7. Approximation periodischer Funktionen

Fourier–Reihen.

Wir betrachten wieder den Raum aller stetigen, $2\,\pi$-periodischen Funktionen

$$\mathrm{C}_{2\pi} \;:=\; \{f \in \mathrm{C}(\mathbb{R}):\ \forall\, t \in \mathbb{R}:\ f(t+2\,\pi) = f(t)\} \tag{7.1}$$

mit dem Standard-Skalarprodukt

$$\langle f, g\rangle \;:=\; \frac{1}{\pi}\int_0^{2\pi} f(t)\, g(t)\, dt. \tag{7.2}$$

Die zugehörige Norm auf $\mathrm{C}_{2\pi}$ werde wieder mit $\|\cdot\|$ oder besser $\|\cdot\|_2$ bezeichnet.

Der folgende Satz fasst nochmals das Beipiel (6.14) zusammen

Satz (7.3)

Die Funktionen $\dfrac{1}{\sqrt{2}},\ \cos t,\ \sin t, \ldots,\ \cos(n\,t),\ \sin(n\,t)$ bilden ein Orthonormalsystem bzgl. $\langle\cdot,\cdot\rangle$. Sie sind somit zugleich eine ONB des von ihnen aufgespannten linearen Teilraumes

$$\mathrm{T}_n \;:=\; \{f \in \mathrm{C}_{2\pi}:\ f(t) = \frac{a_0}{2} + \sum_{k=1}^{n} [a_k \cos(kt) + b_k \sin(kt)],\ a_k, b_k \in \mathbb{R}\}. \tag{7.4}$$

Die L_2–Bestapproximation einer Funktion $f \in \mathrm{C}_{2\pi}$ aus T_n lautet daher

$$\begin{aligned} S_n(f)(t) \;&=\; f_n(t) \;=\; \frac{a_0}{2} + \sum_{k=1}^{n} [a_k \cos(kt) + b_k \sin(kt)] \\ a_k \;&=\; \frac{1}{\pi}\int_0^{2\pi} f(t)\,\cos(kt)\,dt, \qquad b_k \;=\; \frac{1}{\pi}\int_0^{2\pi} f(t)\,\sin(kt)\,dt. \end{aligned} \tag{7.5}$$

Für $n \to \infty$ erhält man hieraus formal die *Fourier-Entwicklung* einer Funktion $f \in \mathrm{C}_{2\pi}$ (Jean-Baptiste-Joseph Fourier; 1768–1830)

$$f(t) \;\sim\; \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos(kt) + b_k \sin(kt)]. \tag{7.6}$$

Wir fragen nach der Konvergenz dieser *Fourier-Reihe*, nach der Konvergenzgeschwindigkeit und untersuchen, unter welchen Voraussetzungen in (7.6) punktweise bzw. gleichmäßige Konvergenz vorliegt.

Zunächst können wir analog zu Satz (6.33) die L_2-Konvergenz der Fourier-Reihe feststellen.

Satz (7.7)

Zu $f \in C_{2\pi}$ bezeichne $T_n(f)$ die (bzw. eine) Bestapproximation von f aus T_n bezüglich $\|\cdot\|_\infty$. Dann gelten

a) $$\|f - T_n(f)\|_\infty \to 0 \qquad (n \to \infty),$$

b) $$\|f - T_n(f)\|_2 \to 0 \qquad (n \to \infty),$$

c) $$\|f - S_n(f)\|_2 \to 0 \qquad (n \to \infty).$$

Beweis: Der Beweis erfolgt analog zu dem von Satz (6.33). Teil a) ergibt sich aus dem zweiten Weierstraßschen Approximationssatz (4.20).
Teil b) folgt aus der Abschätzung

$$\|f - T_n(f)\|_2 \;\le\; \sqrt{2}\, \|f - T_n(f)\|_\infty .$$

Teil c) folgt schließlich aus der Minimaleigenschaft von $S_n(f)$, nämlich
$\|f - S_n(f)\|_2 \le \|f - T_n(f)\|_2$. □

Erinnert sei auch an die *Parseval Gleichung* (6.13), die sich aus der Dichtheit der trigonometrischen Polynome in $C_{2\pi}$ ergibt,

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty} \left(a_k^2 + b_k^2\right) \;=\; \|f\|_2^2. \tag{7.8}$$

Aus (7.8) folgt insbesondere, dass die Fourier-Koeffizienten $a_k,\ b_k$ Nullfolgen bilden.

Nun zur Untersuchung der punktweisen bzw. gleichmäßigen Konvergenz der Fourier-Reihe. Zunächst ist klar, dass Satz (7.7) lediglich die *Konvergenz im quadratischen Mittel* zeigt, woraus nicht auf die punktweise Konvergenz geschlossen werden kann. Tatsächlich hat Paul du Bois-Reymond (1831–1889) eine stetige Funktion $f \in C_{2\pi}$ angegeben, deren Fourier-Reihe in mindestens einem Punkt divergiert.

Die Jackson–Sätze.

Ausgangspunkt für die folgenden Untersuchungen ist das Lemma von Lebesgue (3.4). Hierbei ist zu beachten, dass $S_n : C_{2\pi} \to T_n$ ein stetiger linearer Projektor ist. Die Stetigkeit ergibt sich beispielsweise aus der Integraldarstellung (vgl. (4.16))

$$S_n(f)(t) \;=\; \frac{1}{\pi}\int_0^{2\pi} f(t+\theta)\, D_n(\theta)\, d\theta, \qquad D_n(\theta) \;:=\; \frac{\sin[(n+1/2)\theta]}{2\,\sin[\theta/2]}. \tag{7.9}$$

Das Lemma von Lebesgue ist also anwendbar und lautet hier konkret

$$\|f - S_n(f)\|_\infty \;\le\; \big(1 + \|S_n\|_\infty\big)\, E_n(f), \qquad E_n(f) \;:=\; \inf_{p\in \mathrm{T}_n} \|f - p\,\|_\infty. \tag{7.10}$$

Dabei bezeichnet $E_n(f)$ den Minimalabstand von f zu T_n bezüglich der Maximumsnorm.

