

5. SCHNELLE FOURIERTRANSFORMATION UND ANWENDUNGEN

In der Signal- und Bildverarbeitung tritt folgendes Problem auf: Zu sehr vielen diskreten Zeitpunkten (oft mehr als eine Million) misst man Werte (t_i, y_i) . Typischerweise sind die Zeitpunkte t_i äquidistant und das beobachtete Signal ist periodisch. Beispiele sind Daten der Sonne oder eines EKGs. Häufig ist man am Spektrum dieses Signals interessiert, also der Fouriertransformation oder den Koeffizienten der Fouriertransformation.

Zur Erinnerung: Die Fourierreihe einer 2π -periodischen Funktion f mit Werten in \mathbb{R} oder \mathbb{C} ist

$$f(x) \sim \sum_{k \in \mathbb{Z}} \hat{f}(k) e^{ikx}, \quad (5.1)$$

wobei die Fourierkoeffizienten durch

$$\hat{f}(k) = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx \quad (5.2)$$

definiert sind. Unter bestimmten Bedingungen (vgl. Analysis-Vorlesung) stimmt f mit seiner Fourierreihe (5.1) überein.

In diesem Kapitel stellen wir mit der schnellen Fourier-Transformation einen der wichtigsten Algorithmen für die Signal- und Bildverarbeitung vor und beschreiben zudem einige Anwendungen.

5.1 Trigonometrische Interpolation und diskrete Fouriertransformation

Im Folgenden nehmen wir an, dass wir die Funktion nur an äquidistanten Werten

$$x_j = \frac{2\pi j}{N}, \quad j = 0, 1, \dots, N$$

kennen. Da f 2π -periodisch ist, gilt $f(x_N) = f(x_0)$, so dass die Trapezregel zur Approximation des Integrals in (5.2) die Darstellung

$$\hat{f}_N(k) = \frac{1}{N} \sum_{j=0}^{N-1} f(x_j) e^{-ikx_j} = \frac{1}{N} \sum_{j=0}^{N-1} f(x_j) \omega_N^{kj} \quad (5.3)$$

hat, wobei

$$\omega_N = e^{-\frac{2\pi i}{N}}$$

die primitive N -te Einheitswurzel ist. Wir schreiben ω statt ω_N , wenn N aus dem Zusammenhang klar ist. Folgen $x = [x_0 \ x_1 \ \dots \ x_{N-1}] \in \mathbb{C}^N$ setzen wir im Weiteren bei Bedarf periodisch auf beliebige ganzzahlige Indizes fort:

$$x_k = x_m \quad \text{falls} \quad k \equiv m \pmod{N}.$$

Definition 5.1. (Diskrete Fouriertransformation)

Die lineare Abbildung $\mathcal{F}_N : \mathbb{C}^N \rightarrow \mathbb{C}^N$ definiert durch

$$\mathcal{F}_N x = \hat{x} \quad \text{mit} \quad \hat{x}_k = \sum_{j=0}^{N-1} \omega_N^{kj} x_j \quad (k = 0, 1, \dots, N-1),$$

heißt **diskrete Fouriertransformation**. Die zugehörige Matrix $F_N = (\omega_N^{kj})_{kj}$ heißt **Fouriermatrix**.

Bemerkung. Die direkte Berechnung der diskreten Fouriertransformation kostet $O(N^2)$ arithmetische Operationen. Im nächsten Abschnitt lernen wir eine effizientere Methode, die schnelle Fouriertransformation kennen.

Lemma 5.2. (Orthogonalität)

Es gilt $F_N \overline{F_N} = NI_N$ oder äquivalent

$$\sum_{k=0}^{N-1} \omega_N^{kl} \overline{\omega_N^{km}} = \begin{cases} N, & l \equiv m \pmod{N}, \\ 0, & \text{sonst.} \end{cases}$$

Beweis. Es gilt $\overline{\omega} = \omega^{-1}$ und daraus folgt sofort die Behauptung für $l = m$. Für $l \neq m$ gilt

$$\sum_{k=0}^{N-1} \omega^{kl} \overline{\omega^{km}} = \sum_{k=0}^{N-1} \omega^{k(l-m)} = \frac{1 - \omega^{(l-m)N}}{1 - \omega^{l-m}} = 0,$$

da $\omega^N = 1$. □

Satz 5.3. (Inverse diskrete Fouriertransformation)

Es gilt $F_N^{-1} = \frac{1}{N} \overline{F_N} = \frac{1}{N} F_N^H$. bzw.

$$x_j = \frac{1}{N} \sum_{k=0}^{N-1} \omega_N^{-kj} \hat{x}_k, \quad j = 0, 1, \dots, N-1.$$

Beweis. Folgt aus Lemma 5.2. □

Satz 5.4. (Parseval'sche Gleichung)

In der Euklidischen Norm $\|\cdot\|$ gilt

$$\left\| \frac{1}{\sqrt{N}} \mathcal{F}_N x \right\| = \|x\| \quad \forall x \in \mathbb{C}^N,$$

d. h. $\frac{1}{\sqrt{N}} \mathcal{F}_N$ ist eine Isometrie, oder

$$\frac{1}{N} \sum_{k=0}^{N-1} |\hat{x}_k|^2 = \sum_{k=0}^{N-1} |x_k|^2.$$

Beweis. Mit Hilfe von Satz 5.3 gilt

$$\|\hat{x}\|^2 = \|F_N x\|^2 = x^H F_N^H F_N x = N x^H x = N \|x\|^2.$$

und damit die Behauptung. □

Lemma 5.5. (Symmetrie)

Die Fouriermatrix ist symmetrisch, $F_N = F_N^T$.

Beweis. Es ist $(F_N^T)_{jk} = \omega_N^{kj} = \omega_N^{jk} = (F_N)_{jk}$. □

Bemerkung. Bezeichnen wir mit $(Rx)_j = x_{-j}$, wobei wir x periodisch fortgesetzt haben, so gilt $\overline{F_N} = F_N R$, denn

$$(F_N Rx)_k = \sum_{j=0}^{N-1} \omega_N^{kj} x_{-j} \stackrel{l=N-j}{=} \sum_{l=1}^N \omega_N^{kN} \omega_N^{-kl} x_l \stackrel{x_N=x_0}{=} \sum_{l=0}^{N-1} \omega_N^{-kl} x_l = (\overline{F_N} x)_k.$$

Satz 5.6. (Aliasing)

