op een tijdstip  $t + \Delta t$ . Noteer

$$A = \begin{bmatrix} 0 & B \\ B & 0 \end{bmatrix} \quad U(t) = \begin{bmatrix} u_1(t) \\ \vdots \\ u_N(t) \\ v_1(t) \\ \vdots \\ v_N(t) \end{bmatrix}. \quad (2.37)$$

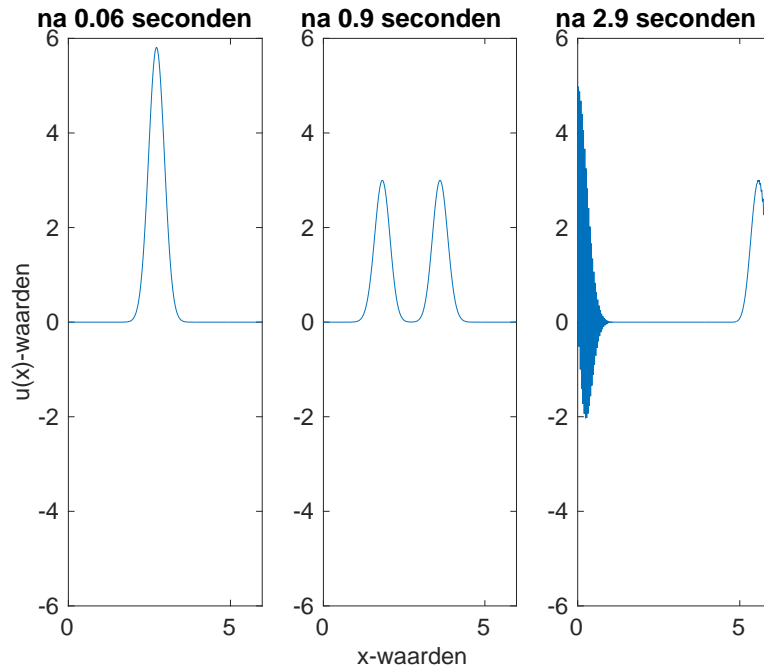
We willen kunnen zeggen dat  $\frac{\partial U}{\partial t}(t) = AU$ . Het probleem is echter dat de matrix  $A$  enkel voor  $i \in \{1, \dots, N-1\}$  een uitdrukking geeft voor  $\dot{u}_i$  en  $\dot{v}_i$ . We missen dus nog vier randvoorwaarden om  $\dot{u}_0, \dot{u}_N, \dot{v}_0$  en  $\dot{v}_N$  te kunnen bepalen. Onderstel even dat we  $A$  kunnen uitbreiden met vier vergelijkingen zodanig dat inderdaad  $\frac{\partial U}{\partial t}(t) = AU$ . Er geldt dan voor een klein getal  $\Delta t > 0$

$$\begin{aligned} U(t + \Delta t) - U(t) &= \int_t^{t+\Delta t} \frac{\partial U}{\partial t}(t) dt = \int_t^{t+\Delta t} AU(t) = A \int_t^{t+\Delta t} U(t) \\ &\approx A \left( \frac{U(t + \Delta t) + U(t)}{2} \right) \Delta t \end{aligned}$$

waarbij we de integraal benaderden met behulp van de trapeziumregel. We vormen dit om naar een stelsel naar  $U(t + \Delta t)$ :

$$\left( I_{2N} - \frac{\Delta t}{2} A \right) U(t + \Delta t) \approx \left( I_{2N} + \frac{\Delta t}{2} A \right) U(t). \quad (2.38)$$





Figuur 2.2: Een oplossing van de tijdsafhankelijke Helmholtzvergelijking, waarin een golf in het midden van het domein wordt ingezet. De oplossing is bruikbaar totdat de golf de randen van het domein bereikt.

Dit laat ons uiteindelijk toe om  $U(t + \Delta t)$  benaderend te berekenen als we  $U(t)$  kennen.

Het ontbreken van goede randvoorwaarden vormt echter een groot probleem. Een naïeve vertaling van Dirichlet of Sommerfeld randvoorwaarden, bijvoorbeeld door  $\dot{u}_0 = \dot{v}_0 = \dot{u}_N = \dot{v}_N = 0$  te stellen, blijkt geen goede resultaten op te leveren, zoals fig. 2.2 toont.

