K O N I N K L I J K   N E D E R L A N D S
M E T E O R O L O G I S C H   I N S T I T U U T

D e   B i l t

WETENSCHAPPELIJK RAPPORT
W.R. 77-9

S.J. Bijlsma

Directe methoden voor het oplossen van elliptische partiële differentiaalvergelijkingen.

De Bilt, 1977.

Publikatienummer: K.N.M.I. W.R. 77-9 (M.B.W.)

U.D.C.: 517.9

Summary.

In this report we give a survey of direct methods for solving Poisson's equation on a rectangle, such as block methods, cyclic reduction methods, matrix decomposition methods and Fourier series methods.
An extension of these methods to more general domains has been given by Buzbee, Dorr, George and Golub (1971).

1. Inleiding.

In dit rapport behandelen we directe methoden voor het oplossen van stelsels lineaire vergelijkingen van de vorm

$$L_0 u = k \,, \tag{1.1}$$

afkomstig van een 5-punts differentieapproximatie van de Poisson vergelijking

$$\Delta U(x,y) = F(x,y) \,, \tag{1.2}$$

$$U = G(R) \quad \text{op de rand } R \ (R = \text{rechthoek}). \tag{1.3}$$

Hoewel het in principe altijd mogelijk is een matrix vergelijking van de vorm (1.1) te inverteren, zijn er pas vrij recent snelle methoden voor dit doel ontwikkeld. Deze directe methoden blijken zeer snel ten opzichte van iteratieve methoden, indien men over een slechte beginschatting van de oplossing van (1.1) beschikt, zodat we een groot aantal iteraties moeten toepassen om de beginfout te reduceren tot een vereiste nauwkeurigheid bereikt is, of indien we grote nauwkeurigheden eisen. Is dit niet het geval, dan zal het effect van de toepassing van directe methoden uiteraard geringer zijn. We geven hier een overzicht van de belangrijkste directe methoden. Sommige van de besproken methoden kunnen gegeneraliseerd worden voor algemenere gebieden dan een rechthoek. Zo niet, dan kunnen we in deze gevallen de methode van Buzbee, Dorr, George and Golub (1971) toepassen, waarbij we de oplossing van (1.2) uitbreiden tot een rechthoekig gebied, dat het oorspronkelijke gebied omvat.
We kunnen de belangrijkste directe methoden verdelen in

Blok methoden,
Cyclische reductie methoden,
Tensorproduct of matrix decompositie methoden,
Fourier reeks methoden,

en we zullen ze in deze volgorde bespreken. Enkele andere methoden zullen onbesproken blijven. Voor informatie hierover en voor verdere referenties verwijzen we naar Dorr (1970) en Leslie and McAvaney (1973). We discretiseren vgl. (1.2) op een rechthoekig rooster, roosterafstand h, bestaande uit N rijen, met M roosterpunten per rij ($N.M$ inwendige roosterpunten). Hiermee gaat vgl. (1.2) over in

$$u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1} - 4u_{i,j} = h^2 f_{i,j} \,, \quad i = 1,\dots,M; \quad j = 1,\dots,N,$$

waar

$$u_{i,j} = U(x_i, y_j) \,, \quad f_{i,j} = F(x_i, y_j).$$

Bovendien zijn $u_{i,0}$, $u_{i,N+1}$, $u_{0,j}$ en $u_{M+1,j}$ gegeven volgens (1.3) in de gediscretiseerde randpunten.
We nummeren de roosterpunten in de leesvolgorde en definieren

$$u_j = \begin{pmatrix} u_{1,j} \\ \vdots \\ u_{M,j} \end{pmatrix}, \quad u = \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix}, \qquad j = 1, \ldots, N. \tag{1.4}$$

Zij verder

$$k_{i,j} = h^2 f_{i,j} - \big(\delta_{i1} G(x_0, y_j) + \delta_{iM} G(x_{M+1}, y_j) + \delta_{j1} G(x_i, y_0) + \delta_{jN} G(x_i, y_{N+1}) \big),$$

waar δ_{ij} de Kronecker delta is, dan definiëren we de vector k analoog aan (1.4). Is L tenslotte een blok tridiagonale matrix

$$L = \begin{pmatrix} D_1 & C_1 & & & \\ B_2 & D_2 & C_2 & & O \\ & \ddots & \ddots & \ddots & \\ & & & & C_{N-1} \\ O & & & B_N & D_N \end{pmatrix}, \tag{1.5}$$

waar B_i, D_i en C_i $(M \times M)$-matrices zijn, dan heeft de matrix L_0 uit (1.1) de vorm (1.5) met

$$D_i = D = \begin{pmatrix} -4 & 1 & & \\ 1 & -4 & 1 & \\ & \ddots & \ddots & 1 \\ & & 1 & -4 \end{pmatrix}_{M \times M}, \quad i = 1, \ldots, N; \quad C_i = B_{i+1} = I, \quad i = 1, \ldots N-1; \tag{1.6}$$

waar I de $(M \times M)$ eenheidsmatrix is.