Ist die Reihe $\sum_{k \in \mathbb{Z}} \widehat{f}(k)$ absolut konvergent, dann gilt

$$\widehat{f_N}(k) = \sum_{j \in \mathbb{Z}} \widehat{f}(k + jN).$$

Beweis. Die Voraussetzung über die Fourierkoeffizienten impliziert die gleichmäßige Konvergenz der Fourierreihe, insbesondere Gleichheit in (5.1). Daher gilt

$$\begin{aligned} \widehat{f_N}(k) &= \frac{1}{N} \sum_{j=0}^{N-1} \left(\sum_{n \in \mathbb{Z}} \widehat{f}(n) \underbrace{e^{inx_j}}_{\omega^{-nj}} \right) \omega^{kj} \\ &= \sum_{n \in \mathbb{Z}} \widehat{f}(n) \left(\frac{1}{N} \sum_{j=0}^{N-1} \omega^{(k-n)j} \right) \\ &= \sum_{j \in \mathbb{Z}} \widehat{f}(k + jN), \end{aligned}$$

wobei die letzte Gleichung aus der Orthogonalität (Lemma 5.2) folgt. □

Beispiel. Zur Veranschaulichung des Aliasing betrachten wir in Abbildung 5.1 das Signal $\sin(\omega t)$ mit $\omega = 19$ auf dem Intervall $[0, 2\pi\omega]$. Dieses Signal tasten wir an $N = 18$ Punkten ab. Das trigonometrische Interpolationspolynom in diesen Stützstellen “täuscht” die Frequenz eins vor, was zur Aussage von Satz 5.6 mit $k = 1$ passt. \diamond

Korollar 5.7. Ist $f : \mathbb{R} \rightarrow \mathbb{C}$ p -mal stetig differenzierbar mit $p \geq 2$ und 2π -periodisch, dann gilt

$$\widehat{f_N}(k) - \widehat{f}(k) = O(N^{-p}) \quad \text{für} \quad |k| \leq N/2.$$

Insbesondere ist für $h = 2\pi/N$

$$h \sum_{j=0}^{N-1} f(x_j) - \int_0^{2\pi} f(x) dx = O(h^p).$$

Beweis. Mit Hilfe partieller Integration folgt

$$\begin{aligned} \widehat{f}(k) &= \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx \\ &= -\frac{(ik)^{-1}}{2\pi} \underbrace{f(x) e^{-ikx}}_0 \Big|_0^{2\pi} + \frac{(ik)^{-1}}{2\pi} \int_0^{2\pi} f'(x) e^{-ikx} dx \\ &= \dots = \frac{(ik)^{-p}}{2\pi} \int_0^{2\pi} f^{(p)}(x) e^{-ikx} dx \end{aligned}$$

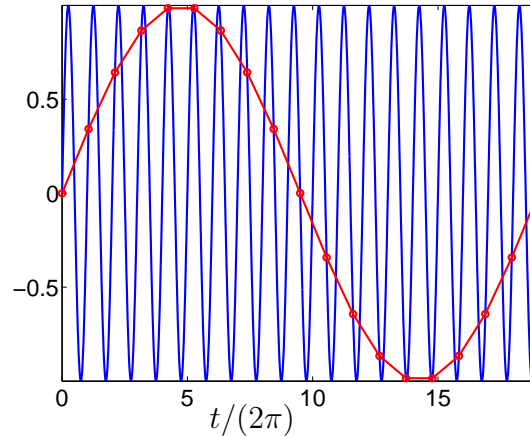


Abb. 5.1: Aliasing Phänomen: $\sin(19t)$ (rot) und das trigonometrische Interpolationspolynom in 20 Knoten (rot).

und damit die Abschätzung

$$|\hat{f}(k)| \leq Ck^{-p}, \quad C = \frac{1}{2\pi} \int_0^{2\pi} |f^{(p)}(x)| dx.$$

Für $|k| \leq N/2$ und $j \neq 0$ ist $|k + jN| \geq (|j| - 1/2)N$ und wir erhalten aus Satz 5.6

$$|\hat{f}_N(k) - \hat{f}(k)| \leq 2 \sum_{j \geq 1} C(j - 1/2)^{-p} N^{-p} = \tilde{C} N^{-p},$$

denn die Reihe konvergiert für $p > 1$.

Die Zusatzbehauptung entspricht der Aussage für $k = 0$. □

Wir betrachten jetzt das trigonometrische Interpolationsproblem, zu Stützstellen x_j und gegebenen Werten y_j ein trigonometrisches Polynom (eine Linearkombination von Funktionen e^{ikx}) zu finden, welches durch die Punkte (x_j, y_j) , $j = 0, \dots, N-1$ geht.

Satz 5.8. Zu gegebenen Werten y_j , $j = 0, \dots, N-1$ und N gerade hat

$$t_N(x) = \frac{1}{N} \sum'_{k=-N/2}^{N/2} \hat{y}_k e^{ikx} := \frac{1}{2N} (\hat{y}_{-N/2} e^{-iNx/2} + \hat{y}_{N/2} e^{iNx/2}) + \frac{1}{N} \sum_{|k| < N/2} \hat{y}_k e^{ikx}$$

die Interpolationseigenschaft $t_N(x_j) = y_j$, $j = 0, \dots, N-1$.

Beweis. Es sei j beliebig aber fest. Wegen der Periodizität folgt

$$t_N(x_j) = \frac{1}{N} \sum_{k=0}^{N-1} \hat{y}_k e^{ikx_j} = \frac{1}{N} \sum_{k=0}^{N-1} \hat{y}_k \omega_N^{-kj} = (F_N^{-1} \hat{y})_j = (F_N^{-1} F_N y)_j = y_j,$$

t_N hat also die Interpolationseigenschaft. □

Satz 5.9. (Fehler bei der trigonometrischen Interpolation)

Falls $(\hat{f}(n))_{n \in \mathbb{Z}}$ absolut summierbar ist, gilt

$$|t_N(x) - f(x)| \leq 2 \sum'_{|n| \geq N/2} |\hat{f}(n)| \quad \forall x.$$

Beweis. Nach Satz 5.6 gilt

$$\begin{aligned}
 |t_N(x) - f(x)| &= \left| \sum'_{|n| \leq N/2} (\hat{f}_N(n) - \hat{f}(n)) e^{inx} - \sum'_{|n| \geq N/2} \hat{f}(n) e^{inx} \right| \\
 &\leq 2 \sum'_{|n| \geq N/2} |\hat{f}(n)|. \quad \square
 \end{aligned}$$

5.2 Schnelle Fouriertransformation (FFT)

Wir leiten jetzt die effiziente Berechnung von $\hat{x} = \mathcal{F}_N x$ mit Hilfe der schnellen Fouriertransformation her. Die Idee von Cooley-Tukey (1965) lautet ähnlich wie bei Danielson-Lanczos 1942, Runge 1925, Gauß und Cäsar: *Divide et impera!*.

$$\begin{array}{ccccccc}
 & & & & \mathcal{F}_{N/4} \begin{pmatrix} x_0 \\ x_4 \\ x_8 \\ \vdots \end{pmatrix} & \mathcal{F}_1 x_0 & = x_0 \\
 & & & & & \mathcal{F}_1 x_{N/2} & = x_{N/2} \\
 & & \mathcal{F}_{N/2} \begin{pmatrix} x_0 \\ x_2 \\ x_4 \\ \vdots \end{pmatrix} & \mathcal{F}_{N/4} \begin{pmatrix} x_2 \\ x_6 \\ x_{10} \\ \vdots \end{pmatrix} & & x_{N/4} & \\
 \mathcal{F}_N \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{pmatrix} & & & & & x_{3N/4} & \\
 & \mathcal{F}_{N/2} \begin{pmatrix} x_1 \\ x_3 \\ x_5 \\ \vdots \end{pmatrix} & \mathcal{F}_{N/4} \begin{pmatrix} x_1 \\ x_5 \\ x_9 \\ \vdots \end{pmatrix} & & & x_{N/8} & \\
 & & & & & x_{5N/8} & \\
 & & & & \mathcal{F}_{N/4} \begin{pmatrix} x_3 \\ x_7 \\ x_{11} \\ \vdots \end{pmatrix} & x_{3N/8} & \\
 & & & & & x_{7N/8} &
 \end{array}$$