## Hoofdstuk 3

# Iteratieve methodes

In dit hoofdstuk bestuderen we hoe de stelsels met grote, ijle coëfficiëntenmatrices uit de vorige hoofdstukken kunnen worden opgelost met behulp van iteratieve methodes. We vergelijken hierbij verschillende methodes op het vlak van reken-tijd en nauwkeurigheid en doen dit zowel op basis van theoretische eigenschappen, als met behulp van numerieke simulaties. We zullen meermaals verwijzen naar de cursus *Iteratieve methodes voor lineaire stelsels en eigenwaarden problemen*. [5]

### 3.1 Eigenwaarden van coëfficiëntenmatrix

Voor enkele iteratieve methoden is het wenselijk om een idee te hebben hoe de eigenwaarden van de coëfficiëntenmatrix (of een hieruit afgeleide iteratiematrix) ongeveer verdeeld liggen, of beter nog, om een expliciete uitdrukking te beko-men voor deze eigenwaarden. We kijken eerst naar een situatie waarin we de eigenwaarden analytisch kunnen uitrekenen.

Beschouw de  $m \times m$ -matrix

$$\hat{T} = \begin{bmatrix} 0 & 1 & & \\ 1 & 0 & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & 0 \end{bmatrix}. \quad (3.1)$$

We beweren dat haar eigenwaarden gegeven worden door

$$\hat{\lambda}_k = 2 \cos\left(\frac{k\pi}{m+1}\right) \quad k = 1, \dots, m \quad (3.2)$$

en, voor een eigenwaarde  $\hat{\lambda}_k$ , een bijbehorende eigenvector  $\hat{v} = (\hat{v}_{k,1}, \dots, \hat{v}_{k,m})^T$  gegeven wordt door

$$\hat{v}_{k,j} = \sin\left(\frac{kj\pi}{m+1}\right) \quad j = 1, \dots, m. \quad (3.3)$$

Om dit in te zien, stel

$$\hat{v}_{k,0} = \sin\left(\frac{0k\pi}{m+1}\right) = 0 \quad \hat{v}_{k,m+1} = \sin\left(\frac{(m+1)j\pi}{m+1}\right) = 0 \quad (3.4)$$

dan is voor iedere  $k, j = 1, \dots, m$

$$(\hat{T}\hat{v}_k)_j = \hat{v}_{k,j-1} + \hat{v}_{k,j+1} = \sin\left(\frac{k(j-1)\pi}{m+1}\right) + \sin\left(\frac{k(j+1)\pi}{m+1}\right) \quad (3.5)$$

$$= \sin\left(\frac{kj\pi}{m+1}\right) \cos\left(\frac{k\pi}{m+1}\right) - \sin\left(\frac{k\pi}{m+1}\right) \cos\left(\frac{kj\pi}{m+1}\right) \quad (3.6)$$

$$+ \sin\left(\frac{kj\pi}{m+1}\right) \cos\left(\frac{k\pi}{m+1}\right) + \sin\left(\frac{k\pi}{m+1}\right) \cos\left(\frac{kj\pi}{m+1}\right) \quad (3.7)$$

$$= 2 \sin\left(\frac{kj\pi}{m+1}\right) \cos\left(\frac{k\pi}{m+1}\right) = \hat{\lambda}_k \hat{v}_{k,j}. \quad (3.8)$$

Beschouwen we dan algemener een matrix

$$T(a, b) = aI + b\hat{T} = \begin{bmatrix} a & b & & \\ b & a & \ddots & \\ & \ddots & \ddots & b \\ & & b & a \end{bmatrix}. \quad (3.9)$$

voor zekere  $a, b \in \mathbb{C}$ , dan worden haar eigenwaarden (wegens [5, Lemma 18]) gegeven door

$$\lambda_k = a + b\hat{\lambda}_k = a + b2 \cos\left(\frac{k\pi}{m+1}\right) \quad k = 1, \dots, m. \quad (3.10)$$

In de meeste gevallen kunnen we de iteratiematrix niet exact in bovenstaande vorm krijgen, maar ze kan wel een indicatie geven van wat te verwachten valt van de eigenwaarden.

## 3.2 Stationaire methoden

Stationaire methoden zijn van de vorm

$$x^{(k+1)} = Mx^{(k)} + c \quad (3.11)$$

waarbij  $(x^{(k)})_{k \in \mathbb{N}}$  een rij vectoren in  $\mathbb{R}^m$  is,  $M$  een  $m \times m$ -matrix (de iteratiematrix genaamd) en  $c$  een vector in  $\mathbb{R}^m$ . De methode convergeert naar het fixpunt  $x^* = (I - M)^{-1}c$  voor eender welke beginvector  $x^{(0)}$  als en slechts als  $\rho(M) < 1$ , waarbij

$$\rho(M) = \max_{\lambda \in \sigma(M)} |\lambda| \quad (3.12)$$

de spectrale straal van  $M$  is ( $\sigma(M)$  is de verzameling eigenwaarden van  $M$ ). [5, Stelling 11] We passen enkele van deze stationaire methoden toe op een paar van onze stelsels.