2. Blokmethoden.

We beschouwen de matrix vergelijking

$$L u = k ,$$

waar L gegeven is door (1.5). Zijn de vectoren u en k geordend volgens (1.4) dan vinden we met behulp van blok Gauss eliminatie (Varga, 1962, p. 196)

$$\begin{aligned}
&W_1 = D_1^{-1} C_1 , \quad W_i = (D_i - B_i W_{i-1})^{-1} C_i , \quad i = 2, \dots, N-1 ; \\
&v_1 = D_1^{-1} k_1 , \quad v_i = (D_i - B_i W_{i-1})^{-1} (k_i - B_i v_{i-1}) , i = 2, \dots, N ; \qquad (2.1) \\
&u_N = v_N , \quad u_i = v_i - W_i u_{i+1} , \quad i = 1, \dots, N-1 .
\end{aligned}$$

In het speciale geval dat D_i, C_i, B_i de vorm (1.6) hebben, kan de procedure (2.1) aanzienlijk worden vereenvoudigd (zie bijvoorbeeld Karlqvist, 1952). We schrijven in dit geval

$$\begin{pmatrix} D & I & & O \\ I & D & I & \\ & \ddots & \ddots & \ddots & \\ & & & & I \\ O & & & I & D \end{pmatrix} = \begin{pmatrix} S_1 & & & O \\ I & \frac{S_2}{S_1} & & \\ & \ddots & \ddots & \\ O & & I & \frac{S_N}{S_{N-1}} \end{pmatrix} \begin{pmatrix} I & \frac{I}{S_1} & & O \\ & \ddots & \ddots & \\ & & & \frac{S_{N-2}}{S_{N-1}} \\ O & & & I \end{pmatrix} ,$$

waar S_i polynomen zijn in de matrix D en bepalen v uit

$$\begin{pmatrix} S_1 & & & O \\ I & \frac{S_2}{S_1} & & \\ & \ddots & \ddots & \\ O & & I & \frac{S_N}{S_{N-1}} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ \vdots \\ v_N \end{pmatrix} = \begin{pmatrix} k_1 \\ \vdots \\ \vdots \\ k_N \end{pmatrix} \qquad (2.2)$$

en vervolgens u met behulp van

$$\begin{pmatrix} I & \frac{I}{S_1} & & O \\ & \ddots & \ddots & \\ & & \ddots & \frac{S_{N-2}}{S_{N-1}} \\ O & & & I \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ \vdots \\ u_N \end{pmatrix} = \begin{pmatrix} v_1 \\ \vdots \\ \vdots \\ v_N \end{pmatrix} . \qquad (2.3)$$

Men kan verifiëren (Karlqvist, 1952) dat geldt

$$S_0(D) = I, \quad S_1(D) = D, \quad S_j(D) = D\, S_{j-1}(D) - S_{j-2}(D), \quad j \geq 2,$$

of

$$S_j(D) = \prod_{r=1}^{j} \left(D - x_r(j) I\right), \quad x_r(j) = 2 \cos \frac{r\pi}{j+1}, \quad j \geq 1. \qquad (2.4)$$

We definiëren

$$H_j = S_j^{-1} S_{j-1} ,$$

zodat

$$H_j H_{j-1} \cdots\cdots H_{j-r+1} = S_j^{-1} S_{j-r} ,$$

waarmee (2.2), analoog aan de tweede regel van (2.1), geschreven kan worden als

$$v_1 = H_1 k_1 , \quad v_i = H_i \left(k_i - v_{i-1}\right) , \quad i = 2, \ldots, N , \qquad (2.5)$$

of

$$v_i = S_i^{-1}(D) \sum_{r=1}^{i} (-1)^{r+i} S_{r-1}(D)\, k_r . \qquad (2.5)'$$

Hiermee gaat (2.3), analoog aan de derde regel van (2.1), over in

$$u_N = v_N = S_N^{-1}(D) \sum_{r=1}^{N} (-1)^{r+N} S_{r-1}(D) k_r , \quad u_i = -H_i u_{i+1} + v_i \qquad (2.6)$$

of met (2.5)'

$$u_i = S_i^{-1}(D) \left(\sum_{r=1}^{i} (-1)^{r+i} S_{r-1}(D) k_r - S_{i-1}(D)\, u_{i+1} \right), \quad i = N-1, \ldots, 1.$$