Abb. 5.2: Schema für die schnelle Fouriertransformation für $N = 8$

Satz 5.10. (Rekursionsformeln)

Zu geradem N und gegebenen Folgen $u, v \in \mathbb{C}^{N/2}$ sei

$$x = [u_0 \ v_0 \ u_1 \ v_1 \ \dots \ u_{N/2-1} \ v_{N/2-1}] \in \mathbb{C}^N,$$

definiert. Dann gilt für $k = 0, 1, \dots, N/2 - 1$

$$\begin{aligned}
 (\mathcal{F}_N x)_k &= (\mathcal{F}_{N/2} u)_k + \omega_N^k (\mathcal{F}_{N/2} v)_k, \\
 (\mathcal{F}_N x)_{k+N/2} &= (\mathcal{F}_{N/2} u)_k - \omega_N^k (\mathcal{F}_{N/2} v)_k.
 \end{aligned}$$

Bemerkung. Sind $\mathcal{F}_{N/2} u$ und $\mathcal{F}_{N/2} v$ bekannt, benötigt man nur noch $N/2$ Multiplikationen und N Additionen/Subtraktionen zur Berechnung von \mathcal{F}_N .

0	000		000	0
1	001		100	4
2	010		010	2
3	011	Bitumkehr	110	6
4	100		001	1
5	101		101	5
6	110		011	3
7	111		111	7

Abb. 5.3: Bitumkehr der Binärdarstellung der Indizes.

Beweis. Nach Definition von \mathcal{F}_N ist für $k = 0, 1, \dots, N/2 - 1$

$$\begin{aligned}
 (\mathcal{F}_N x)_k &= \sum_{l=0}^{N-1} \omega_N^{kl} x_l \\
 &= \sum_{j=0}^{N/2-1} \omega_N^{k \cdot 2j} u_j + \sum_{j=0}^{N/2-1} \omega_N^{k \cdot (2j+1)} v_j \\
 &= \sum_{j=0}^{N/2-1} \omega_{N/2}^{kj} u_j + \omega_N^k \sum_{j=0}^{N/2-1} \omega_{N/2}^{kj} v_j \\
 &= (\mathcal{F}_{N/2} u)_k + \omega_N^k (\mathcal{F}_{N/2} v)_k.
 \end{aligned}$$

Wegen der Periodizität gilt

$$(\mathcal{F}_{N/2} u)_{k+N/2} = (\mathcal{F}_{N/2} u)_k$$

und ebenso für v . Die Behauptung folgt nun aus $\omega_N^{k+N/2} = -\omega_N^k$, denn $e^{-i\pi} = -1$. \square

Wir setzen jetzt $N = 2^L$, $L \in \mathbb{N}$ voraus. Die Berechnung von $\mathcal{F}_N x$ für $x \in \mathbb{C}^N$ kann dann rekursiv wie in Abbildung 5.2 dargestellt erfolgen. Diese Schema läßt sich leicht rekursiv programmieren.

Sollte eine rekursive Implementierung nicht möglich sein, so geht man in Abbildung 5.2 von rechts nach links vor. Hierbei muss man jedoch beachten, dass man die Komponenten des Vektors x in einer anderen als der ursprünglichen Reihenfolge benötigt. Wie aus Abbildung 5.3 zu erkennen, wird die Umsortierung durch Bitumkehr der Binärdarstellung der Indizes erreicht.

In MATLAB kann man dies zum Beispiel durch

```
function b = bitrev(N);
% function bitsrev = bitrev(N);
b = dec2bin([0:N-1]);
b = fliplr(b);
b = bin2dec(b);
```

erreichen.

Satz 5.11. $\mathcal{F}_N x$ kann für $N = 2^L$ mit $\frac{1}{2}N \log_2 N$ komplexen Multiplikationen und $N \log_2 N$ komplexen Additionen berechnet werden.

Beweis. Nach obigem Schema ist die Anzahl der Multiplikationen

$$\sum_{j=1}^L \frac{N}{2^j} 2^{j-1} = \frac{N}{2} L = \frac{N}{2} \log_2 N. \quad \square$$

Ein Vergleich mit der direkten Berechnung liefert

N	N^2	$N \log_2 N$	Quotient
$2^5 = 32$	10^3	160	6.4
$2^{10} \approx 10^3$	10^6	10^4	100
$2^{20} \approx 10^6$	10^{12}	$2 \cdot 10^7$	$5 \cdot 10^4$

Auf einem Rechner mit einer Leistung von 3300 MHz dauert eine schnelle Fouriertransformation mit $N = 2^{24} = 16.777.216$ mit der FFT-Funktion von Matlab nur etwa eine Sekunde.

Eine wichtige Anwendung der Fouriertransformation ist die Multiplikation von Folgen $x, y \in \mathbb{C}^N$. Mit $x \bullet y$ bezeichnen wir das punktweise Produkt von x und y , $(x \bullet y)_k := x_k y_k$. In den Anwendungen wichtiger ist die Faltung:

Definition 5.12. (*Faltung*)

Es seien $x, y \in \mathbb{C}^N$ (periodisch fortgesetzt). Die **Faltung** $x \star y \in \mathbb{C}^N$ ist definiert durch

$$(x \star y)_k = \sum_{j=0}^{N-1} x_{k-j} y_j.$$

Satz 5.13. (*Faltungssatz*) $\mathcal{F}_N(x \star y) = \mathcal{F}_N x \bullet \mathcal{F}_N y$.

Beweis. Nach Definition der Faltung ist

$$\begin{aligned} (\mathcal{F}_N(x \star y))_m &= \sum_{k=0}^{N-1} \omega^{mk} \sum_{j=0}^{N-1} x_{k-j} y_j \\ &= \sum_{l=k-j}^{N-1} \sum_{j=0}^{N-1-j} \omega^{m(l+j)} x_l y_j \\ &= \sum_{j=0}^{N-1} \omega^{mj} y_j \cdot \sum_{l=0}^{N-1} \omega^{ml} x_l \\ &= \hat{y}_m \hat{x}_m, \end{aligned}$$

wobei $\hat{x}_m = \mathcal{F}_N x$ und $\hat{y}_m = \mathcal{F}_N y$. □

Der Satz zeigt auch, dass die Faltung kommutativ und assoziativ ist.

Korollar 5.14. $x \star y = \frac{1}{N} \overline{\mathcal{F}_N} (\mathcal{F}_N x \bullet \mathcal{F}_N y)$.