## De gewogen methode van Jacobi

Beschouw een lineair, vierkant stelsel van de vorm  $Ax = b$  ( $A \in \mathbb{R}^{m \times m}$  regulier,  $x, y \in \mathbb{R}^m$ ). Definieer de matrix  $D$  als de  $m \times m$  diagonaalmatrix met als hoofddiagonaal juist de hoofddiagonaal van  $A$ . De gewogen methode van Jacobi met parameter  $\omega \in \mathbb{R}$  wordt dan gegeven door

$$x^{(k+1)} = [I - \omega D^{-1}A]x^{(k)} + \omega D^{-1}b. \quad (3.13)$$

Met de notatie van hiervoor hebben we dan  $M_\omega = I - \omega D^{-1}A$  en  $c_\omega = \omega D^{-1}b$ . In de onderstelling dat  $\rho(M_\omega) < 1$ , convergeert deze methode dan naar

$$(I - M_\omega)^{-1}c = (\omega D^{-1}A)^{-1}\omega D^{-1}b = A^{-1}DD^{-1}b = A^{-1}b \quad (3.14)$$

hetgeen de oplossing is van ons oorspronkelijke stelsel.

De methode zal niet convergeren indien  $D^{-1}A$  indefiniet is. Indien de matrix  $D^{-1}A$  symmetrisch en positief definiet is, dan convergeert de methode indien  $\omega$  juist gekozen wordt; de keuze voor  $\omega$  die  $\rho(M_\omega)$  minimaliseert, is

$$\omega^* = \frac{2}{\lambda_{\min} + \lambda_{\max}} \quad (3.15)$$

en de dan bereikte spectrale straal is

$$\rho^* = \rho(M_{\omega^*}) = \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\min} + \lambda_{\max}} < 1 \quad (3.16)$$

waarbij  $\lambda_{\min}, \lambda_{\max} > 0$  respectievelijk de kleinste en grootste eigenwaarden van  $D^{-1}A$  zijn. [5, p. 26-27]

De uitdaging is natuurlijk het berekenen van  $\lambda_{\min}$  en  $\lambda_{\max}$ . Dit lukt in het geval van de eendimensionale, tijdsonafhankelijke Helmholtzvergelijking met een plaatsonafhankelijk golfgetal en Dirichlet randvoorwaarden. Dit is, zoals eerder vermeld, niet het probleem dat in het praktijk het vaakst moet worden opgelost, maar haar oplossing zal ons een idee geven van de convergentie-eigenschappen van de gewogen Jacobi methode. De  $(N+1) \times (N+1)$  matrix  $A$  is in dit geval (zie sectie 1.4):

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ \frac{1}{h^2} & \frac{-2}{h^2} + k^2 & \frac{1}{h^2} & 0 & \dots & 0 \\ 0 & \frac{1}{h^2} & \frac{-2}{h^2} + k^2 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & \frac{1}{h^2} & 0 \\ \vdots & \vdots & \ddots & \frac{1}{h^2} & \frac{-2}{h^2} + k^2 & \frac{1}{h^2} \\ 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix} \quad (3.17)$$

Deze matrix staat nog niet in de vorm (3.9), maar we kunnen dit wel bekomen door op te merken dat de eerste en laatste vergelijkingen niets anders dan  $x_0 = x_N = 0$  stellen, zodat we deze gewoon uit het stelsel mogen weglaten. We

bekomen een stelsel in de  $N - 1$  veranderlijken  $x_1, \dots, x_{N-1}$ , bepaald door coëfficiëntenmatrix

$$\tilde{A} = \begin{bmatrix} \frac{-2}{h^2} + k^2 & \frac{1}{h^2} & & & \\ \frac{1}{h^2} & \frac{-2}{h^2} + k^2 & \ddots & & \\ & & \ddots & \ddots & \\ & & & \frac{1}{h^2} & \frac{-2}{h^2} + k^2 \end{bmatrix} \quad (3.18)$$