De matrix inversie $y = S_j^{-1} z$ kan uitgevoerd worden door gebruik te maken van (2.4)

$$y = \prod_{r=1}^{j} (D - x_r(j) I)^{-1} z ,$$

zodat iedere operatie $(D - x_r(j) I)^{-1}$ het inverteren van een diagonaal dominante tridiagonale matrix inhoudt (zie hiervoor Varga, 1962, p. 195). Schlechter (1960, p. 287) heeft in plaats van (2.5) en (2.6), een gesimplifiseerde methode voorgesteld, die hij een H-proces noemt. We schrijven daartoe

$$\begin{pmatrix} D & I & & & O \\ I & D & I & & \\ & \ddots & \ddots & \ddots & \\ & & & & I \\ O & & & I & D \end{pmatrix} = \begin{pmatrix} I & & & & O \\ H_1 & I & & & \\ & \ddots & \ddots & & \\ & & & & \\ O & & & H_{N-1} & I \end{pmatrix} \begin{pmatrix} H_1^{-1} & I & & & O \\ & \ddots & \ddots & & \\ & & & & I \\ O & & & & H_N^{-1} \end{pmatrix},$$

waar H_j gedefinieerd is zoals boven. Er geldt

$$H_j + H_{j+1}^{-1} = D .$$

Passen we een proces toe analoog aan (2.2) en (2.3) dan krijgen we in plaats van (2.5)

$$v_1 = k_1 , \quad S_{i-2}(D)\, v_{i-1} = S_{i-1}(D)\,(k_i - v_i) .$$

Definiëren we

$$v_i^* = S_{i-1}\, v_i ,$$

dan geldt

$$v_1^* = k_1 , \quad v_i^* = S_{i-1} k_i - v_{i-1}^* , \quad i = 2, \ldots, N ,$$

zodat (2.6) in dit geval wordt

$$S_N u_N = v_N^* , \quad u_{N-1} = k_N - D u_N , \quad u_i = k_{i+1} - D u_{i+1} - u_{i+2} , \quad i = N-2, \ldots, 1. \tag{2.7}$$

Methode (2.7) kan gegeneraliseerd worden voor matrices van een algemenere bloktridiagonale vorm. Er zij wel opgemerkt dat procedure (2.7) bijzonder instabiel is met betrekking tot afrondingsfouten (voor verdere discussie zie Dorr (1970)).

3. Cyclische reductie methoden.

We beschouwen vgl.(1.1)

$$\mathcal{L}_0 u = k .$$

Door een geschikte ordening van de roosterpunten heeft $\mathcal{L}_0$ de vorm

$$\mathcal{L}_0 = \begin{pmatrix} I & F \\ F^T & I \end{pmatrix} , \tag{3.1}$$

waar de eenheidsmatrices vierkant zijn. De matrix $\mathcal{L}_0$ heet 2-cyclisch (Varga, 1962, p. 99). Ordenen we u en k overeenkomstig (3.1) dan vinden we

$$\begin{aligned} u_1 + F u_2 &= k_1 , \\ F^T u_1 + u_2 &= k_2 , \end{aligned} \tag{3.2}$$

of

$$(I - F^T F) u_2 = k_2 - F^T k_1 . \tag{3.3}$$

Hiermee hebben we het aantal vergelijkingen gereduceerd. We noemen (3.3) de cyclische reductie van matrix vergelijking (3.2). (Zie Hageman and Varga (1964) voor de toepassing van iteratieve methoden op (3.3)).
We kunnen bovenstaande reductie methode toepassen op matrices van de vorm (1.5) indien geldt $D_i = A$, $B_i = C_i = T$ met

T, A $(M \times M)$-matrices,
$TA = AT,$
$N = 2^{q+1} - 1$ voor elke integer $q \geq 1$.

We ordenen u en k weer volgens (1.4) en schrijven

$$\begin{pmatrix} A & T & & & O \\ T & A & T & & \\ & \ddots & \ddots & \ddots & \\ & & & & T \\ O & & & T & A \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ \vdots \\ u_N \end{pmatrix} = \begin{pmatrix} k_1 \\ \vdots \\ \vdots \\ k_N \end{pmatrix} , \tag{3.4}$$

of

$$A u_1 + T u_2 = k_1 , \tag{3.5}$$

$$T u_{j-1} + A u_j + T u_{j+1} = k_j , \quad j = 2, \ldots, N-1 , \tag{3.5)'$$

$$T u_{N-1} + A u_N = k_N . \tag{3.5)''$$

We herschrijven (3.5)' als

$$
\begin{aligned}
Tu_{j-2} + Au_{j-1} + Tu_j \qquad\qquad &= k_{j-1} ,\\
Tu_{j-1} + Au_j + Tu_{j+1} \qquad &= k_j ,\\
Tu_j + Au_{j+1} + Tu_{j+2} &= k_{j+1} .
\end{aligned}
$$