Wir beginnen mit einer Abschätzung für $\|S_n\|_\infty$.

Satz (7.11)

a) $$\|S_n\|_\infty \;=\; \frac{1}{\pi}\int_0^\pi \Big|\,\frac{\sin[(n+1/2)\,\theta\,]}{\sin(\theta/2)}\,\Big|\,d\,\theta$$

b) $$\frac{4}{\pi^2}\;\ln(1+n) \;\le\; \|S_n\|_\infty \;\le\; 1 \;+\; \ln(2\,n+1).$$

Beweis:

zu a) Dies folgt direkt aus der Integraldarstellung (7.9). Man beachte, dass $D_n(\theta)$ eine gerade, $2\,\pi$-periodische Funktion ist.

zu b) Seien $\theta_k := (k\,\pi)/(n+1/2)$ die Nullstellen von $D_n(\theta)$. Damit gilt

$$\begin{aligned}
\|S_n\|_\infty \;&\ge\; \frac{1}{\pi}\sum_{k=0}^{n-1}\int_{\theta_k}^{\theta_{k+1}} \Big|\,\frac{\sin[(n+1/2)\,\theta\,]}{\theta/2}\,\Big|\,d\,\theta \\
&\ge\; \frac{2}{\pi}\sum_{k=0}^{n-1}\frac{1}{\theta_{k+1}}\int_{\theta_k}^{\theta_{k+1}} |\,\sin[(n+1/2)\,\theta\,]\,|\,d\,\theta \\
&=\; \frac{4}{\pi^2}\sum_{k=0}^{n-1}\frac{1}{k+1} \;\ge\; \frac{4}{\pi^2}\;\ln(n+1).
\end{aligned}$$

Für die rechte Seite verwenden wir die folgenden beiden Abschätzungen

$$\Big|\,\frac{\sin[(n+1/2)\,\theta\,]}{\sin(\theta/2)}\,\Big| \;=\; 2\,\Big|\,{\sum_{k=0}^{n}}'\cos(k\,\theta)\,\Big| \;\le\; 2\,n+1,$$

sowie $$\Big|\,\frac{\sin[(n+1/2)\,\theta\,]}{\sin(\theta/2)}\,\Big| \;\le\; \frac{1}{\theta/\pi} \;=\; \frac{\pi}{\theta}.$$

Für irgendein $\mu \in\,]0,1[$ folgt somit

$$\|S_n\|_\infty \;\le\; \frac{1}{\pi}\,\Big(\int_0^\mu (2\,n+1)\;d\theta \;+\; \int_\mu^\pi \frac{\pi}{\theta}\;d\theta\Big) \;=\; \frac{(2\,n+1)\,\mu}{\pi} \;+\; \ln\frac{\pi}{\mu}.$$

Speziell für $\mu = \pi/(2\,n+1)$ ergibt sich die angegebene obere Schranke. $\square$

Um nun mittels (7.10) Konvergenz zu zeigen, benötigen wir die hinreichend schnelle Konvergenz von $E_n(f) \to 0 \quad (n \to \infty)$. Hierzu dienen die verschiedenen Sätze von Jackson, benannt nach Dunham Jackson (1888–1946), einem Schüler von Edmund Landau.

Satz (7.12) (Jackson I)

Für $f \in \mathrm{C}_{2\pi}^{(1)} := \mathrm{C}_{2\pi} \cap \mathrm{C}^1(\mathbb{R})$ und $n \in \mathbb{N}_0$ gilt

$$E_n(f) \;\leq\; \frac{\pi}{2(n+1)}\, \|f'\|_\infty.$$

Beweis: Zunächst zeigt man mittels partieller Integration die Darstellung

$$f(t) \;=\; \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\,d\theta \;+\; \frac{1}{2\pi}\int_{-\pi}^{\pi} \theta\, f'(\theta+t+\pi)\,d\theta. \tag{7.13}$$

Der erste Summand ist konstant. Da T_n die konstanten Funktionen enthält, genügt es, den zweiten Summanden durch trigonometrische Polynome aus T_n zu approximieren.

$$E_n(f) \;=\; \inf\,\{\, \|\frac{1}{2\pi}\int_{-\pi}^{\pi} \theta\, f'(\theta+t+\pi)\,d\theta \;-\; q(t)\|_\infty :\; q \in \mathrm{T}_n\}. \tag{7.14}$$

Nun ist für $g \in \mathrm{C}_{2\pi}$ und $p \in \mathrm{T}_n$ die folgende Funktion

$$q(t) \;:=\; \int_{-\pi}^{\pi} p(\theta)\, g(\theta+t)\,d\theta$$

auch stets wieder ein trigonometrisches Polynom in T_n. Aufgrund der Periodizität von p und g hat man nämlich

$$\begin{aligned} q(t) \;&=\; \int_{-\pi}^{\pi} p(\theta - t)\, g(\theta)\,d\theta \\ p(\theta - t) \;&=\; \frac{a_0(\theta)}{2} \;+\; \sum_{k=1}^{n} a_k(\theta)\,\cos(kt) \;+\; b_k(\theta)\,\sin(kt). \end{aligned}$$

Wir reduzieren also die Infimumsbildung in (7.14) auf die trigonometrischen Polynome q der obigen Form, wobei wir $g(t) := f'(t+\pi)$ wählen. Damit erhalten wir aus (7.14)

$$\begin{aligned} E_n(f) \;&\leq\; \inf\,\{\, \|\frac{1}{2\pi}\int_{-\pi}^{\pi} (\theta \;-\; p(\theta))\, f'(\theta+t+\pi)\,d\theta\,\|_\infty :\; p \in \mathrm{T}_n\} \\ &\leq\; \frac{1}{2\pi}\,\|f'\|_\infty\; \inf\,\{ \int_{-\pi}^{\pi} |\theta \;-\; p(\theta)|\,d\theta \;:\; p \in \mathrm{T}_n\}. \end{aligned} \tag{7.15}$$

Damit ist nun unser Problem, eine obere Schranke für $E_n(f)$ zu finden, zurückgeführt worden auf eine L_1-Approximationsaufgabe, nämlich auf das Problem, die Funktion $g(t) = t$ in L_1-Sinn auf dem Intervall $[-\pi, \pi]$ durch ein trigonometrisches Polynom $p \in T_n$ zu approximieren.