De matrix  $\tilde{D}^{-1}\tilde{A}$  wordt dan bepaald door in vorige matrix iedere rij te delen door het element dat op de hoofdiagonaal in die rij staat:

$$\tilde{D}^{-1}\tilde{A} = \begin{bmatrix} 1 & \frac{-1}{2-h^2k^2} & & & \\ \frac{-1}{2-h^2k^2} & 1 & \ddots & & \\ & & \ddots & \ddots & \\ & & & \frac{-1}{2-h^2k^2} & 1 \end{bmatrix}. \quad (3.19)$$

Gezien deze matrix in de vorm (3.9) staat, worden haar eigenwaarden wegens (3.10) gegeven door

$$\lambda_j = 1 + 2 \frac{\cos\left(\frac{j\pi}{N}\right)}{2 - k^2h^2} \quad j = 1, \dots, N - 1 \quad (3.20)$$

De matrix is positief definitief indien alle eigenwaarden  $\lambda_j$  strikt positief zijn. Voor voldoende grote  $N$  zal  $2 - k^2h^2 > 0$ , zodat kleinste en grootste eigenwaarden gegeven worden door

$$\lambda_{\min} = 1 - 2 \frac{\cos\left(\frac{\pi}{N}\right)}{2 - k^2h^2} \quad \lambda_{\max} = 1 + 2 \frac{\cos\left(\frac{\pi}{N}\right)}{2 - k^2h^2} \quad (3.21)$$

Gebruik makend van de Taylorreeks van  $\cos(x)$  vinden we

$$(2 - k^2h^2)\lambda_{\min} = 2 - h^2k^2 - 2 \cos\left(\frac{\pi}{N}\right) = -h^2k^2 + \left(\frac{\pi}{N}\right)^2 + \mathcal{O}(N^{-4}) \quad (3.22)$$

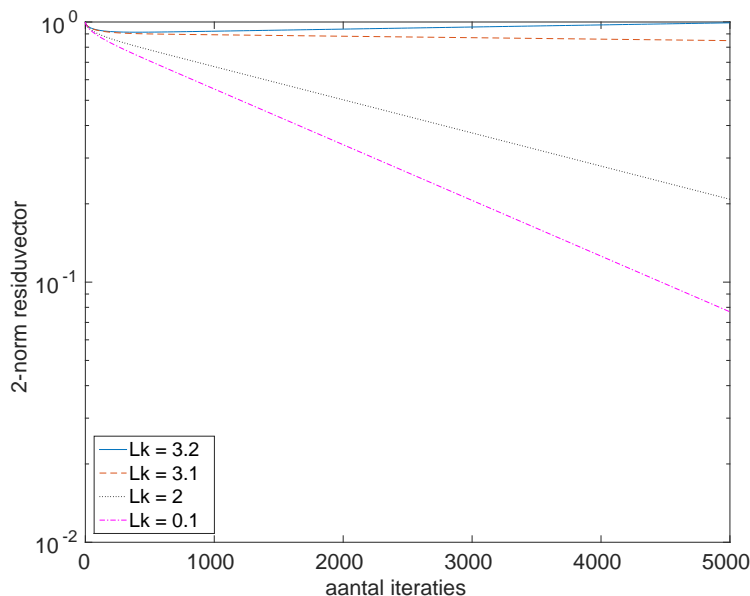
zodat  $\lambda_{\min} > 0$  als en slechts als

$$-L^2k^2 + \pi^2 + \mathcal{O}(N^{-2}) > 0 \Leftrightarrow 1 < \left(\frac{\pi}{Lk}\right)^2 + \mathcal{O}(N^{-2}). \quad (3.23)$$

Willen we dat, voor vaste  $k$  en  $L$ , de matrix positief definitief is voor willekeurig (maar voldoende) grote  $N$ , dan moeten we dus eisen dat  $Lk < \pi$ . In dat geval hebben we uit (3.15) en (3.16)

$$\omega^* = \frac{2}{\lambda_{\min} + \lambda_{\max}} = 1 \quad (3.24)$$

$$\rho^* = \frac{\lambda_{\min} - \lambda_{\max}}{\lambda_{\min} + \lambda_{\max}} = \frac{2 \cos\left(\frac{\pi}{N}\right)}{2 - k^2h^2}. \quad (3.25)$$



Figuur 3.1: Uitvoering van de Jacobi iteratie voor  $Lk = 0.1$ ,  $Lk = 2$ ,  $Lk = 3.1$  en  $Lk = 3.2$  (en telkens  $N = 100$ ). We zien trage convergentie in de eerste drie gevallen en divergentie in het laatste geval. Rekeningtijd is telkens ongeveer 0.06 seconden<sup>1</sup>. De spectrale straal van de iteratiematrix is achtereenvolgens 0.99951, 0.99971, 0.99999 en 1.00002.