Vermenigvuldigen we de eerste en derde vergelijking met T , de tweede vgl. met $-A$ en tellen we op dan krijgen we

$$T^2 u_{j-2} + (2T^2 - A^2) u_j + T^2 u_{j+2} = Tk_{j-1} - Ak_j + Tk_{j+1} .$$

Dus als j even is, bevat het nieuwe systeem u_j's met even indices. Dit procés, waarbij we de vergelijkingen op bovenstaande manier reduceren, staat bekend als cyclische (oneven - even) reductie. We kunnen vgl. (3.4) equivalent schrijven als ($n = N+1$ = even)

$$
\begin{pmatrix}
(2T^2-A^2) & T^2 & & & O \\
T^2 & (2T^2-A^2) & T^2 & & \\
& \ddots & \ddots & \ddots & \\
& & & & T^2 \\
O & & & T^2 & (2T^2-A^2)
\end{pmatrix}
\begin{pmatrix} u_2 \\ \vdots \\ \vdots \\ u_{n-2} \end{pmatrix}
=
\begin{pmatrix} Tk_1 + Tk_3 - Ak_2 \\ \vdots \\ \vdots \\ Tk_{n-1} + Tk_{n-3} + Ak_{n-2} \end{pmatrix}
\qquad (3.6)
$$

en

$$
\begin{pmatrix}
A & & & O \\
& A & & \\
& & \ddots & \\
O & & & A
\end{pmatrix}
\begin{pmatrix} u_1 \\ \vdots \\ \vdots \\ u_{n-1} \end{pmatrix}
=
\begin{pmatrix} k_1 - Tu_2 \\ k_3 - Tu_2 - Tu_4 \\ \vdots \\ k_{n-1} - Tu_{n-2} \end{pmatrix}
\qquad (3.6)'
$$

Daar $n = 2^{q+1}$ en het nieuwe systeem (3.6) alleen u_j's met even indices bevat is de blokdimensie van het nieuwe systeem $\frac{n-2}{2} = 2^q - 1$.
Zijn de u_j's met even indices bekend, dan kunnen de u_j's met oneven indices gemakkelijk uit (3.6)' worden opgelost.
Systeem (3.6)' wordt aangeduid als de geëlimineerde vergelijkingen. Daar de matrix uit (3.6) wederom blok tridiagonaal is en van de vorm (3.4) kunnen we het reductieproces voortzetten tot we één blok hebben. Om het proces recursief te kunnen beschrijven definiëren we

$$A^{(0)} = A \ , \ T^{(0)} = T, \ k_j^{(0)} = k_j \quad , \ j = 1, \dots, n-1 \ ;$$

zodat voor $r = 0, 1, \dots\, q-1$ geldt

$$A^{(r+1)} = 2\left(T^{(r)}\right)^2 - \left(A^{(r)}\right)^2 ,$$

$$T^{(r+1)} = \left(T^{(r)}\right)^2 ,$$

$$k_j^{(r+1)} = T^{(r)}\left(k_{j-2^r}^{(r)} + k_{j+2^r}^{(r)}\right) - A^{(r)} k_j^{(r)} .$$

Bij iedere stap hebben we een nieuw systeem vergelijkingen

$$\begin{pmatrix} A^{(r)} & T^{(r)} & & & \\ T^{(r)} & A^{(r)} & T^{(r)} & & O \\ & \ddots & \ddots & \ddots & \\ & & & & T^{(r)} \\ O & & & T^{(r)} & A^{(r)} \end{pmatrix} \begin{pmatrix} u_{2^r} \\ \vdots \\ u_{j \cdot 2^r} \\ \vdots \\ u_{(2^{q+1-r}-1)2^r} \end{pmatrix} = \begin{pmatrix} k_{2^r}^{(r)} \\ \vdots \\ k_{j \cdot 2^r}^{(r)} \\ \vdots \\ k_{(2^{q+1-r}-1)2^r}^{(r)} \end{pmatrix} .$$

De geëlimineerde vgln zijn

$$\begin{pmatrix} A^{(r-1)} & & O \\ & \ddots & \\ O & & A^{(r-1)} \end{pmatrix} \begin{pmatrix} u_{2^{r-1}} \\ \vdots \\ u_{(2j-1)2^{r-1}} \\ \vdots \\ u_{(2^{q+2-r}-1)2^{r-1}} \end{pmatrix} =$$

$$\begin{pmatrix} k^{(r-1)}_{2^{r-1}} - T^{(r-1)} u_{2^r} \\ \vdots \\ k^{(r-1)}_{(2j-1)2^{r-1}} - T^{(r-1)}\left(u_{j\cdot 2^r} + u_{(j-1)2^r}\right) \\ \vdots \\ k^{(r-1)}_{(2^{q+2-r}-1)2^{r-1}} - T^{(r-1)} u_{(2^{q+1-r}-1)2^r} \end{pmatrix} .$$