Wir verzichten darauf, die L_1-Optimalität der folgenden Konstruktion zu beweisen, und geben statt dessen nur die Lösung dieser Approximationsaufgabe an. Dazu bedarf es noch einiger Vorbemerkungen über trigonometrischer Interpolation.

(a) Zu $(2n+1)$ Knoten $t_0 < t_1 < \ldots < t_{2n} \in [-\pi, \pi[$ und Funktionswerten f_j gibt es genau ein trigonometrisches Polynom $p \in \mathrm{T}_n$ mit $p(t_j) = f_j, \; j = 0, \ldots, 2n.$
(Beweis: Transformation in ein komplexes Polynom.)

(b) Ein trigonometrisches Polynom $p \in \mathrm{T}_n, \; p \neq 0,$ kann in $[-\pi, \pi[$ höchstens $2n$ Nullstellen haben.
(Beweis: Gäbe es mehr als $2n$ Nullstellen, hätte man einen Widerspruch zu (a).)

(c) Wir setzen $t_k := \dfrac{k\pi}{n+1}, \, k = 1, \ldots, n.$ Dann gibt es genau ein trigonometrisches Polynom $p \in \mathrm{T}_n$ der Form $p(t) = \sum_1^n b_k \sin(kt)$ mit $p(t_k) = t_k, \; k = 1, \ldots, n.$
(Beweis: Wende (a) an auf die Stützstellen $(\pm t_k, \pm t_k)$ und $(0, 0)$ und zeige ebenfalls mittels (a), dass alle a_k verschwinden müssen.)

(d) Die Fehlerfunktion $e(t) := t - p(t)$ hat in $]0, \pi[$ genau die Nullstellen $t_k, \; k = 1, \ldots, n.$
(Beweis: e hat nach Konstruktion die Nullstellen t_k und 0. Hätte e noch eine weitere Nullstelle in $]0, \pi[$, so wäre $e'(t) = 1 - p'(t)$ ein *gerades* trigonometrisches Polynom in T_n und hätte nach dem Satz von Rolle wenigstens $(n+1)$ Nullstellen in $]0, \pi[$, also wenigstens $(2n+2)$ Nullstellen in $[-\pi, \pi[$; Widerspruch!)

Für die obige Konstruktion werten wir nun die L_1–Norm aus

$$\|e\|_1 \; = \; \int_{-\pi}^{\pi} |\theta \; - \; p(\theta)| \, d\theta \; = \; 2 \int_0^{\pi} |\theta \; - \; p(\theta)| \, d\theta.$$

Dazu setzen wir $\sigma_{n+1}(t) \; := \; \mathrm{sign}[\sin((n+1)t)]$ und finden

$$\begin{aligned}
\int_0^{\pi} |t \; - \; p(t)| \, dt \; &= \; \Big| \int_0^{\pi} \sigma_{n+1}(t) \, (t \; - \; p(t)) \, dt \, \Big| \\
&= \; \Big| \int_0^{\pi} \sigma_{n+1}(t) \, t \, dt \; - \; \sum_{k=1}^{n} b_k \int_0^{\pi} \sigma_{n+1}(t) \; \sin(k\,t) \, dt \, \Big| \\
&= \; \Big| \sum_{k=0}^{n} (-1)^k \int_{t_k}^{t_{k+1}} t \, dt \, \Big| \\
&= \; \frac{\pi^2}{2(n+1)} \, .
\end{aligned}$$

Zur vorletzten Gleichung zeigt man explizit, dass sämtliche Integrale $\int \sigma_{n+1}(t)\,\sin(k\,t)\,dt$ verschwinden. Die letzte Gleichheit ist ebenfalls durch explizite Berechnung der Integrale $\int t\,dt$ zu zeigen.

Insgesamt ist damit eine obere Schranke, nämlich $\|e\|_1 \le \pi^2/(n+1)$ bewiesen, wobei diese Schranke (ohne Beweis) bestmöglich ist. Setzt man diese nun in (7.15) ein, so ergibt sich schließlich die Behauptung des Satzes. □

Bemerkung (7.16)

Der in der Abschätzung (7.12) auftretende Faktor $\pi/(2n+2)$ lässt sich nicht verbessern.

Beweis: Zu $\varepsilon > 0$ lässt sich eine Funktion $f_\varepsilon \in \mathrm{C}_{2\pi}^{(1)}$ konstruieren mit den Eigenschaften

$$f_\varepsilon(k\pi/(n+1)) \;=\; (-1)^k,\; k = 0, \pm 1, \ldots, \pm(n+1), \quad \|f_\varepsilon'\|_\infty \;\le\; \frac{2(n+1)}{\pi}\,(1+\varepsilon).$$

Sei p_ε nun eine Bestapproximation von f_ε aus T_n bzgl. $\|\cdot\|_\infty$. Wäre $\|f_\varepsilon - p_\varepsilon\|_\infty < 1$, so hätte p_ε in den t_k das gleiche Vorzeichen wie f_ε und somit nach Zwischenwertsatz wenigstens $2n+2$ Nullstellen in $]-\pi,\pi[$. Widerspruch!
Damit folgt aber

$$E_n(f_\varepsilon) \;\ge\; 1 \;\ge\; \frac{\pi}{2\,(n+1)\,(1+\varepsilon)}\,\|f_\varepsilon'\|_\infty.$$

Für $\varepsilon \downarrow 0$ geht die Abschätzung gegen die des ersten Jackson-Satzes. □

Satz (7.17) (Jackson II)