Die direkte Berechnung der Faltung kostet $O(N^2)$ Operationen. Mit FFT kann die Berechnung in $O(N \log N)$ Operationen erfolgen.

Bemerkung. Die Berechnung der Faltung mit Hilfe von Korollar 5.14 ist ab $N = 32$ mit FFT schneller als direkt.

Häufig ist die betrachtete Folge reell. Dann kann der Aufwand zur Berechnung der Fouriertransformation durch folgende Symmetrieeigenschaft weiter reduziert werden:

Lemma 5.15. (Symmetrie)

- (a) Für
- $x = (x_0, x_1, \dots, x_{N-1}) \in \mathbb{R}^N$
- periodisch fortgesetzt gilt

$$\hat{x}_{-k} = \overline{\hat{x}_k} \quad \text{für } k \in \mathbb{Z}.$$

- (b) Falls
- $x \in \mathbb{C}^N$
- eine gerade Folge ist (d.h.,
- $x_{-k} = x_k$
- für alle
- $k \in \mathbb{Z}$
-), so ist auch die Fourier-Transformierte
- \hat{x}
- gerade. Falls
- x
- ungerade ist (d.h.,
- $x_{-k} = -x_k$
- für alle
- $k \in \mathbb{Z}$
-), so ist auch die Fourier-Transformierte
- \hat{x}
- ungerade.

Bemerkung. Eine gerade Folge hat für N gerade die Form

$$\begin{bmatrix} x_0 & x_1 & \cdots & x_{N/2-1} & | & x_{N/2} & \cdots & x_2 & x_1 \end{bmatrix}.$$

Bei einer ungeraden Folge ist $x_0 = -x_0$, also $x_0 = 0$ und ebenso $x_{N/2} = -x_{-N/2} = -x_{N/2} = 0$, woraus sich die Struktur

$$\begin{bmatrix} 0 & x_1 & \cdots & x_{N/2-1} & | & 0 & -x_{N/2-1} & \cdots & -x_1 \end{bmatrix}$$

ergibt.

Beweis. Übung. □

Wegen Lemma 5.15(a) kann man also die Fouriertransformation der reellen Vektoren x und y simultan berechnen, denn für $z = x + iy$ mit $x, y \in \mathbb{R}^N$ folgt aus

$$\hat{z}_k = \hat{x}_k + i\hat{y}_k, \quad k \in \mathbb{Z}$$

wegen $\overline{\hat{x}_{-k}} = \hat{x}_k$, $\overline{\hat{y}_{-k}} = \hat{y}_k$

$$\overline{\hat{z}_{-k}} = \overline{\hat{x}_{-k}} - i\overline{\hat{y}_{-k}} = \hat{x}_k - i\hat{y}_k$$

und daraus

$$\hat{x}_k = \frac{1}{2} \left(\hat{z}_k + \overline{\hat{z}_{-k}} \right), \quad \hat{y}_k = \frac{1}{2i} \left(\hat{z}_k - \overline{\hat{z}_{-k}} \right).$$

5.3 Inverses Faltungsproblem, Regularisierung, Filter

In der Bildverarbeitung stellt sich häufig folgendes Problem: Aus einem gemessenen Signal b , welches durch einen “Apparat” (z. B. eine Linse) aus einem (unbekannten) Eingangssignal u erzeugt wurde, soll das Eingangssignal u bestimmt werden.

Unter der Annahme, dass die Abbildung $u \mapsto b$ linear und verschiebungsinvariant ist, kann man den Apparat durch die Faltungsgleichung

$$\int_{-\infty}^{\infty} a(x-y)u(y)dy + \varepsilon(x) = b(x), \quad x \in \mathbb{R}$$

modellieren. Hierbei ist a die Apparatfunktion, u das unbekannte Eingangssignal, ε ein Modell für Störungen (“Rauschen”), die aus Mess-, Rundungs- oder Modellfehlern resultieren können, und b ist das beobachtete Signal. Die Störung ist ebenfalls unbekannt, bestenfalls weiß man

$$|\varepsilon(x)| \leq M \quad \text{oder} \quad \left(\int |\varepsilon(x)|^2 dx \right)^{1/2} \approx \delta \quad \text{Streuung.}$$

Im Allgemeinen haben a, u, b kompakten Träger ($\text{supp } a = \overline{\{x : a(x) \neq 0\}}$). Nach Variablentransformation nehmen wir ohne Einschränkung

$$\text{supp } a \subseteq \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \quad \text{supp } u \subseteq \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

an. Aus $b(x) - \varepsilon(x) \neq 0$ folgt, dass ein y existiert mit $a(x-y)u(y) \neq 0$. Nach obiger Annahme über die Träger ist das nur für $x-y \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ und $y \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ möglich, also folgt

$$\text{supp}(b - \varepsilon) \subseteq [-\pi, \pi]$$

und somit

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} a(x-y)u(y)dy + \varepsilon(x) = b(x), \quad x \in [-\pi, \pi].$$

Die Fourierreihe von u ist durch

$$u(x) = \sum_{n=-\infty}^{\infty} \hat{u}(n)e^{inx}, \quad \hat{u}(n) = \frac{1}{2\pi} \int_0^{2\pi} u(t)e^{-int}dt, \quad n \in \mathbb{Z}$$

gegeben, wobei $\hat{u}(n)$ der n -te Fourierkoeffizient ist

Setzen wir a, u, ε, b 2π -periodisch auf \mathbb{R} fort, so kann man mit dem kontinuierlichen Faltungsoperator

$$(a \star u)(t) = \frac{1}{2\pi} \int_0^{2\pi} a(t-\tau)u(\tau)d\tau$$

kurz

$$a \star u + \varepsilon = b \quad \text{bzw.} \quad Au + \varepsilon = b$$

mit dem Operator $Au = a \star u$ schreiben.

Der übliche Lösungsweg wäre, die Störung zu vernachlässigen und das System $Av = b$ zu lösen. Der Fehler $\|v - u\| = \|A^{-1}\varepsilon\|$ ist akzeptabel, falls $\|A^{-1}\varepsilon\| \ll \|v\|$ gilt.

Wir wollen nun genauer betrachten, wie sich in unserem Fall Störungen der Daten auf die Lösung auswirken. Mit dem kontinuierlichen Faltungssatz

$$\widehat{(a \star u)}(n) = \hat{a}(n)\hat{u}(n)$$

(entsprechend dem diskreten Faltungssatz) gilt

$$\hat{a}(n)\hat{u}(n) + \hat{\varepsilon}(n) = \hat{b}(n) \iff \hat{u}(n) = \frac{\hat{b}(n)}{\hat{a}(n)} - \frac{\hat{\varepsilon}(n)}{\hat{a}(n)} =: \hat{v}(n) - \frac{\hat{\varepsilon}(n)}{\hat{a}(n)},$$

wobei $\hat{v}(n)$ Lösung von $\hat{a}(n)\hat{v}(n) = \hat{b}(n)$ ist.