We vinden

$$\rho^* = 1 - \frac{\pi^2}{2N^2} + \frac{k^2 L^2}{2N^2} + \mathcal{O}(N^{-4}) = 1 + \frac{-\pi^2 + k^2 L^2}{2N^2} + \mathcal{O}(N^{-4}) \quad (3.26)$$

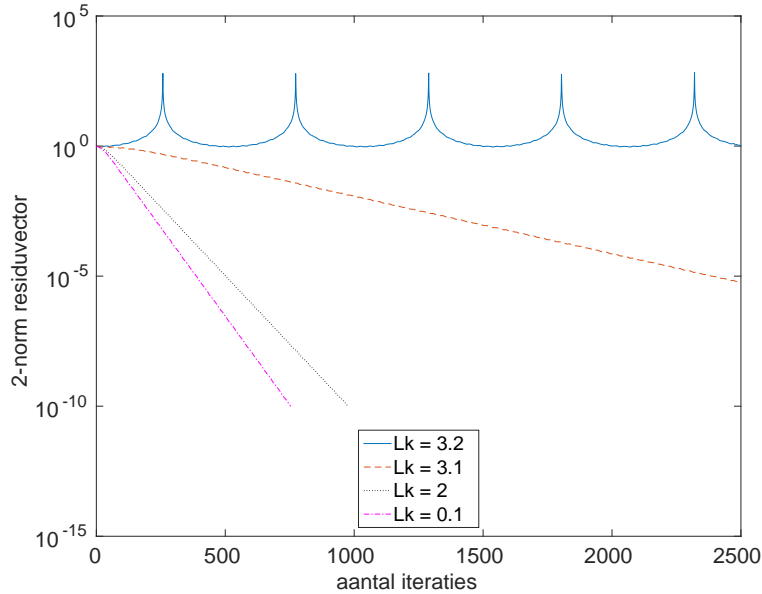
en besluiten dat de gewogen methode van Jacobi enkel een degelijke convergentiesnelheid gaat hebben indien  $N$  niet te groot is en  $Lk < \pi$ . In de praktijk is ze dus niet erg bruikbaar.

## Chebyshev versnelling

Indien we de spectrale straal van de iteratiematrix van een stationaire methode expliciet kennen (of, algemener, indien we een betere bovengrens op de straal kennen dan 1), dan kunnen we deze gebruiken om de convergentiesnelheid van de methode substantieel te verbeteren. We verwijzen voor de theorie hierrond naar [5, hoofdstuk 3].

Bekijken we opnieuw het  $1D$  probleem met Dirichlet randvoorwaarden, dan hadden we eerder afgeleid dat de spectrale straal gegeven wordt door (3.25). We

<sup>1</sup>Rekeningtijd werd altijd op (dezelfde) Linux computer in de Universiteit Antwerpen gemeten.



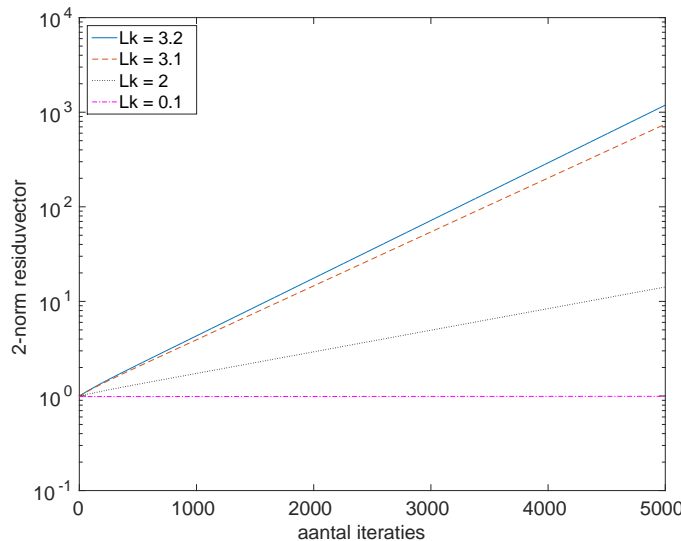
Figuur 3.2: Chebyshev versnelling van een Jacobi iteratie voor  $Lk = 0.1$ ,  $Lk = 2$ ,  $Lk = 3.1$  en  $Lk = 3.2$  (en telkens  $N = 100$ ). Rekeningtijd is 0.03 seconden voor de eerste twee en 0.08 seconden voor de derde.

zien dat de convergentie in de gevallen  $Lk = 0.1$ ,  $Lk = 2$  en  $Lk = 3.1$  gevoelig verbetert ten opzichte van de overeenkomstige situaties in de primaire gewogen Jacobi iteratie. In deze gevallen, zeker voor de eerste twee, is de methode aldus bruikbaar. Voor grotere  $N$  zal de convergentie echter steeds trager worden.