Ist $f \in \mathrm{C}_{2\pi}$ Lipschitz-stetig mit einer Lipschitz-Konstanten L, so gilt für $n \in \mathbb{N}_0$

$$E_n(f) \;\le\; \frac{\pi}{2(n+1)}\,L.$$

Beweis: Für $\delta > 0$ setze man $F_\delta(t) \;:=\; \frac{1}{2\,\delta}\int\limits_{t-\delta}^{t+\delta} f(\theta)\,d\theta.$ Dann ist $F_\delta \in \mathrm{C}_{2\pi}^{(1)}$

und es gilt $F_\delta'(t) \;=\; \frac{1}{2\,\delta}\big(f(t+\delta) - f(t-\delta)\big)$, also $\|F_\delta'\|_\infty \le L.$

Ist $p \in \mathrm{T}_n$ nun Bestapproximation von F_δ aus T_n bzgl. $\|\cdot\|_\infty$, so folgt

$$E_n(f) \;\le\; \|f - p\,\|_\infty \;\le\; \|f - F_\delta\|_\infty + \|F_\delta - p\,\|_\infty \;\le\; \|f - F_\delta\|_\infty + \frac{\pi}{2(n+1)}\,L.$$

Ferner lässt sich abschätzen

$$\|f - F_\delta\|_\infty \;=\; \max_t\,\Big|\frac{1}{2\delta}\int\limits_{t-\delta}^{t+\delta}(f(t) - f(\theta))d\theta\Big| \;\le\; \max_t\,\frac{L}{2\delta}\int\limits_{t-\delta}^{t+\delta}|t-\theta|d\theta \;=\; \frac{L\delta}{2}$$

und somit $\|f - F_\delta\|_\infty \to 0 \quad (\delta \downarrow 0)$. □

Satz (7.18) (Jackson III)

Ist $f \in C_{2\pi}$, so gilt für $n \in \mathbb{N}_0$

$$E_n(f) \;\leq\; \frac{3}{2}\,\omega(\frac{\pi}{n+1}).$$

Dabei ist $\omega(\delta) := \sup\{|f(x) - f(y)| : \; |x - y| \leq \delta\}$ der Stetigkeitsmodul von f.

Beweis: Für $\delta > 0$ sei F_δ wie im Beweis zu (7.17) erklärt. Dann folgt mit $F_\delta'(t) = \big(f(t+\delta) - f(t-\delta)\big)/(2\,\delta)$ die Abschätzung

$$\|F_\delta'\|_\infty \;\leq\; \frac{\omega(2\,\delta)}{2\,\delta} \;\leq\; \frac{\omega(\delta)}{\delta}.$$

Die letzte Ungleichung ist eine direkte Folge aus der Definition des Stetigkeitsmoduls. Nach dem Jackson-Satz I gilt damit $E_n(F_\delta) \;\leq\; \dfrac{\pi}{2(n+1)\delta}\,\omega(\delta)$. Weiterhin ist

$$\begin{aligned}
\|f - F_\delta\|_\infty \;&=\; \max_t \frac{1}{2\,\delta}\,\Big|\int\limits_{t-\delta}^{t+\delta} \big(f(t) - f(\theta)\big)\,d\theta\Big| \\
&\leq\; \max_t \frac{1}{2\,\delta} \int\limits_{t-\delta}^{t+\delta} \omega(\delta)\,d\theta \;=\; \omega(\delta).
\end{aligned}$$

Insgesamt ergibt sich somit wieder

$$E_n(f) \;\leq\; \|f - F_\delta\|_\infty \;+\; E_n(F_\delta) \;\leq\; \big(1 + \frac{\pi}{2(n+1)\,\delta}\big)\,\omega(\delta).$$

Speziell für $\delta = \pi/(n+1)$ ergibt sich die Behauptung des Satzes. □

Bemerkung (7.19)

Da für eine stetige Funktion $f \in C_{2\pi}$ der Stetigkeitsmodul gegen Null konvergiert $\omega(\pi/(n+1)) \to 0$ $(n \to \infty)$ folgt aus dem dritten Jackson-Satz auch der zweite Weierstraßsche Approximationssatz zurück: Jede stetige, 2π-periodische Funktion lässt sich bzgl. der Maximumsnorm beliebig genau durch trigonometrische Polynome approximieren.

Kombiniert man nun diese Abschätzung für die Minimalabweichung $E_n(f)$ mit der Abschätzung (7.11b) für die Operatornorm $\|S_n\|_\infty$, so ergibt sich der folgende Satz von Dini und Lipschitz.

Satz (7.20) (Dini, Lipschitz)

Erfüllt $f \in C_{2\pi}$ die *Dini-Lipschitz-Bedingung*,

$$\lim_{\delta \downarrow 0}\,[\,\omega(\delta)\;\ln\delta\,] \;=\; 0,$$

so konvergiert die Fourier-Reihe $S_n(f)$ gleichmäßig gegen f für $n \to \infty$.

Beweis: Aufgrund des Lemmas von Lebesgue (7.10) sowie der Abschätzungen (7.11) und (7.18) ergibt sich

$$\begin{aligned}\|f - S_n(f)\|_\infty &\leq \big(1 + \|S_n\|_\infty\big)\, E_n(f)\\ &\leq \big(2 \,+\, \ln(2\,n+1)\big)\ \frac{3}{2}\,\omega(\frac{\pi}{n+1}).\end{aligned}$$

Für $n \geq 3$ stellt man fest, dass $\ln(2\,n+1) \leq \ln(2\,\pi) - \ln(\dfrac{\pi}{n+1})$ und $\ln(2\,\pi) < 2$ gelten.