Das Rauschen ε enthält im Allgemeinen hochfrequente Anteile, $|\hat{\varepsilon}(n)| \approx \delta$ auch für große n . Ist a eine glatte Funktion, dann kann man zeigen, dass die Fourierkoeffizienten $|\hat{a}(n)|$ für $|n| \rightarrow \infty$ schnell abfallen (je glatter a ist, desto schneller). Im Allgemeinen wird also $|\hat{\varepsilon}(n)|/|\hat{a}(n)|$ groß für wachsendes $|n|$:

$$u(x) = \sum_{n=-\infty}^{\infty} \hat{u}(n)e^{inx} = \sum_{n=-\infty}^{\infty} \hat{v}(n)e^{inx} - \sum_{n=-\infty}^{\infty} \frac{\hat{\varepsilon}(n)}{\hat{a}(n)}e^{inx},$$

was zu einer katastrophalen Fehlerverstärkung führt (*ill-posed problem*).

Um dies zu vermeiden, kann man verschiedene Ansätze wählen. Die erste Überlegung ist, dass bei glattem Eingangssignal u die Fourierkoeffizienten von u für wachsendes $|n|$ rasch klein werden. Daher kann man die Fourierentwicklung einfach abschneiden und

$$u_M(x) = \sum_{n=-M}^M \frac{\widehat{b}(n)}{\widehat{a}(n)} e^{inx}$$

statt u verwenden. Das Abschneiden entspricht einer Filterung:

$$u_M(x) = \sum_{n=-\infty}^{\infty} \phi_M(n) \frac{\widehat{b}(n)}{\widehat{a}(n)} e^{inx}$$

mit

$$\phi_M(n) = \begin{cases} 1 & |n| \leq M, \\ 0 & \text{sonst.} \end{cases}$$

ϕ_M bezeichnet man als Tiefpassfilter, denn er lässt die niedrigen Frequenzen “passieren” und filtert die hohen heraus.

Ein anderer Zugang beruht darauf, dass man gar nicht $Au = b$ lösen möchte, sondern sinnvollerweise nur

$$\|Au - b\| \leq \|\varepsilon\|$$

fordern kann, wobei $\|\cdot\|$ die (skalierte) L^2 -Norm ist:

$$\|f\| = \left(\frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx \right)^{1/2} = \|f\|_{L^2}.$$

Ist bekannt, dass $\|\varepsilon\| \approx \delta$ gilt, so fordern wir

$$\|Au - b\| \leq \delta.$$

Im Allgemeinen wird es viele Signale u geben, die dies erfüllen, und unter all diesen wählen wir jenes mit

$$\|u''\| = \min \quad \text{oder} \quad \|u\| = \min,$$

also ein möglichst glattes (vergleiche mit einem Spline) oder ein möglichst kleines u . Allgemein ergibt sich das Minimierungsproblem

$$\begin{aligned} \|Lu\| &= \min & L \text{ linear,} \\ \|Au - b\| &\leq \delta & \text{Nebenbedingung.} \end{aligned} \tag{5.4}$$

Im Folgenden nehmen wir $\|b\| > \delta$ an. Dies ist eine sinnvolle Annahme, da sonst das beobachtete Signal nicht vom Rauschen unterschieden werden kann.

Lemma 5.16. *Es sei $\|b\| > \delta$. Dann gilt für die Lösung u von (5.4)*

$$u \neq 0, \quad \|Au - b\| = \delta.$$

Beweis. $u \neq 0$ ist klar. Angenommen, $\|Lu\| = \min$ für ein u mit $\|Au - b\| = \delta - \epsilon$, wobei $0 < \epsilon \leq \delta$. Dann definieren wir für ein $0 < \rho < 1$

$$\tilde{u} = (1 - \rho)u.$$

Hierfür gilt $\|L\tilde{u}\| = (1 - \rho)\|Lu\| < \|Lu\|$ und

$$\begin{aligned}\|A\tilde{u} - b\| &= \|(1 - \rho)(Au - b) - \rho b\| \\ &\leq (1 - \rho)\|Au - b\| + \rho\|b\| \\ &= (1 - \rho)(\delta - \epsilon) + \rho\|b\| \\ &= \delta - \epsilon + \rho(\|b\| - \delta + \epsilon)\end{aligned}$$

Für $0 < \rho < 1$ so klein, dass $\rho(\|b\| - \delta + \epsilon) < \epsilon$ gilt, folgt dann $\|A\tilde{u} - b\| \leq \delta$. \tilde{u} erfüllt also die Nebenbedingung. Dies steht im Widerspruch dazu, dass u das Minimierungsproblem (5.4) löst. \square

Für die praktische Implementierung verwendet man die

Tychonoff-Regularisierung

Für festes $\alpha > 0$ (Regularisierungsparameter) löse das Minimierungsproblem

$$\|Au - b\|^2 + \alpha\|Lu\|^2 = \min! \quad (5.5)$$

ohne Nebenbedingung. Im Grenzfall reduziert sich das Minimierungsproblem zu

$$\begin{array}{lll}\alpha \rightarrow 0 : & \|Au - b\| = 0, & \|Lu\| \text{ bel. groß,} \\ \alpha \rightarrow \infty : & \|Lu\| = 0, & \|Au - b\| \text{ bel. groß.}\end{array}$$

Das optimale α für das Minimierungsproblem (5.4) wäre nach Lemma 5.16 jenes, für das $\|Au - b\| \approx \|\epsilon\|$. Im Allgemeinen wird $\|\epsilon\|$ nicht bekannt sein. Dann wählt man α so, dass die Lösung von (5.5) "optisch brauchbar" ist.

Im endlichdimensionalen Fall hängen die Minimierungsprobleme (5.4) und (5.5) eng zusammen. Es seien $A, L \in \mathbb{R}^{m,n}$, $b \in \mathbb{R}^m$, $u \in \mathbb{R}^n$ und $\|\cdot\|$ sei die Euklidnorm. Dann gilt

$$\begin{aligned}\|Lu\|^2 &= (Lu)^T(Lu) = u^T L^T L u \\ \|Au - b\|^2 &= (Au - b)^T(Au - b) = u^T A^T A u - 2u^T A^T b + b^T b\end{aligned}$$

Die allgemeine Lösung von

$$f(u) = \min, \quad g(u) = 0$$

erfüllt

$$f'(u) + g'(u)\lambda = 0, \quad g(u) = 0,$$

wobei λ der Lagrange-Multiplikator ist. In unserem Spezialfall (5.4) gilt also

$$2L^T L u + (2A^T A u - 2A^T b)\lambda = 0, \quad \|Au - b\|^2 - \delta^2 = 0$$

oder

$$(L^T L + \lambda A^T A)u = \lambda A^T b.$$

Für festes λ ist dies die Lösung des Minimierungsproblems

$$\|Lu\|^2 + \lambda\|Au - b\|^2 = \min$$

oder mit $\alpha = \lambda^{-1}$ die Lösung von (5.5).