Na q stappen moeten we het systeem vergelijkingen

$$A^{(q)} u_{2^q} = k^{(q)}_{2^q}$$

oplossen, waarna we via de geëlimineerde vergelijkingen de oplossing van (3.4) kunnen vinden. Voor een voorbeeld, zie Rosmond and Faulkner (1976, p. 643). Bovenstaande berekeningen worden sterk vereenvoudigd door gebruik te maken van de volgende identiteiten (zie Buzbee, Golub and Nielson (1970, p. 634))

$$T^{(r)} = T^{2^r} ,$$

$$A^{(r)} = - \prod_{j=1}^{2^r} \left(A + x_j(r) T\right) ,$$

$$x_j(r) = 2 \cos \frac{(2j-1)\pi}{2^{r+1}} .$$

Buzbee, Golub and Nielson (1970) hebben bovenstaande methode gegeneraliseerd voor meer algemene elliptische partiële differentiaal vergelijkingen, met verschillende typen randvoorwaarden en aangetoond dat de methode instabiel is. Rosmond and Faulkner (1976) hebben voor vierkante gebieden ($M = N$) als resultaat gevonden dat de methode instabiel is voor $N > 31$. In Buzbee, Golub and Nielson (1970) wordt aangegeven hoe de methode te wijzigen om een stabiele procedure te krijgen, wat resulteert in de cyclische reductie methode van Buneman (1969).

Sweet (1974) beschrijft een procedure waarin $N = 2^{\ell} 3^m 5^n - 1$, waar ℓ, m en n willekeurige integers zijn. Zie ook Schumann and Sweet (1976) voor een procedure die geen enkele beperking oplegt aan N .

Voor een verzameling rekenprogramma's op dit gebied, zie Swarztrauber and Sweet (1975). Het zal duidelijk zijn dat de hier boven beschreven methode aanzienlijk wordt vereenvoudigd door toepassing op de Poisson vergelijking zodat $A = D$, $T = I$.

4. Tensorproduct of matrix decompositie methoden.

We beschouwen weer de matrix vgl.

$$\begin{pmatrix} A & T & & & \\ T & A & T & & O \\ & \ddots & \ddots & \ddots & \\ & & & & T \\ O & & & T & A \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ \vdots \\ u_N \end{pmatrix} = \begin{pmatrix} k_1 \\ \vdots \\ \vdots \\ k_N \end{pmatrix}, \tag{4.1}$$

waar A en T wederom $(M \times M)$-matrices zijn, waarvan we veronderstellen dat ze symmetrisch zijn en commuteren d.w.z. $AT = TA$.
In dit geval (Bellman (1960)) bestaat er een matrix Q zodanig dat

$$Q^T A Q = \Lambda\,,\quad Q^T T Q = \Omega\,,\quad Q^T Q = I\,, \tag{4.2}$$

waar Λ en Ω reële diagonaal matrices zijn. De matrix Q heeft als kolommen de eigenvectoren van A en T en de diagonaalelementen van Λ en Ω zijn de eigenwaarden van A en T. We ordenen u_j en k_j als in hoofdstuk 1

$$u_j = \begin{pmatrix} u_{1,j} \\ \vdots \\ u_{M,j} \end{pmatrix},\quad k_j = \begin{pmatrix} k_{1,j} \\ \vdots \\ k_{M,j} \end{pmatrix}. \tag{4.3}$$

Schrijven we (4.1) als

$$\begin{aligned} A u_1 + T u_2 &= k_1\,, \\ T u_{j-1} + A u_j + T u_{j+1} &= k_j\,, \qquad j = 2, \ldots, N-1, \\ T u_{N-1} + A u_N &= k_N\,, \end{aligned}$$

dan gaat dit, met (4.2), over in

$$\begin{aligned} \Lambda \bar{u}_1 + \Omega \bar{u}_2 &= \bar{k}_1\,, \\ \Omega \bar{u}_{j-1} + \Lambda \bar{u}_j + \Omega \bar{u}_{j+1} &= \bar{k}_j\,, \qquad j = 2, \ldots, N-1\,, \\ \Omega \bar{u}_{N-1} + \Lambda \bar{u}_N &= \bar{k}_N\,, \end{aligned} \tag{4.4}$$