Somit folgt

$$\|f - S_n(f)\|_\infty \;\leq\; 6\,\omega(\frac{\pi}{n+1}) \;-\; 1.5\,\ln(\frac{\pi}{n+1})\,\omega(\frac{\pi}{n+1}) \;\to\; 0 \quad (n \to \infty).$$ □

Bemerkung (7.21)

Über die Konvergenzgeschwindigkeit der Fourier-Entwicklung, wenn weitere Glattheitsvoraussetzungen an die Funkton f gestellt werden, gibt schließlich der vierte Jackson-Satz Auskunft, den wir nur kurz zitieren wollen

$$f \;\in\; \mathrm{C}_{2\,\pi}^{(k)} \;:=\; \mathrm{C}_{2\,\pi} \cap \mathrm{C}^k(\mathbb{R}) \;\Rightarrow\; E_n(f) \;\leq\; \big(\frac{\pi}{2n+2}\big)^k\; \|f^{(k)}\|_\infty.$$

Berechnung von Fourier–Koeffizienten.

Wir beginnen mit der **komplexen Darstellung** der Partialsumme

$$S_n(f) \;=\; \frac{a_0}{2} \;+\; \sum_{k=1}^{n}\,[a_k\cos(kt) + b_k\sin(kt)] \tag{7.22}$$

der Fourier-Entwicklung einer Funktion $f \in \mathrm{C}_{2\,\pi}$. Stellt man $\cos t$ und $\sin t$ als Real- und Imaginärteil von e^{it} dar, so ergibt sich

$$\begin{aligned}S_n(f) &= \frac{a_0}{2} + \sum_{k=1}^{n}\,[\,\frac{a_k}{2}\,(\mathrm{e}^{ikt} + \mathrm{e}^{-ikt}) \;+\; \frac{b_k}{2i}\,(\mathrm{e}^{ikt} - \mathrm{e}^{-ikt})\,]\\ &= \frac{a_0}{2} + \sum_{k=1}^{n}\,[\,\frac{a_k - i\,b_k}{2}\,\mathrm{e}^{ikt} \;+\; \frac{a_k + i\,b_k}{2}\,\mathrm{e}^{-ikt}\,] \qquad (7.23)\\ &= \sum_{k=-n}^{n} c_k\mathrm{e}^{ikt}\,.\end{aligned}$$

Umrechnung der Koeffizienten: $(k = 1, 2, \ldots, n)$

$$\begin{aligned} c_0 &= \frac{a_0}{2}, & c_k &= \frac{1}{2}(a_k - i\,b_k), & c_{-k} &= \frac{1}{2}(a_k + i\,b_k), \\ a_0 &= 2\,c_0, & a_k &= c_k + c_{-k}, & b_k &= i\,(c_k - c_{-k}). \end{aligned} \tag{7.24}$$

Fourier–Koeffizienten: $(k \in \mathbb{N}_0,\ j \in \mathbb{Z})$

$$\begin{aligned} a_k &= \frac{1}{\pi} \int_0^{2\pi} f(t)\ \cos(kt)\,dt, \\ b_k &= \frac{1}{\pi} \int_0^{2\pi} f(t)\ \sin(kt)\,dt, \\ c_j &= \frac{1}{2\pi} \int_0^{2\pi} f(t)\,\mathrm{e}^{-i\,j\,t}\,dt. \end{aligned} \tag{7.25}$$

Die Grundidee zur numerischen Berechnung der Fourier-Koeffizienten besteht darin, die Integrale in (7.25) durch Trapezsummen zu approximieren. Wegen der Periodizität der Integranden sind diese zur numerischen Quadratur besonders gut geeignet.

Für hinreichend große $N \in \mathbb{N}$ setzen wir

$$h := \frac{2\pi}{N}, \quad t_k := k\,h, \quad f_k := f(t_k), \quad k = 0, 1, \ldots, N. \tag{7.26}$$

Mit $f_0 = f_N$ (Periodizität) ergeben sich dann die folgenden Näherungen durch die Trapezsumme

$$\begin{aligned} a_k &\approx \frac{h}{\pi} \left\{ \frac{f_0}{2} + \sum_{j=1}^{N-1} f_j\ \cos(k\,t_j) + \frac{f_N}{2} \right\} \\ &= \frac{2}{N} \sum_{j=0}^{N-1} f_j\ \cos(j\,k\,h) \ =:\ A_k, \quad k = 0, \ldots, n, \\ b_k &\approx \frac{2}{N} \sum_{j=0}^{N-1} f_j\ \sin(j\,k\,h) \ =:\ B_k, \quad k = 1, \ldots, n, \\ S_n(f) &\approx \frac{A_0}{2} + \sum_{k=1}^{n} [\,A_k\ \cos(k\,t) + B_k\ \sin(k\,t)\,] \ =:\ \widetilde{S_n}(f)(t). \end{aligned} \tag{7.27}$$

Man beachte, dass die Approximationen $A_k \approx a_k$ und $B_k \approx b_k$ i. Allg. nur für kleine k brauchbar sind. So sind die $A_k,\ B_k$ gemäß Definition periodisch im Index, $A_{k+N} = A_k,\ B_{k+N} = B_k,$ während die tatsächlichen Fourier-Koeffizienten Nullfolgen bilden. Eine *Faustregel* besagt, dass man $N \geq 2\,n$ wählen sollte.

Andere Interpretationen:

a) Man kann die diskreten Fourier-Koeffizienten A_k, B_k (eigentlich $A_{k,N}$, $B_{k,N}$) auch als exakte Fourier-Koeffizienten einer *diskretisierten Approximationsaufgabe* interpretieren. Zu $N \geq 2n+1$ werde definiert

$$\begin{aligned} R &:= \mathrm{C}_{2\pi}(B), \qquad B := \left\{ \frac{2\pi}{N} k : \; k = 0, 1, \ldots, N-1 \right\}, \\ \langle f, g \rangle &:= \frac{2}{N} \sum_{k=0}^{N-1} f(k\,h)\, g(k\,h). \end{aligned} \tag{7.28}$$

Die Funktionen $\dfrac{1}{\sqrt{2}}$, $\cos t, \sin t, \ldots, \cos(n\,t), \sin(n\,t)$ bilden eine ONB des von ihnen aufgespannten (diskreten) Raumes $\widetilde{\mathrm{T}}_n$ bzgl. $\langle\cdot,\cdot\rangle$. Die A_k, B_k liefern dann die Lösung der L_2–Approximationsaufgabe, ein vorgebenes $f \in R$ durch ein Element aus $\widetilde{\mathrm{T}}_n$ zu approximieren. Genauer ist die Lösung der Approximationsaufgabe gegeben durch

$$\widetilde{S_n}(f)(t) = \frac{A_0}{2} + \sum_{k=1}^{n} [\, A_k \cos(k\,t) + B_k \sin(k\,t)\,],$$

wobei die A_k, B_k durch (7.27) gegeben sind.

b) Die A_k, B_k lassen sich auch als Lösung einer *trigonometrischen Interpolationaufgabe* interpretieren.