Angewandt auf $Au = a \star u$ und $Lu = u^{(p)}$ gilt

Satz 5.17. Für festes $\alpha > 0$ ist die Lösung des Minimierungsproblems

$$\|a \star u - b\|_{L^2}^2 + \alpha \|u^{(p)}\|_{L^2}^2 = \min, \quad (5.6)$$

(für u 2π -periodisch mit $u^{(p)} \in L^2$) gegeben durch die Fourierkoeffizienten

$$\widehat{u}(n) = \begin{cases} \frac{|\widehat{a}(n)|^2}{|\widehat{a}(n)|^2 + \alpha n^{2p}} \frac{\widehat{b}(n)}{\widehat{a}(n)} =: \phi_\alpha(n) \frac{\widehat{b}(n)}{\widehat{a}(n)}, & \widehat{a}(n) \neq 0, \\ 0 & \widehat{a}(n) = 0. \end{cases}$$

ϕ_α heißt *Regularisierungsfiler*.

Beweis. Nach der Parseval'schen Gleichung ist (5.6) äquivalent zu

$$\sum_{n \in \mathbb{Z}} \left(|\widehat{a}(n)\widehat{u}(n) - \widehat{b}(n)|^2 + \alpha |(in)^p \widehat{u}(n)|^2 \right) = \min,$$

denn der n -te Fourierkoeffizient von $u^{(p)}$ ist $(in)^p \widehat{u}(n)$. Die Summe wird minimal, wenn jeder einzelne Summand minimal wird. Der n -te Summand ist von der Form

$$\begin{aligned} & |\widehat{a}(n)|^2 |\widehat{u}(n)|^2 - 2 \operatorname{Re} \left(\overline{\widehat{a}(n)\widehat{u}(n)} \widehat{b}(n) \right) + |\widehat{b}(n)|^2 + \alpha n^{2p} |\widehat{u}(n)|^2 \\ &= (|\widehat{a}(n)|^2 + \alpha n^{2p}) |\widehat{u}(n)|^2 - 2 \operatorname{Re} \left(\overline{\widehat{a}(n)\widehat{u}(n)} \widehat{b}(n) \right) + |\widehat{b}(n)|^2 \\ &=: r |z|^2 - 2 \operatorname{Re} \bar{z} s + |\widehat{b}(n)|^2, \quad r > 0, \end{aligned}$$

mit $z = \widehat{u}(n)$ und $s = \overline{\widehat{a}(n)} \widehat{b}(n)$. Mit $q = s/r$ ist die Minimierung des n -ten Summanden äquivalent zur Minimierung von

$$|z|^2 - 2 \operatorname{Re}(\bar{z}q) \geq |z|^2 - 2|z| \cdot |q| \geq -|q|^2.$$

In dieser Ungleichungskette gilt Gleichheit für $z = q = s/r$, so dass das Minimum für $z = q$ angenommen wird. Somit ist die Lösung von (5.6) durch

$$\widehat{u}(n) = \frac{\overline{\widehat{a}(n)} \widehat{b}(n)}{|\widehat{a}(n)|^2 + \alpha n^{2p}}$$

gegeben. □

Korollar 5.18. Es sei u_α die Lösung von (5.6) zu gegebenem α . Dann ist

$$\begin{aligned} \alpha \mapsto \|a \star u_\alpha - b\|_{L^2} & \quad \text{monoton wachsend,} \\ \alpha \mapsto \|u_\alpha^{(p)}\|_{L^2} & \quad \text{monoton fallend.} \end{aligned}$$

Beweis. Nach der Parseval'schen Gleichung ist

$$\begin{aligned} \|a \star u_\alpha - b\|_{L^2}^2 &= \sum_{n \in \mathbb{Z}} |\widehat{a}(n)\widehat{u}_\alpha(n) - \widehat{b}(n)|^2 \\ &= \sum_{n \in \mathbb{Z}} (1 - \phi_\alpha(n))^2 |\widehat{b}(n)|^2. \end{aligned}$$

Die Abbildung

$$\alpha \mapsto 1 - \phi_\alpha(n) = \frac{\alpha n^{2p}}{|\widehat{a}(n)|^2 + \alpha n^{2p}}$$

ist für alle n monoton wachsend (betrachte $x \mapsto x/(c+x)$ für $c > 0$).

Ferner ist wiederum nach der Parseval'schen Gleichung

$$\|u_\alpha^{(p)}\|_{L^2}^2 = \sum_{n \in \mathbb{Z}} |(in)^p \widehat{u}_\alpha(n)|^2 = \sum_{n \in \mathbb{Z}} (n^p \phi_\alpha(n))^2 \frac{|\widehat{b}(n)|^2}{|\widehat{a}(n)|^2}.$$

Da die Abbildung $\alpha \mapsto \phi_\alpha(n)$ monoton fallend ist, ist damit auch der zweite Teil der Behauptung gezeigt. \square

Nun wird im Allgemeinen b nur in diskreten Punkten gemessen werden, sagen wir in $x_j = j \frac{2\pi}{N}$. Die Nebenbedingung in (5.4) ist dann so nicht mehr sinnvoll und wird durch

$$\frac{1}{N} \sum_{j=0}^{N-1} |(a \star u)(x_j) - b(x_j)|^2 \leq \delta^2$$

ersetzt.

Lemma 5.19. *Es seien*

$$u_N(x) = \sum_{n=-N/2}^{N/2-1} \widehat{u}_N(n) e^{inx}, \quad b_N(x) = \sum_{n=-N/2}^{N/2-1} \widehat{b}_N(n) e^{inx}$$

trigonometrische Interpolationspolynome der 2π -periodischen Funktionen u und b . Dann ist

$$\|a \star u_N - b_N\|_{L^2}^2 = \frac{1}{N} \sum_{j=0}^{N-1} |(a \star u_N)(x_j) - b(x_j)|^2.$$

Beweis. Zunächst liefert die (kontinuierliche) Parseval'sche Gleichung

$$\|a \star u_N - b_N\|_{L^2}^2 = \sum_{n=-N/2}^{N/2-1} |\widehat{a}(n) \widehat{u}_N(n) - \widehat{b}_N(n)|^2,$$

da nur diese Fourierkoeffizienten von \widehat{u} und \widehat{b} von Null verschieden sind. Wegen der Interpolationseigenschaft gilt

$$(a \star u_N)(x_j) = \sum_{n=-N/2}^{N/2-1} \widehat{a}(n) \widehat{u}_N(n) \underbrace{e^{inx_j}}_{\overline{\omega}_N^{nj}} = (-1)^j (\overline{\mathcal{F}}_N(\widehat{a} \widehat{u}_N))_j$$

und analog

$$b(x_j) = b_N(x_j) = \sum_{n=-N/2}^{N/2-1} \widehat{b}_N(n) e^{inx_j} = (-1)^j (\overline{\mathcal{F}}_N(\widehat{b}_N))_j$$

Die Behauptung folgt nun direkt aus der (diskreten) Parseval'schen Gleichung in Satz 5.4. \square