### Toepassing op Sommerfeld randvoorwaarden

Gezien voor grote  $N$  de coëfficiëntenmatrix bij het eendimensionale probleem met Dirichlet randvoorwaarden weinig verschilt van die met Sommerfeld randvoorwaarden, zou men naïef kunnen verwachten dat ook het convergentiegedrag van de daarop toegepaste stationaire methode weinig zal veranderen. Bekijken we echter opnieuw hetzelfde probleem als bij de vorige twee methoden, dan zien we dat er in geen van de vier gevallen nog convergentie is.

Omdat de convergentie-eigenschappen van stationaire methoden zeer gevoelig zijn voor kleine veranderingen in het spectrum van de iteratiematrix, zijn ze enkel bruikbaar indien we veel informatie beschikken over dit spectrum. Over het algemeen zal dit niet zo zijn, niet enkel omdat we niet met Dirichlet randvoorwaarden werken, maar ook omdat het golfgetal variabel kan zijn. We besluiten dat deze methoden in het algemeen niet bruikbaar zijn voor ons probleem.



Figuur 3.3: Jacobi iteratie toegepast op matrix met Sommerfeld randvoorwaarden,  $N = 51$  en  $Lk = 0.1$ ,  $Lk = 2$ ,  $Lk = 3.1$  en  $Lk = 3.2$ . De spectrale straal van de iteratiematrix is achtereenvolgens 1.00000, 1.00053, 1.00131 en 1.00141.

### 3.3 GMRES

GMRES (*Generalized Minimal Residual*) is een Krylov-deelruimtemethode voor niet-symmetrische matrices (zie [5, hoofdstuk 4]). De  $k$ -de Krylovruimte ten opzichte van een  $n \times n$  matrix en een vector  $v \in \mathbb{R}^n$  wordt gedefinieerd als

$$\mathcal{K}_k(A, v) = \text{span}\{v, Av, A^2v, \dots, A^{k-1}v\}. \quad (3.27)$$

De reeks benaderingen voor de oplossing van een stelsel  $Ax = b$  wordt dan gegeven door een rij vectoren  $(x^{(k)})_{k \in \mathbb{N}}$ , die men bepaalt door te eisen dat

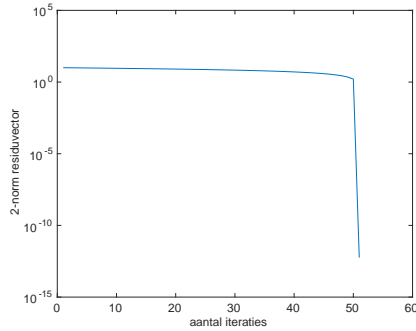
$$x^{(k)} \in x^{(0)} + \mathcal{K}_k(A, r^{(0)}) \quad \text{en} \quad \|r^{(k)}\|_2 = \min_{v \in x^{(0)} + \mathcal{K}_k(A, r^{(0)})} \|b - Av\|_2. \quad (3.28)$$

waarbij  $x^{(0)}$  gegeven is en voor iedere  $k \in \mathbb{N}$  we  $r^{(k)} = b - Ax^{(k)}$  stellen.  $r^{(k)}$  noemen we het  $k$ -de residu, wat de naam van de methode verklaart: in iedere stap wordt de norm van het residu over een bepaalde (steeds groter wordende) deelruimte van  $\mathbb{R}^n$  geminimaliseerd.

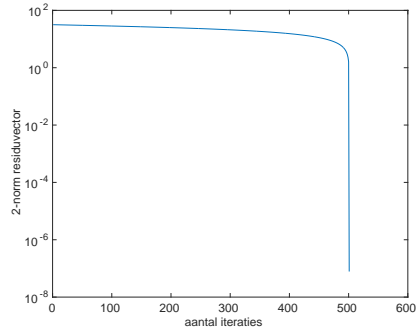
Het is duidelijk dat, als  $r^{(k)} = 0$ ,  $x^{(k)}$  de gezochte oplossing is van het stelsel. We kunnen zelfs meer zeggen: stel dat  $x^*$  de exacte oplossing is van het stelsel  $Ax = b$ , dan volgt dat  $r^{(k)} = A(x^* - x^{(k)})$  en aldus

$$\|x^* - x^{(k)}\|_2 = \|A^{-1}r^{(k)}\|_2 \leq \|A^{-1}\|_2 \|r^{(k)}\|_2 \quad (3.29)$$