Zu Interpolationsdaten (t_k, f_k), $k = 0, 1, \ldots, N-1$, mit t_k aus (7.26), ist eine interpolierende trigonometrische Summe der Form

$$q(t) = \begin{cases} \dfrac{A_0}{2} + \displaystyle\sum_{k=1}^{n} [\, A_k \cos(k\,t) + B_k \sin(k\,t)\,], & \text{für } N = 2n+1, \\ \dfrac{A_0}{2} + \displaystyle\sum_{k=1}^{n-1} [\, A_k \cos(k\,t) + B_k \sin(k\,t)\,] + A_n \cos(n\,t), & \text{für } N = 2n. \end{cases} \tag{7.29}$$

Es lässt sich dann zeigen, dass diese Interpolationsaufgabe eine eindeutig bestimmte Lösung besitzt, welche gerade durch die A_k, B_k gemäß (7.27) gegeben ist.

Komplexe Variante: Man sucht ein (komplexes) trigonometrisches Polynom der Form

$$p(t) = \sum_{k=0}^{N-1} C_k \, \mathrm{e}^{ikt}, \tag{7.30}$$

das die gegebenen Daten (t_k, f_k), $k = 0, \ldots, N-1$, interpoliert.

Man erhält wieder eine eindeutig bestimmte Lösung, nämlich

$$C_k = \frac{1}{N} \sum_{j=0}^{N-1} f_j \, \mathrm{e}^{-i\,j\,k\,h}, \quad k = 0, 1, \ldots, N-1. \tag{7.31}$$

Auch hier lässt sich die reelle Lösung (7.27) und die komplexe Lösung (7.31) ineinander umrechnen. Man erhält analog zum kontinuierlichen Approximationsproblem, vgl. (7.22), $(k = 1, \dots, n)$

$$\begin{aligned} C_0 &= \frac{A_0}{2}, & C_k &= \frac{1}{2}(A_k - i\,B_k), & C_{N-k} &= \frac{1}{2}(A_k + i\,A_k), \\ A_0 &= 2\,C_0, & A_k &= C_k + C_{N-k}, & B_k &= i\,(C_k - C_{N-k}). \end{aligned} \tag{7.32}$$

Fazit: Sowohl für die numerische Berechnung der Fourier-Koeffizienten wie auch für die Auswertung der trigonometrischen Partialsummen sind jeweils trigonometrische Summen der Form (7.27) bzw. (7.31) auf einem festen Gitter - oder kontinuierlich (7.22) bzw. (7.23) auszuwerten. Dabei ist es gleichgültig, ob man die reelle oder die komplexe Darstellung verwendet, da beide inneinander umgerechnet werden können.

Der Algorithmus von Goertzel und Reinsch.

Der Algorithmus berechnet zu gegebenen Daten $f_0, \dots, f_{N-1} \in \mathbb{R}$ und $t \in \mathbb{R}$ die trigonometrischen Summen

$$A(t) := \sum_{k=0}^{N-1} f_k \cos(k\,t), \quad B(t) := \sum_{k=0}^{N-1} f_k \sin(k\,t). \tag{7.33}$$

Die Grundidee des Algorithmus ist – ganz analog zum Clenshaw Verfahren – die Verwendung einer Dreiterm-Rekursion für die Terme $c_k := \cos(k\,t)$ und $s_k := \sin(k\,t)$. Mittels trigonometrischer Umformung zeigt man

$$\begin{aligned} c_0 &= 1, & c_1 &= \cos t, & c_{k+1} &= 2\,c_1\,c_k - c_{k-1},\ k \geq 1, \\ s_0 &= 0, & s_1 &= \sin t, & s_{k+1} &= 2\,c_1\,s_k - s_{k-1},\ k \geq 1. \end{aligned} \tag{7.34}$$

Mit der gleichen Technik, die wir für den Algorithmus von Clenshaw angewendet haben, vgl. (6.30), erhalten wir den folgenden Algorithmus von Goertzel (1958):

Algorithmus von Goertzel (7.35)

$$U_N := 0, \quad U_{N-1} := f_{n-1},$$

$$c := \cos t, \quad F := 2\,c,$$

für $k = N-2,\ N-3,\ \dots,\ 1$

$$U_k := f_k + F\,U_{k+1} - U_{k+2}$$

$$A(t) := f_0 + c\,U_1 - U_2, \quad B(t) := U_1 \sin t.$$

Anmerkungen: Zu Berechnung aller A_k, B_k, $t = t_k = 2\pi k/N$, benötigt der Algorithmus $O(N^2)$ wesentliche Operationen.

Für kleine $|t|$ ist der Algorithmus numerisch instabil. Der Grund liegt darin, dass der Algorithmus mit $c = \cos t$ arbeitet. Kleine Änderungen in $c \approx 1$ entsprechen dabei großen Änderungen in t.