Daher gilt

Satz 5.20. *Unter allen trigonometrischen Polynomen*

$$u_N(x) = \sum_{n=-N/2}^{N/2-1} \hat{u}_N(n) e^{inx}$$

ist u_N die Lösung von (5.6) genau dann, wenn für $n = -N/2, \dots, N/2 - 1$ gilt

$$\begin{cases} \frac{|\hat{a}(n)|^2}{|\hat{a}(n)|^2 + \alpha n^{2p}} \frac{\hat{b}(n)}{\hat{a}(n)} =: \phi_\alpha(n) \frac{\hat{b}(n)}{\hat{a}(n)}, & \hat{a}(n) \neq 0, \\ 0 & \hat{a}(n) = 0. \end{cases}$$

Beweis. Wie für Satz 5.17. □

5.4 Schnelle Poisson-Löser

Als weitere Anwendung der schnellen Fouriertransformation betrachten wir folgendes Problem: Gesucht ist eine Funktion $u : \Omega \rightarrow \mathbb{R}$, $\Omega \subset \mathbb{R}^2$ mit

$$\begin{array}{lll} -\Delta u = f & \text{in } \Omega & \text{Poissongleichung,} \\ u = v & \text{auf } \partial\Omega & \text{Dirichlet-Randbedingungen,} \end{array}$$

wobei $\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$ der Laplaceoperator in \mathbb{R}^2 ist.

Die Lösung dieser Gleichung ist ein fundamentales Problem der Potentialtheorie, welche Anwendungen in der Wärmeleitung (u Temperatur, f Wärmequellendichte), der Strömungslehre, der Elektro- und Magnetostatik, usw. hat.

Die Diskretisierung der Poissongleichung mittels finiter Differenzen führt auf ein endlich-dimensionales lineares Gleichungssystem hoher Dimension. Ziel dieses Abschnitts ist es zu zeigen, dass solche Gleichungssysteme mit FFT schnell gelöst werden können, wenn Ω ein Rechteck ist.

Allgemeinere Gebiete kann man durch Gebietszerlegung und anschließende Iteration lösen oder man kann das Gebiet konform auf ein Rechteck abbilden, dort schnell lösen und die Lösung mit der inversen Abbildung wieder auf das Ausgangsgebiet transformieren.

Betrachten wir die Diskretisierung mit finiten Differenzen. Es sei $\Omega = (0, Mh) \times (0, Nh)$, wobei h die Gitterweite ist. Mit

$$u_{mn} \approx u(mh, nh)$$

bezeichnen wir die Näherungslösungen an den Gitterpunkten. Am Rand des Gitters ($m = 0$ oder M und $n = 0$ oder N) sind die Werte durch die Randbedingung gegeben. Im Inneren ersetzen wir $-\Delta u$ durch Differenzenquotienten:

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2}(mh, nh) &\approx \frac{u_{m-1,n} - 2u_{m,n} + u_{m+1,n}}{h^2} \\ \frac{\partial^2 u}{\partial y^2}(mh, nh) &\approx \frac{u_{m,n-1} - 2u_{m,n} + u_{m,n+1}}{h^2} \end{aligned}$$

Die diskretisierte Gleichung lautet dann

$$\frac{1}{h^2} (4u_{m,n} - u_{m-1,n} - u_{m+1,n} - u_{m,n-1} - u_{m,n+1}) = f(mh, nh) \quad \text{in } \Omega. \quad (5.7)$$

oder als Differenzenstern

$$\begin{array}{ccc} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{array}$$

Auf dem Rand gilt

$$u_{m,n} = v(mh, nh) \quad \text{für } (mh, nh) \in \partial\Omega.$$

Daraus ergibt sich für die unbekannten Werte

$$u_{m,n}, \quad m = 1, \dots, M-1, \quad n = 1, \dots, N-1$$

das lineare Gleichungssystem der Dimension $(M-1)(N-1)$:

$$\begin{bmatrix} 4 & -1 & 0 & -1 & & & \\ -1 & 4 & -1 & & & & \\ 0 & -1 & 4 & & & \ddots & \\ -1 & & & 4 & -1 & 0 & \ddots \\ & -1 & & -1 & 4 & -1 & \ddots \\ & & \ddots & 0 & -1 & 4 & -1 \\ & & & \ddots & & 4 & -1 & 0 \\ & & & & \ddots & -1 & 4 & -1 \\ & & & & & -1 & 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{21} \\ u_{31} \\ u_{12} \\ u_{22} \\ u_{32} \\ u_{13} \\ u_{23} \\ u_{33} \end{bmatrix} = h^2 \begin{bmatrix} f_{11} \\ f_{21} \\ f_{31} \\ f_{12} \\ f_{22} \\ f_{32} \\ f_{13} \\ f_{23} \\ f_{33} \end{bmatrix} + \begin{bmatrix} v_{01} + v_{10} \\ v_{20} \\ v_{30} + v_{41} \\ v_{02} \\ 0 \\ v_{42} \\ v_{03} + v_{14} \\ v_{24} \\ v_{43} + v_{34} \end{bmatrix}$$

Wir werden zeigen, dass sich dieses Gleichungssystem mit Hilfe von FFT in $O(MN \log M)$ Operationen lösen lässt.

Zur Vorbereitung betrachten wir das eindimensionale Problem:

$$\begin{array}{lll} -u'' = f & \text{in } \Omega = (0, 1) & \text{Poissongleichung,} \\ u(x) = v(x) & \text{für } x \in \{0, 1\} & \text{Dirichlet-Randbedingungen,} \end{array}$$

welches nach Diskretisierung die Form

$$-u_{n+1} + 2u_n - u_{n-1} = h^2 f(nh), \quad n = 1, \dots, N-1$$

mit den Randwerten $u_0 = v(0)$ und $u_N = v(1)$ hat. Dies führt auf ein lineares Gleichungssystem mit tridiagonaler Koeffizientenmatrix:

$$\begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{N-2} \\ u_{N-1} \end{bmatrix} = \begin{bmatrix} h^2 f_1 + v(0) \\ h^2 f_2 \\ \vdots \\ h^2 f_{N-2} \\ h^2 f_{N-1} + v(1) \end{bmatrix} =: \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{N-2} \\ b_{N-1} \end{bmatrix}$$

Die Tridiagonalmatrix ist symmetrisch und positiv definit. Das Gleichungssystem lässt sich durch Gauß-Elimination unter Ausnutzung der Bandstruktur wie in Algorithmus 5.1 angegeben lösen. Der Aufwand für ein System der Dimension n beträgt $5n$ Multiplikationen und $3n$ Additionen für beliebige α, β und $3n$ Multiplikationen und Additionen im Spezialfall $\beta_j = -1$.