waar $\bar{u}_j = Q^T u_j$ en $\bar{k}_j = Q^T k_j$ geordend zijn als in (4.3).
Vgl. (4.4) kan, voor $i = 1, \dots, M$ worden geschreven als

$$
\begin{aligned}
\lambda_i \bar{u}_{i,1} + \omega_i \bar{u}_{i,2} &= \bar{k}_{i,1}, \\
\omega_i \bar{u}_{i,j-1} + \lambda_i \bar{u}_{i,j} + \omega_i \bar{u}_{i,j+1} &= \bar{k}_{i,j}, \quad j = 2, \dots, N-1, \\
\omega_i \bar{u}_{i,N-1} + \lambda_i \bar{u}_{i,N} &= \bar{k}_{i,N}.
\end{aligned}
\tag{4.5}
$$

Schrijven we verder

$$
\Gamma_i = \begin{pmatrix} \lambda_i & \omega_i & & & 0 \\ \omega_i & \lambda_i & \omega_i & & \\ & \ddots & \ddots & \ddots & \\ & & & & \omega_i \\ 0 & & & \omega_i & \lambda_i \end{pmatrix}_{N \times N},
$$

$$
\tilde{u}_i = \begin{pmatrix} \bar{u}_{i,1} \\ \vdots \\ \bar{u}_{i,N} \end{pmatrix}, \quad \tilde{k}_i = \begin{pmatrix} \bar{k}_{i,1} \\ \vdots \\ \bar{k}_{i,N} \end{pmatrix},
$$

dan is vgl. (4.5) equivalent met

$$
\Gamma_i \tilde{u}_i = \tilde{k}_i, \qquad i = 1, \dots, M,
$$

waar Γ_i een eenvoudig te inverteren tridiagonale, symmetrische matrix is.
Is $\tilde{u}_i$ ($i = 1, \dots, M$) bekend, dan vinden we u_j uit $u_j = Q \bar{u}_j$, $j = 1, \dots, N$.
In het geval $A = D$, $T = I$ geldt voor $Q = (q_{i,j})$

$$
q_{i,j} = \sqrt{\frac{2}{M+1}} \sin \frac{i j \pi}{M+1}, \quad Q^2 = I,
$$
$$
\lambda_j = 2\left(\left(\cos \frac{j\pi}{M+1} - 1\right) - 1\right).
$$

Het hierboven beschreven proces om vgl. (4.1) op te lossen staat bekend als matrix decompositie (Buzbee, Golub and Nielson, 1970).
Nauw verwant hiermee is de tensorproduct methode (Lynch, Rice and Thomas, 1964). Zij $B = (b_{i,j})$ een $(N \times N)$-matrix en A een $(M \times M)$-matrix dan is het tensor product $B \times A$ de blok N -dimensionale matrix, waarvan het (i,j)-de element gelijk is aan $b_{i,j} A$. Dus als

$$B = \begin{pmatrix} -2 & 1 & & & 0 \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ 0 & & & 1 & -2 \end{pmatrix}_{N \times N}, \quad A = \begin{pmatrix} -2 & 1 & & & 0 \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ 0 & & & 1 & -2 \end{pmatrix}_{M \times M}$$

en I_N, I_M de $(N \times N)$ respectievelijk $(M \times M)$-eenheidsmatrices zijn, dan kan vgl. (1.1) worden geschreven als

$$\left[(I_N \times A) + (B \times I_M) \right] u = k. \qquad (4.6)$$

Zij

$$C = \begin{pmatrix} -2 & 1 & & & 0 \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ 0 & & & 1 & -2 \end{pmatrix}_{J \times J}$$

dan bestaat er een matrix $R = (r_{i,j})$ met

$$r_{i,j} = \sqrt{\frac{2}{J+1}} \sin \frac{ij\pi}{J+1},$$
$$\lambda_j = 2\left(\cos \frac{j\pi}{J+1} - 1\right), \quad R^2 = I_J,$$

zodat

$$RCR = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_J \end{pmatrix} = D.$$

Noemen we de overeenkomstige matrices voor B en A , R, D respectievelijk R_1, D_1 en R_2, D_2 , dan schrijven we analoog aan het voorgaande

$$R_1 B R_1 = D_1, \quad R_1^2 = I_N,$$
$$R_2 A R_2 = D_2, \quad R_2^2 = I_M.$$

Hiermee wordt de oplossing van (4.6) gegeven door

$$u = (R_1 \times R_2) S^{-1} (R_1 \times R_2) k,$$

waar

$$S = \left[(I_N \times D_2) + (D_1 \times I_M) \right],$$

zodat de matrix S een diagonaal matrix is. Lynch, Rice and Thomas (1964) hebben de methode gegeneraliseerd voor meer algemene differentiaal vergelijkingen, maar hun methode is slechts toepasbaar in rechthoekige gebieden.