Stabilisierung nach Reinsch (1968)

Man ersetzt die Rekursion (7.34) durch eine solche, die mit $\sin^2(t/2)$ bzw. $\cos^2(t/2)$ anstelle von $\cos t$ arbeitet. Hintergrund sind die trigonometrischen Relationen

$$\sin^2(\frac{t}{2}) = \frac{1}{2}(1-\cos t), \qquad \cos^2(\frac{t}{2}) = \frac{1}{2}(1+\cos t).$$

Fall 1: $\cos t > 0$.

$$\begin{aligned} U_k &= f_k + 2c_1 U_{k+1} - U_{k+2} \\ &= f_k + 2(c_1-1)U_{k+1} + 2U_{k+1} - U_{k+2} \\ \Rightarrow \ (U_k - U_{k+1}) &= f_k + 2(c_1-1)U_{k+1} + (U_{k+1}-U_{k+2}). \end{aligned}$$

Mit $D_k := U_k - U_{k+1}$ gilt somit

$$D_k = f_k - 4\sin^2(t/2)\,U_{k+1} + D_{k+1}$$

und $A = f_0 + c_1 U_1 - U_2 = U_0 - c_1 U_1 = D_0 + 2\sin^2(t/2)\,U_1.$

Fall 2: $\cos t \le 0$.

$$\begin{aligned} U_k &= f_k + 2c_1 U_{k+1} - U_{k+2} \\ &= f_k + 2(c_1+1)U_{k+1} - 2U_{k+1} - U_{k+2} \\ \Rightarrow \ (U_k + U_{k+1}) &= f_k + 2(c_1+1)U_{k+1} - (U_{k+1}+U_{k+2}). \end{aligned}$$

Mit $D_k := U_k + U_{k+1}$ gilt somit

$$D_k = f_k + 4\cos^2(t/2)\,U_{k+1} - D_{k+1}$$

und $A = f_0 + C_1 U_1 - U_2 = U_0 - C_1 U_1 = D_0 - 2\cos^2(t/2)\,U_1.$

Algorithmus von Goertzel und Reinsch (7.36)

Falls $\cos t > 0$: $\sigma := 1, \quad F := -4\sin^2(t/2),$

sonst: $\sigma := -1, \quad F := 4\cos^2(t/2),$

$U_N := 0, \quad D_{N-1} := f_{N-1},$

$$\text{für } k = N-2,\ N-3,\ \dots,\ 0$$

$$U_{k+1} := D_{k+1} + \sigma\, U_{k+2},$$

$$D_k := f_k + F\, U_{k+1} + \sigma\, D_{k+1},$$

$$A(t) := D_0 - 0.5\, F\, U_1, \quad B(t) := U_1 \sin t.$$

In dieser Form ist der Algorithmus numerisch stabil und nur wenig aufwendiger als die ursprüngliche Version von Goertzel.

Die schnelle Fourier-Transformation (FFT).

Der Algorithmus von Goertzel und Reinsch wertet die trigonometrischen Summen an beliebigen Stellen $t \in \mathbb{R}$ aus. Er ist jedoch mit $O(N^2)$ Operationen immer noch recht aufwendig.

Ist man jedoch nur an den diskreten Fourier-Koeffizienten A_k, B_k interessiert, oder möchte man das trigonometrische Polynom nur auf dem Gitter $k\,h$, $h = 2\,\pi/N$ auswerten, so lassen sich effizientere Algorithmen angeben, die mit $O(N \log_2 N)$ Operationen auskommen. Solche Verfahren sind unter dem Namen FFT (fast Fourier transform) bekannt. Sie sind durch eine Arbeit von J.W. Cooley und J.W. Tuckey (Math. Comput. 19, 1965) populär geworden und bilden auch heute ein aktuelles Forschungsgebiet.

Wir gehen im Folgenden von der *komplexen Darstellung* aus und nehmen an, dass N eine Zweierpotenz ist. Der reelle Fall wird mittels der Rücktransformation (7.32) erledigt.

Gegeben: $N = 2^{r+1}, \quad r \in \mathbb{N}_0$

$$h = 2\,\pi/N, \quad t_k = k\,h, \quad f_k = f(t_k),\ k = 0, \dots, N-1,$$

Gesucht: $C_k = \dfrac{1}{N} \displaystyle\sum_{j=0}^{N-1} f_j\, \mathrm{e}^{-i\,j\,k\,h}, \quad k = 0, \dots, N-1.$

Die Grundidee besteht darin, die Summe in zwei Summanden aufzuteilen, die selber als Fourier-Koeffizienten zu einem gröberen Gitter mit doppelter Schrittweite interpretiert werden können.

Setzen wir $m := N/2$ und sortieren die Summe C_k nach geraden und ungeraden Indizes, so ergibt sich für $k = 0, \dots, N-1$

$$
\begin{aligned}
C_k &= \frac{1}{N}\left[\sum_{j=0}^{m-1} f_{2j}\,\mathrm{e}^{-i\,(2j)\,k\,h} \;+\; \sum_{j=0}^{m-1} f_{2j+1}\,\mathrm{e}^{-i\,(2j+1)\,k\,h}\right] \\
&= \left[\frac{1}{N}\sum_{j=0}^{m-1} f_{2j}\,\mathrm{e}^{-i\,j\,k\,(2h)}\right] \;+\; \mathrm{e}^{-i\,k\,\pi/m}\left[\frac{1}{N}\sum_{j=0}^{m-1} f_{2j+1}\,\mathrm{e}^{-i\,j\,k\,(2h)}\right] \\
&= G_k \;+\; \mathrm{e}^{-i\,k\,\pi/m}\;U_k, \qquad \text{wobei definiert wird}
\end{aligned}
$$

$$
G_k \;:=\; \frac{1}{N}\sum_{j=0}^{m-1} f_{2j}\,\mathrm{e}^{-i\,j\,k\,(2\pi/m)}, \qquad U_k \;:=\; \frac{1}{N}\sum_{j=0}^{m-1} f_{2j+1}\,\mathrm{e}^{-i\,j\,k\,(2\pi/m)}.
$$

Nun sind die G_k und U_k periodisch im Index:

$$
G_{k+m} \;=\; G_k, \quad U_{k+m} \;=\; U_k, \quad k \;=\; 0,\,1,\,\ldots\,,\,m-1,
$$

so dass sich der Rechenaufwand zur Berechnung der G_k und U_k gegenüber der Berechnung der C_k im Wesentlichen halbiert. Wir fassen den Reduktionsschritt noch einmal zusammen.