Algorithmus 5.1 Lösung des Tridiagonalsystems $\text{tridiag}(\beta, \alpha, \beta)u = b$ {Hauptdiagonale: $\alpha_1, \dots, \alpha_n$, Nebendiagonale: $\beta_1, \dots, \beta_{n-1}$ }{Lösung überschreibt b }**for** $k = 2 : n$ **do**

$$t = \beta_{k-1}, \beta_{k-1} = \frac{t}{\alpha_{k-1}}, \alpha_k = \alpha_k - t\beta_{k-1}$$

$$b_k = b_k - \beta_{k-1}b_{k-1}$$

end for

$$b_n = \frac{b_n}{\alpha_n}$$

for $k = n - 1 : -1 : 1$ **do**

$$b_k = \frac{b_k}{\alpha_k} - \beta_k b_{k+1}$$

end for

Beim zweidimensionalen Problem hat das Gleichungssystem die Bandbreite $M - 1$. Die Lösung durch Gauß-Elimination würde also $O(M^2 \cdot NM)$ Operationen kosten. Die Grundidee eines schnellen Löser basiert darauf, die Gleichungen als Faltung zu interpretieren. In (5.7) tritt der Ausdruck

$$-u_{m-1,n} + 4u_{m,n} - u_{m+1,n} \quad (5.8)$$

auf, welcher Teil einer Faltung bezüglich der ersten Variablen ist. Wie in obigem Gleichungssystem schieben wir die Randbedingungen in die rechte Seite der Differenzgleichung. Daraus ergibt sich

$$4u_{m,n} - u_{m-1,n} - u_{m+1,n} - u_{m,n-1} - u_{m,n+1} = b_{m,n} \quad m = 1, \dots, M - 1 \quad (5.9)$$

wenn wir

$$u_{0,n} = u_{M,n} = 0, \quad b_{0,n} = b_{M,n} = 0$$

setzen. Wir erweitern jetzt $u_n := u_{\cdot,n}$ und $b_n := b_{\cdot,n}$ so, dass (5.9) sogar für alle $m \in \mathbb{Z}$ gilt. Für $m = 0$ folgt wegen der Randbedingungen

$$-u_{-1,n} - u_{1,n} = 0, \quad n = 1, \dots, N - 1$$

und mit den so definierten $u_{-1,n}$, $n = 1, \dots, N - 1$ folgt für $m = -1$ und $b_{-1,n} := -b_{1,n}$

$$-4u_{1,n} - u_{-2,n} + u_{1,n-1} + u_{1,n+1} = -b_{1,n}, \quad n = 1, \dots, N - 1.$$

Diese Gleichung entspricht für $u_{-2,n} = u_{2,n}$ gerade (5.9) für $m = 1$. Allgemein gilt (5.9) also, wenn wir die Vektoren u_n und b_n zunächst ungerade

$$u_{-m,n} = -u_{m,n}, \quad b_{-m,n} = -b_{m,n}, \quad m = 1, \dots, M - 1$$

und dann $2M$ -periodisch fortsetzen. Dann entspricht (5.8) der Faltung

$$u_{\cdot,n} * d, \quad d_{-1} = d_1 = -1, \quad d_0 = 4, \quad d_j = 0 \text{ sonst,}$$

denn mit dieser Folge d gilt

$$(u_{\cdot,n} * d)_k = \sum_{j=-M+1}^{M-1} u_{k-j,n} d_j = u_{k+1,n} d_{-1} + u_{k,n} d_0 + u_{k-1,n} d_1.$$

Damit lautet (5.9)

$$-u_{n+1} + d * u_n - u_{n-1} = b_n.$$

Diese Gleichung transformieren wir mit \mathcal{F}_{2M} :

$$-\widehat{u}_{n+1} + \widehat{d} \bullet \widehat{u}_n - \widehat{u}_{n-1} = \widehat{b}_n, \quad n = 1, \dots, N-1, \quad \widehat{u}_0 = \widehat{u}_N = 0$$

oder komponentenweise für $k = 0, 1, \dots, 2M-1$

$$-(\widehat{u}_{n+1})_k + \widehat{d}_k (\widehat{u}_n)_k - (\widehat{u}_{n-1})_k = (\widehat{b}_n)_k, \quad n = 1, \dots, N-1 \quad (5.10)$$

Hierzu sind $N-1$ Fouriertransformationen der Länge $2M$ erforderlich. Der Aufwand hierfür ist nach Satz 5.11 in $O(2MN \log(2M))$.

Die Fourierkoeffizienten von d_k lassen sich leicht berechnen:

$$\widehat{d}_k = \sum_{j=0}^{2M-1} \omega_{2M}^{jk} d_j = \sum_{j=-M}^{M-1} \omega_{2M}^{jk} d_j = -\omega_{2M}^{-k} + 4\omega_{2M}^0 - \omega_{2M}^k = 4 - 2 \cos \frac{\pi k}{M}.$$

Für jedes einzelne k entspricht (5.10) einem Tridiagonalsystem der Dimension $N-1$ mit positiv definiter Matrix

$$\begin{bmatrix} \widehat{d}_k & -1 & & & \\ -1 & \widehat{d}_k & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & \widehat{d}_k & -1 \\ & & & -1 & \widehat{d}_k \end{bmatrix} \begin{bmatrix} (\widehat{u}_1)_k \\ (\widehat{u}_2)_k \\ \vdots \\ (\widehat{u}_{N-2})_k \\ (\widehat{u}_{N-1})_k \end{bmatrix} = \begin{bmatrix} (\widehat{b}_1)_k \\ (\widehat{b}_2)_k \\ \vdots \\ (\widehat{b}_{N-2})_k \\ (\widehat{b}_{N-1})_k \end{bmatrix}.$$

Der Aufwand zur Lösung dieser $2M$ Tridiagonalsysteme mit Algorithmus 5.1 ist $6NM$. Zuletzt transformieren wir mit der inversen Fouriertransformation

$$u_n = \frac{1}{2M} \overline{\mathcal{F}_{2M}} \widehat{u}_n, \quad n = 1, \dots, N-1,$$

in $O(MN \log M)$ Operationen zurück.

Zur weiteren Reduktion des Rechenaufwandes kann man Lemma 5.15 ausnutzen, indem man – wie in der Bemerkung nach dem Lemma erläutert – die Fouriertransformation zweier reeller Vektoren simultan berechnet.

Da die Lösungen \widehat{u}_n des Tridiagonalsystems ungerade sind, genügt es, diese Gleichungssysteme für $k = 0, \dots, M$ statt für $k = 0, \dots, 2M-1$ zu lösen.

Bei der Rücktransformation kann man wieder \widehat{u}_n und \widehat{u}_{n+1} simultan behandeln:

$$\mathcal{F}_{2N}^{-1} (\widehat{u}_n + i\widehat{u}_{n+1}) = u_n + iu_{n+1}.$$

Bemerkung. Mit Mehrgitterverfahren kann die Lösung in $O(MN)$ (zum Beispiel $50MN$) berechnet werden. Mehrgitterverfahren sind auch für allgemeine Gebiete Ω anwendbar.

Die Idee zur Konstruktion eines schnellen Poissonlösers lässt sich auch auf Neumannrandbedingungen anwenden:

$$\begin{aligned} -\Delta u + u &= f & \text{in } \Omega &= (0, 1) \times (0, 1) \\ \frac{\partial u}{\partial n} &= 0 & \text{auf } \partial\Omega & \quad (\text{Neumannrandbedingung}). \end{aligned}$$

Die Randbedingung wird durch

$$u_{1j} = u_{0j} \quad (j = 1, \dots, N-1) \quad \text{auf dem Randteil } x = 0$$

und analog auf dem Rest des Randes diskretisiert.