5. Fourier reeks methoden.

We keren terug naar de matrix vergelijking uit de inleiding

$$L_0 u = k. \tag{5.1}$$

Schrijven we

$$\begin{pmatrix} D & I & & O \\ I & D & I & \\ & \ddots & \ddots & I \\ O & & I & D \end{pmatrix} = \begin{pmatrix} D_0 & & O \\ & \ddots & \\ O & & D_0 \end{pmatrix} + \begin{pmatrix} -2I & I & & O \\ I & -2I & I & \\ & \ddots & \ddots & I \\ O & & I & -2I \end{pmatrix}, \tag{5.2}$$

waar

$$D_0 = \begin{pmatrix} -2 & 1 & & O \\ 1 & -2 & 1 & \\ & \ddots & \ddots & 1 \\ O & & 1 & -2 \end{pmatrix},$$

dan kan de oplossing langs de j-de lijn geschreven worden als

$$u_{i,j} = \sqrt{\frac{2}{M+1}} \sum_{r=1}^{M} \hat{u}_{r,j} \sin \frac{r\pi i}{M+1}, \quad i = 1, \ldots, M, \tag{5.3}$$

waar

$$\sin \frac{r\pi i}{M+1}, \qquad i = 1, \ldots, M$$

een eigenfunctie is van D_0 (gediscretiseerd analogon van $h^2 \frac{\partial^2}{\partial x^2}$) met eigenwaarde

$$\lambda_r = 2\left(\cos \frac{r\pi}{M+1} - 1\right).$$

Substitutie van (5.3) in (5.1) (waar L_0 geschreven is volgens (5.2)) geeft voor $r = 1, \ldots, M$ een $(N \times N)$ differentievergelijking

$$\hat{u}_{r,j-1} - 2\hat{u}_{r,j} + \hat{u}_{r,j+1} + \lambda_r \hat{u}_{r,j} = \hat{k}_{r,j}, \quad j = 1, \ldots, N,$$

waar

$$\hat{k}_{r,j} = \sqrt{\frac{2}{M+1}} \sum_{l=1}^{M} k_{l,j} \cdot \sin \frac{l \pi r}{M+1} .$$

Hockney (1965) heeft aangetoond dat deze methode bijzonder effectief kan zijn, door het aantal roosterlijnen te beperken tot $N = 2^q$ of $N = 3.2^q$. Een reductie van het aantal vergelijkingen wordt verkregen door éénmaal cyclische reductie toe te passen, volgens de methode van hoofdstuk 3. Hockney's methode laat zich als volgt omschrijven

a. Pas oneven-even reductie toe om het probleem slechts op de even lijnen te beschouwen.
b. Bepaal $\hat{k}_{r,j}$ op de even lijnen.
c. Bepaal $\hat{u}_{r,j}$ op de even lijnen.
d. Bepaal $u_{i,j}$ op de even lijnen.
e. Bepaal de oplossing op de oneven lijnen.

Toepassing van de "fast Fourier transform" is bijzonder voordelig (Hockney (1969)). Le Bail (1972) heeft Fourier methoden toegepast op meer algemene lineaire tweede-orde partiële differentiaalvergelijkingen. Voor verdere toepassing van Fourier methoden, zie ook Sköllermo (1975).

6. Besluit.

Dorr (1970) heeft voorgaande directe methoden, wat betreft het aantal arithmetische operaties, vergeleken met de punt successieve overrelaxatie methode (SOR) en de Peaceman-Rachford "alternating-direction implicit" methode (ADI). Hierbij is het aantal operaties bij de iteratieve methoden bepaald door het aantal iteraties, dat nodig is om de beginfout met een factor h^2 (h = roosterafstand) te reduceren. Enkele resultaten zijn (vierkant rooster $N=M$).

methode	aantal operaties
Blok (veelterm)	$6\, N^3$
Blok (Schechter)	$\frac{3}{2}\, N^3$
cyclische reductie (Buneman)	$\frac{9}{2}\, N^2 \log_2 N$
Tensor product (Lynch et al)	$8\, N^3$
Fourier reeks (Hockney)	$5\, N^2 \log_2 N$
SOR	$\frac{3}{2}\, N^3 \log_2 N$
ADI	$4\, N^2 (\log_2 N)^2$