Reduktionsschritt (7.37) Mit $m \;:=\; N/2$ berechne man

$$
\begin{aligned}
&\text{für} \quad k \;=\; 0,\,1,\ldots,\,m-1 \\
&\qquad G_k \;:=\; \frac{1}{N}\sum_{j=0}^{m-1} f_{2j}\,\mathrm{e}^{-i\,j\,k\,(2\pi/m)}, \\
&\qquad U_k \;:=\; \frac{1}{N}\sum_{j=0}^{m-1} f_{2j+1}\,\mathrm{e}^{-i\,j\,k\,(2\pi/m)}, \\
&\qquad C_k \;:=\; G_k \;+\; \mathrm{e}^{-i\,k\,\pi/m}\;U_k, \\
&\qquad C_{k+m} \;:=\; G_k \;-\; \mathrm{e}^{-i\,k\,\pi/m}\;U_k. \\
&\text{end } k.
\end{aligned}
$$

Dieser Reduktionsschritt wird nun iteriert bis nur noch triviale Fourier-Transformationen (also $m \;=\; 1$) auszuführen sind. Die einzelnen Fourier-Transformationen für $m \;=\; 1,\,2,\,4,\ldots,2^r$ werden aus diesem gemäß (7.37) zusammengesetzt.

Der Algorithmus besteht aus zwei Phasen: Zunächst werden die Daten so umsortiert, wie sie nach Durchführung aller Reduktionsschritte auftreten. In der zweiten Phase werden diese Daten dann gemäß (7.37) wieder zusammengesetzt.

Phase 1: Sortieren der Daten:

Wir betrachten zunächst den Sortierschritt für $N = 8$:

Ausgangsdaten:	f_0	f_1	f_2	f_3	f_4	f_5	f_6	f_7
Schritt 1:	f_0	f_2	f_4	$f_6\|$	f_1	f_3	f_5	f_7
Schritt 2:	f_0	$f_4\|$	f_2	$f_6\|$	f_1	$f_5\|$	f_3	f_7
Schritt 3:	$f_0\|$	$f_4\|$	$f_2\|$	$f_6\|$	$f_1\|$	$f_5\|$	$f_3\|$	f_7

Betrachtet man die Dualdarstellung der Indizes bei den obigen Sortierschritten, so sieht man für den ersten Schritt

$$\begin{array}{cccc}
(2j) & |\ast \ldots \ast|0| & (2j+1) & |\ast \ldots \ast|1| \\
\downarrow & \searrow\searrow & \downarrow & \searrow\searrow \\
(j) & |0|\ast \ldots \ast| & (m+j) & |1|\ast \ldots \ast|
\end{array}$$

Für sämtliche r Sortierschritte erhält man die Index-Zuordnung

$$j \;=\; (j_r\, j_{r-1}\, \ldots\, j_1\, j_0)_2 \quad\mapsto\quad \overline{j} \;=\; (j_0\, j_1\, \ldots\, j_{r-1}\, j_r)_2,$$

d.h. der Index des f-Wertes in Anfangsposition j ergibt sich durch *Bitumkehr* der Ziffernfolge in der Dualdarstellung des Index j.

Für das Beispiel $N = 8$ ergibt sich

j	$(j)_2$	$(\overline{j})_2$	$\overline{j}$
0	(0 0 0)	(0 0 0)	0
1	(0 0 1)	(1 0 0)	4
2	(0 1 0)	(0 1 0)	2
3	(0 1 1)	(1 1 0)	6
4	(1 0 0)	(0 0 1)	1
5	(1 0 1)	(1 0 1)	5
6	(1 1 0)	(0 1 1)	3
7	(1 1 1)	(1 1 1)	7

Algorithmus FFT – 1.Teil (7.38)

```
C_0 = f_0/N;   j̄ = 0;
for  j = 1,2,...,N−1
     ℓ = N/2;
     while  ℓ + j̄ ≥ N,
          ℓ = ℓ/2;
     end
     j̄ = j̄ + 3ℓ − N;
     C_j̄ := f_j/N;
end j
```

Erläuterung: In Schritt j wird zum Index j der neues Index $\overline{j}$ berechnet. Dabei wird in der while–Schleife der alte Index $\overline{j}$ auf führende Einsen getestet

$$\begin{aligned}\overline{j} &= (1,\dots,1,0,j_{\ell+2},\dots,j_r)_2\\ \ell &= (0,\dots,0,1,0,\ \dots\ ,0)_2\\ \overline{j}_{\text{neu}} &= (0,\dots,0,1,j_{\ell+2},\dots,j_r)_2\\ \overline{j}+3\,\ell-n &= (0,\dots,0,1,j_{\ell+2},\dots,j_r)_2\end{aligned}$$

Phase 2: Nachdem die f_j–Werte nun in die richtige Reihenfolge gebracht worden sind, erfolgen die Reduktionsschritte (7.37)

Algorithmus FFT – 2.Teil (7.39)

$$\begin{array}{l}
\text{for}\quad \ell \;=\; 0,1,\dots,r\\
\qquad m \;=\; 2^{\ell};\quad \tilde{m} \;=\; 2\cdot m;\\
\qquad \text{for}\quad k \;=\; 0,\,1,\dots,m-1\\
\qquad\qquad c \;=\; \exp(-i\,k\,\pi/m);\\
\qquad\qquad \text{for}\quad j \;=\; 0 \,:\, \tilde{m} \,:\, N-\tilde{m}\\
\qquad\qquad\qquad G \;=\; C_{j+k};\\
\qquad\qquad\qquad U \;=\; c\cdot C_{j+k+m};\\
\qquad\qquad\qquad C_{j+k} \;=\; G\,+\,U;\\
\qquad\qquad\qquad C_{j+k+m} \;=\; G\,-\,U;\\
\qquad\qquad \text{end}\; j\\
\qquad \text{end}\; k\\
\text{end}\; \ell
\end{array}$$

Joseph Fourier (1768 – 1830)