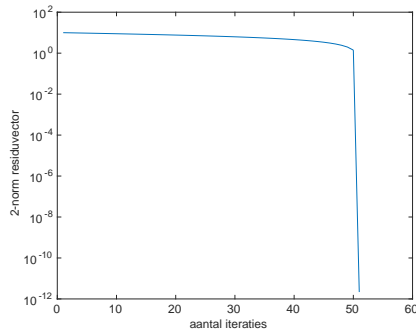




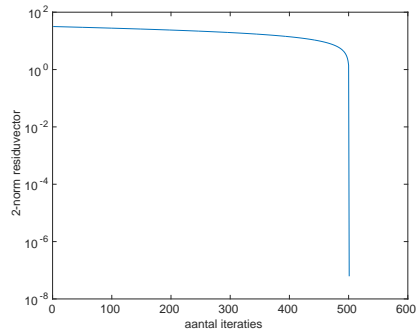
(a) Dirichlet,  $N = 100$ , 0.01 sec.



(b) Dirichlet,  $N = 1000$ , 1.36 sec.



(c) Sommerfeld,  $N = 100$ , 0.02 sec.



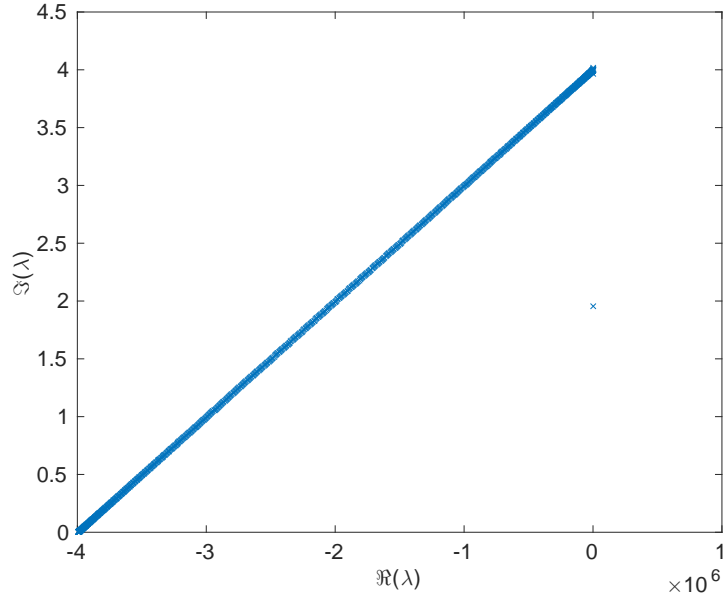
(d) Sommerfeld,  $N = 1000$ , 2.56 sec.

Figuur 3.4: Convergentiegedrag en rekestijd van GMRES toegepast op verschillende problemen. Telkens zien we dat er pas redelijke nauwkeurigheid wordt bereikt na  $N/2$  iteraties.

zodat  $x^{(k)} \rightarrow x^*$  indien  $r^{(k)} \rightarrow 0$ . Men kan tonen dat, voor een willekeurige beginvector  $x^{(0)}$ , inderdaad  $(\|r^{(k)}\|_2)_{k \in \mathbb{N}}$  naar 0 convergeert (en dan noodzakelijk naar 0 daalt, wegens de definitie van  $x^{(k)}$ ). De fout  $\|x^* - x^{(k)}\|$  zal ook een dalende trend vertonen (wegens bovenstaande afchatting), maar deze daling hoeft niet strikt te zijn.

Een van de voornaamste nadelen van GMRES is dat het geheugengebruik met  $n$  scalairen toeneemt met iedere iteratie. Het valt daarom te hopen dat de methode, omdat ze gebaseerd is op een minimalisatieprobleem, snel naar de oplossing zal convergeren. Passen we haar echter toe op de Helmholtzvergelijking (met  $L = 1$  en  $f(x) = k(x) = 1$  voor alle  $x$ ), dan zien we dat er zeer trage convergentie is, zowel bij Dirichlet randvoorwaarden als met Sommerfeld randvoorwaarden.

We willen het slechte convergentiegedrag uit fig. 3.4 kunnen verklaren. De coëfficiëntenmatrix bepaalt in grote mate het convergentiegedrag. In het geval van Dirichlet randvoorwaarden wordt deze gegeven door de matrix  $\tilde{A}$  uit (3.18).