Experimenten, die een variant van de tensor product methode en de cyclische reductie methode van Buneman vergelijken met de Sheldon versnelling van de SOR methode, voor verschillende reductiefactoren, vindt men in Lesley and McAvaney (1973).
Vergelijkende experimenten tussen de Buneman methode en de SOR methode vindt men ook in Rosmond and Faulkner (1976). Bovengenoemde experimenten tonen duidelijk de superioriteit van de directe methoden (speciaal de Buneman methode) vooral wanneer er grote fout-reductie factoren in het geding zijn. Hierbij moeten we wel de beperkingen voor ogen houden, die de toepassing van deze methoden met zich mee brengt. Laten we deze beperkingen vallen (b.v. wat betreft rechthoekige gebieden) dan worden de resultaten een stuk ongunstiger, zie Buzbee, Dorr, George and Golub (1971, p. 734 en 735), hoewel de hierin beschreven methode efficiënter toegepast kan worden (Buzbee and Dorr (1974)), zie ook Temperton (1977), speciaal voor het oplossen van een Poisson vergelijking op een regelmatige achthoek.
Een discussie over het gebruik van een combinatie van directe en iteratieve methoden bij het oplossen van algemenere elliptische vergelijkingen vindt men in Concus and Golub (1973).

7. Literatuur.

R. Bellman (1960)	Introduction to Matrix Analysis, McGraw-Hill, N.Y.
O. Buneman (1969)	A compact non-iterative Poisson solver, Rep. SU-IPR-294, Institute for plasma research, Stanford University, Stanford, Calif.
B.L. Buzbee, G.H. Golub and C.W. Nielson (1970)	On direct methods for solving Poisson's equations, SIAM J. Numer. Anal. 7, pp. 627-656.
B.L. Buzbee, F.W. Dorr, J.A. George and G.H. Golub (1971)	The direct solution of the discrete Poisson equation on irregular regions, SIAM J. Numer. Anal., 8, pp. 722-736.
B.L. Buzbee and F.W. Dorr (1974)	The direct solution of the biharmonic equation on rectangular regions and the Poisson equation on irregular regions, SIAM J. Numer. Anal., 11, pp. 753-763.
P. Concus and G.H. Golub (1973)	Use of fast direct methods for the efficient numerical solution of non-separable elliptic equations, SIAM J. Numer. Anal., 10, pp. 1103-1120.
F.W. Dorr (1970)	The direct solution of the discrete Poisson equation on a rectangle, SIAM Review, 12, pp. 248-263.
L.A. Hageman and R.S. Varga (1964)	Block iterative methods for cyclically reduced matrix equations, Numer. Math., 6, pp. 106-119.
R.W. Hockney (1965)	A fast direct solution of Poisson's equation using Fourier analysis, J. Assoc. Comp. Mach., 12, pp. 95-113.
R.W. Hockney (1969)	The potential calculation and some applications in: Methods in computational physics, 9, Acad. Press, Inc. N.Y.
O. Karlqvist (1952)	Numerical solution of elliptic difference equations by matrix methods, Tellus, 4, pp. 374-384.

R.C. Le Bail (1972) — Use of fast Fourier transforms for solving partial differential equations in physics, J. Comp. Phys, 9, pp. 440-465.

L.M. Leslie and B.J. McAvaney (1973) — Comparative test of direct and iterative methods for solving Helmholtz-type equations, Mon. Wea. Rev., 101, pp. 235-239.

R.E. Lynch, J.R. Rice and D.H. Thomas (1964) — Direct solution of partial difference equations by tensor product methods, Numer. Math. 6, pp. 185-199.

T.E. Rosmond and F.D. Faulkner (1976) — Direct solution of elliptic equations by block cyclic reduction and factorization, Mon. Wea. Rev. , 104, pp. 641-649.

S. Schechter (1960) — Quasi-tridiagonal matrices and type-insensitive difference equations, Quart. Appl. Math., 18, pp. 285-295.

U. Schumann and R.A. Sweet (1976) — Direct Poisson equation solver for potential and pressure fields on a staggered grid with obstacles, in: Lecture Notes in Physics 59, Proceedings of the fifth international conference on numerical methods in fluid dynamics, A.I. van de Vooren and P.J. Zandbergen eds, Springer Verlag, Heidelberg.

G. Sköllermo (1975) — A Fourier method for the numerical solution of Poisson's equation, Math. Comp., 29, pp 697 - 711.

P.N. Swarztrauber and R.A. Sweet (1975) — Efficient Fortran subprograms for the solution of elliptic partial differential equations, NCAR Tech. Note IN/IA-109, 139 pp, Boulder, Colo.

R.A. Sweet (1974) — A generalized cyclic reduction algorithm, SIAM J. Numer. Anal., 11, pp. 506-520.

C. Temperton (1977) — An improved algorithm for the direct solution of Poisson's equation over irregular regions, Internal Report 5, Research Dept., E.C.M.W.F., Bracknell.

R.S. Varga (1962) — Matrix Iterative Analysis, Prentice-Hall, Englewood Cliffs, N.J.