Figuur 3.5: Spectrum van de coëfficiëntenmatrix bij  $N = 1000$  en Sommerfeld randvoorwaarden

Deze matrix staat in de vorm (3.9) staat en we kennen dus haar eigenwaarden uit (3.10):

$$\lambda_j = \frac{-2}{h^2} + k^2 + 2 \cos\left(\frac{j\pi}{N}\right) / h^2 = k^2 - \frac{2}{h^2} \left(1 - \cos\left(\frac{j\pi}{N}\right)\right) \quad j = 1, \dots, N-1 \quad (3.30)$$

Dit betekent dat er eigenwaarden zijn die  $\mathcal{O}(N^2)$  zijn, terwijl er mogelijk ook eigenwaarden zeer dicht bij nul liggen. Dit tezamen suggereert een zeer groot conditiegetal; een numerieke berekening levert een conditiegetal van  $5.5935 \cdot 10^5$  voor  $N = 1000$  en Dirichlet randvoorwaarden. Met Sommerfeld randvoorwaarden is het conditiegetal bij  $N = 1000$  gelijk aan  $1.9980 \cdot 10^6$ . Hier ziet het spectrum er uit als in fig. 3.5.

Het conditiegetal is een sterke indicator voor de convergentiesnelheid van Krylov-methoden. Door het enorm grote conditiegetal zal de convergentie zeer traag zijn met GMRES en het geheugengebruik voor grotere problemen niet meer beheersbaar.

# Besluit en outlook

De Helmholtzvergelijking kan eenvoudig gediscretiseerd worden om tot een lineair stelsel te komen en deze benadering is consistent van orde twee. Gegeven de functies  $f(x)$  en  $k(x)$  en het domein, worden de grootste uitdagingen gevormd door het vinden van geschikte randvoorwaarden en het oplossen van het grote, ijle stelsel. Immers, waar we voor een  $1D$ -probleem een stelsel van  $N + 1$  vergelijkingen krijgen indien we de lengte  $L$  van het domein in  $N$  intervallen verdelen, hebben we voor dezelfde nauwkeurigheid  $\mathcal{O}(N^2)$  vergelijkingen nodig voor een  $2D$ -probleem met een domein  $[0, L]^2$  en  $\mathcal{O}(N^3)$  vergelijkingen voor het  $3D$ -probleem. Methodes vinden die dit soort problemen in een redelijke tijdsduur kunnen oplossen, is aangewezen.

Voor het probleem van de randvoorwaarden blijkt de oplossing van Sommerfeld goed in het eendimensionale en tijdsafhankelijke geval. In het tweedimensionale probleem zorgen deze randvoorwaarden echter al voor onwenselijke effecten. In dit geval kan ECS een uitweg bieden; deze wordt verkend in de bachelorproef van Teppo Mertens [2]. Het probleem met de tijdsafhankelijke vergelijking vereist nog meer onderzoek.

Iteratieve methodes lijken aangewezen om de optredende ijle, lineaire stelsels op te lossen. De eenvoudige stationaire methoden blijken echter een convergentiegedrag te hebben dat sterk afhangt van de specifieke instantie van de Helmholtzvergelijking (in het algemeen hoeft er geen convergentie te zijn) en de populaire Krylov-methode GMRES convergeert slechts zeer langzaam en na vele iteraties, zodat geen van beide in het algemeen bruikbaar zijn. Dit gedrag kunnen we verklaren aan de hand van het spectrum en het grote conditiegetal van de coëfficiëntenmatrix en we verwachten dus gelijkaardige effecten bij andere iteratieve methodes. Preconditioners zouden hier een mogelijke oplossing kunnen bieden.

# Bibliografie

- [1] David, M. *Gewone Differentiaalvergelijkingen*. cursus uit najaar 2014 te Universiteit Antwerpen.
- [2] Mertens, T. *Wifi signalen en golven*. bachelorproef uit 2016 te Universiteit Antwerpen.
- [3] Sommerfeld, A. vertaald door Straus, A.G. *Partial Differential Equations in Physics*. Lectures on Theoretical Physics Volume VI. New York: Academic Press; 1949.
- [4] Vanroose, W. en Cools, S. *Applications of Partial Differential Equations: The Helmholtz equation*. 2014.
- [5] Vanroose, W. *Iteratieve methodes voor lineaire stelsels en eigenwaarden problemen*. cursus uit voorjaar 2016 te Universiteit Antwerpen.